In: Stoyanova & Kuyukchiev (eds.) Psychology and Linguistics. Papers in Honor of Prof. Encho Gerganov. Sofia: Prosveta, 2011: 221-239

COMPREHENSIVE MINIMAL DEPENDENCY APPROACH TO LEAN ANNOTATION OF MORPHOSYNTACTIC PHENOMENA

Tania Avgustinova, DFKI Language Technology Lab & Saarland University

Viewing linguistic modelling as annotation of a variety of phenomena becomes an increasingly popular perspective in natural language processing and grammar engineering, and it changes the way grammar modularity is understood nowadays. In this contribution we shall concentrate on a generalisation of the notion of dependency in combination with a rich typology of observable relationships holding between linguistically relevant items. The result is a comprehensive minimal dependency approach to lean annotation of morphosyntactic phenomena. Although the pre-theoretical ontology of relational types has been developed on the basis of Slavic morphosyntax, it provides a meta-annotation scheme which — by design — is meant to be compatible with theory-specific annotation schemes. Due to a phenomena-driven setting, the lean annotation approach represents one plausible strategy of introducing systematicity into the interpretation of linguistic data while remaining non-committal in theoretical disputes.

KEYWORDS: linguistic modelling, meta-annotation, systematic relations

1 INTRODUCTION

To achieve systematic interpretation of linguistic data, let us make a minimalist assumption about the nature of linguistic modelling and view it, essentially, as annotation of morphosyntactic phenomena. Such a perspective, being methodologically related to phenomena-oriented practice of test-suite building and corpus-annotation, can provide us with innovative insights into the way grammar modularity is to be understood. If we consider a constraint-based linguistic theory like the Head-driven Phrase Structure Grammar, HPSG (Pollard and Sag 1994; Sag, Wasow et al. 2001), we shall see that various grammatical properties of linguistic entities are typically revealed in complex taxonomies. Every linguistic entity can be considered from a variety of perspectives which are modelled as distinct dimensions of hierarchical classification. As a matter of fact, nothing in the formal apparatus of HPSG would actually exclude the possibility to organise also the relations holding in syntactic constructions in a type hierarchy. The type subsumption in such a multiple inheritance hierarchy is then interpretable as modelling a continuum from general – and presumably universal – systematic relations to more and still more specific instances of these relations resulting from admissible cross-classifications.

From a broader perspective, there is a growing movement within linguistics to promote the use of ontologies for linguistic description. Quite indicative in this context is the Surrey Morphology Group initiative on creating infrastructure for canonical typology in the form of a Community of Practice Extension (COPE) within the GOLD ontology for linguistics (Farrar and Langendoen 2003)[1]. However, differences in terminology and the underlying logic assumed are major stumbling blocks. One way of addressing these problems is to adopt the canonical approach to typology by taking defining properties and placing them in a multidimensional space. With regard to the motivation, (Corbett 2005) writes:

> This approach sidesteps two potential dangers in typology, namely 'premature statistics' and 'not comparing like with like'. The first danger is that something which is frequently found may be treated as uninteresting, whereas there are linguistic phenomena which are common yet which, I believe, should surprise us. The second danger is that we fail to take sufficient care over our terminology and so do not see that phenomena labelled identically are in fact distinct (conversely we miss identities because of different traditions of labeling).

Linguists' intuitions about what are particular instances of a phenomenon, such as a case or agreement, can differ because of differences in the choice of criteria which they take to be definitional. The canonical approach allows for addressing these differences by taking defining properties and placing them in a multidimensional space. In this way, one can treat, for example, issues of whether particular constructions fit under the rubric 'agreement' or 'case' as a matter of greater or lesser proximity to a canonical ideal. An ontology for this approach would therefore require a mapping out of the criteria that linguistic typologists use for defining linguistic constructs, and relevant issues would include canonical criteria for defining different morphosyntactic features – like case, gender, number, etc. – as well as specifying canonical criteria for syntax-morphology interaction – e.g. in classical terms of agreement, government, head, modifier, etc. [2]

The generalized dependency theory adopted here can be seen as an extension of classical dependency grammar approaches with additional elements of linguistic and cognitive sophistication such as a typed feature system, multiple dimensions of cross-classification and the consequent and consistent use of the multiple-inheritance type hierarchy as the sole basis for all encoded linguistic knowledge. The notion of linguistic construction is commonly used to refer to the syntactic arrangement of patterning within a grammatical unit. This traditional view has been refined by (Avgustinova and Uszkoreit 2000; Avgustinova and Uszkoreit 2003; Avgustinova and Uszkoreit 2006), where arrays of systematic relations motivate shared patterns of variation cross-linguistically as well as across constructions. A key to formalisation is the observation that syntagmatic regularities in morphosyntax reveal essential relations between properties of linguistic objects (Figure 1). In fact, all grammatical representations have to eventually identify linguistic items of different motivation and complexity, encode

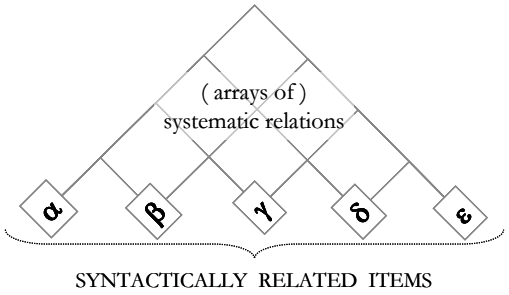
---

[1] See also: www.linguistics-ontology.org/gold.html

[2] Cf. the workshop site at: http://www.ias.surrey.ac.uk/workshops/typology/

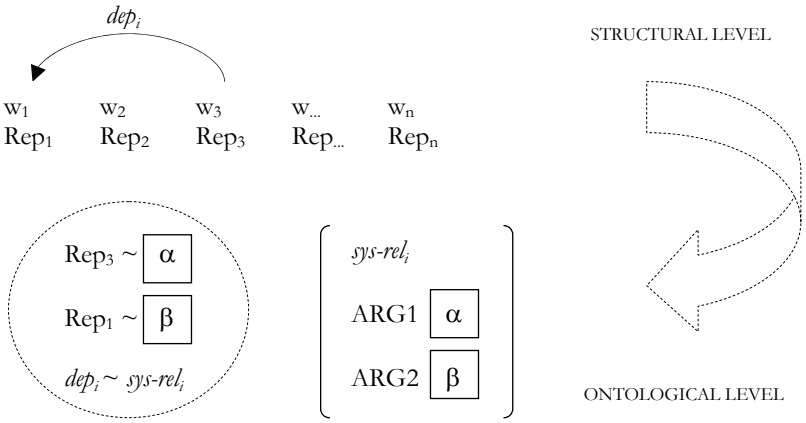
properties of linguistic items, and specify explicit or implicit relationships between properties of linguistic items.

Figure 1: Domain of interest



A comprehensive minimalist notion of dependency would refer to a systematic relation observable in morphosyntax, with no explicit directionality. Yet formally we need a convention for identifying (or referring to) a given minimal dependency or systematic relation. For that reason, each binary relation should be characterised by two attributes (ARG1 and ARG2) and by its type (*sys-rel*), as summarised in (Figure 2).

Figure 2: A generalised notion of dependency

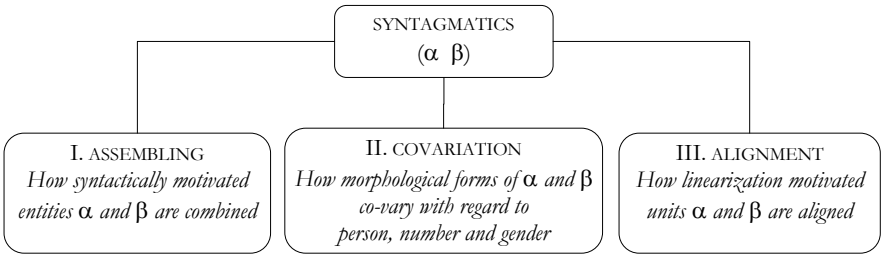


Let $\{w_1, w_2, \ldots, w_n\}$refer to concrete words in a string and $\{Rep_1, Rep_2, \ldots, Rep_n\}$ to the respective linguistic representations of these words, and a relationship dep_i can be identified between the words w_1 and w_3, such that the former depends on the latter. This structural level analysis is generalized at an ontological level as a systematic relation of type $sys\text{-}rel_i$ (corresponding to dep_i) associated with a feature structure with two attributes ARG1 and ARG2 whose respective values α and β (corresponding to Rep_3

and Rep_1) are identified with the linguistic items participating in the relationship. Eventually, we obtain a well-formed typed feature structure describing a grammatical phenomenon.

Let us further assume — quite in the spirit of (Avgustinova 2000; Avgustinova 2003; Avgustinova and Uszkoreit 2003; Avgustinova 2004; Avgustinova and Uszkoreit 2006; Avgustinova 2007; Avgustinova 2007) — that cross-linguistically observable SYNTAGMATICS, pertaining to a given pair of linguistic entities and essentially accounting for morphosyntax and word order, is manifested along tree dimensions (Figure 3), namely: ASSEMBLING (i.e. how two syntactically motivated entities are combined), COVARIATION (i.e. how two morphological forms co-vary in person number, gender), as well as ALIGNMENT (i.e. how two linearization motivated units are aligned).

Figure 3: Observable syntagmatics



Attempts to derive the range of possible syntactic combinations of two linguistic items from a multidimensional inheritance network of basic relational types open new ways of classifying relations that are non-standard or marginal in most theories. A number of these relations fall out from the proposed cross-classification in natural and theoretically rewarding ways. Our comprehensive minimal dependency approach to lean annotation of morphosyntactic phenomena straightforwardly matches the canonical methodological setting outlined in (Corbett 2005):

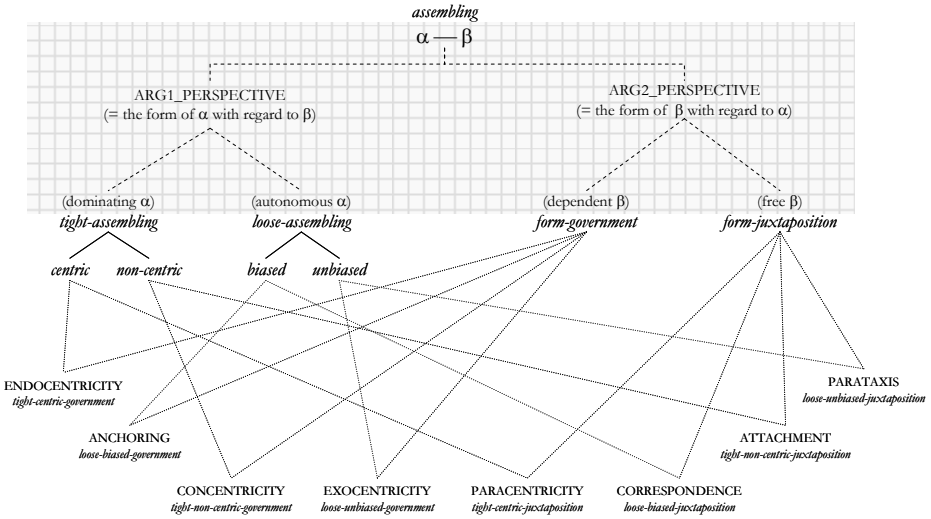
> In a canonical approach, we take definitions to their logical end point and build theoretical spaces of possibilities. Only then do we ask how this space is populated. … It follows that canonical instances (the best examples, those most closely matching the canon) may well not be the most frequent. They may indeed be extremely rare, or even non-existent. However, they fix a point from which occurring phenomena can be calibrated, and it is then significant and interesting to investigate frequency distributions.

The classification presented below has been originally designed to systematise the inventory of syntactic relationships found across Slavic languages. Yet, a far-reaching outcome of this study is to promote the view that systematic relations holding between the components of syntactic constructions must be treated as research objects in their own right.

2 CLASSIFICATION OF ASSEMBLING RELATIONS

An assembling relation which holds between an element α and an element β can always be viewed from two basic perspectives, namely, with regard to ARG1, i.e. taking into primary consideration the formal properties of α as they are exhibited in the relation, or with regard to ARG2 and respectively focusing on β, as sketched in (Figure 4). Considering, on the one hand, the form of α with regard to β, we shall differentiate *tight-assembling* when the form of α is dominating in the relation and *loose-assembling* when it is autonomous in this respect, i.e. not explicitly dominating in the relation. It appears typical of α in tight assembling to be dominating in form and yet either syntagmatically prior (in *centric* assembling) or neutral (in *non-centric* assembling). In contrast, the form of α in loose assembling is autonomous, being again either syntagmatically prior (resulting in *biased* assembling) or neutral (resulting in *unbiased* assembling). Considering, on the other hand, the form of β with regard to α, let us make explicit that in a given relation it is either dependent (i.e. determined by the other element) and indicates *government*, or free and indicates *juxtaposition*. Eventually, as a result of admissible cross-classifications of relation types, we obtain eight groups of related phenomena (Figure 5).

Figure 4: The assembling dimension



The ENDOCENTRICITY class, interpreted as tight centric government, includes phenomena of "strict selection" like the traditional subcategorisation, for instance. For representative examples cf. (Figure 6). The CONCENTRICITY class, interpreted as tight non-centric government, includes phenomena viewed traditionally as "mutual selection" like in the case of modification. For representative examples cf. (Figure 7). The PARACENTRICITY class, interpreted as tight centric juxtaposition, includes "head-functor" phenomena like the traditional marking, for instance. For representative

examples cf. (Figure 8). The ATTACHMENT class, interpreted as tight non-centric juxtaposition, includes phenomena like the traditional adjunction, for instance. For representative examples cf. (Figure 9).

Figure 5: Distinguished classes of assembling phenomena

		Tight Assembling		Loose Assembling	
	α →	centric	non-centric	biased	unbiased
Form Government	β → dependent	ENDOCENTRICITY "strict selection" subcategorisation relational_case cross-referencing object_cliticisation agglomerate	CONCENTRICITY "mutual selection" modification concordial_case ascriptive_predication - attributive - classificational	ANCHORING control identificational_predication - equative - specificational	EXOCENTRICITY co-predication co-dependence coupling - existential - localisation - possessive
Form Juxtaposition	free	PARACENTRICITY "head – functor" marking morphosyntactic_marking inflectional_marking	ATTACHMENT adjunction case_adjunction secondary_predication predicative_case_adjunction	CORRESPONDENCE co-marking resumption relativisation	PARATAXIS coordination composite detachment - isolating - parenthetical - expository

The ANCHORING class, interpreted as loose biased government, includes phenomena corresponding to the traditional notion of control, for instance. For representative examples cf. (Figure 10). The EXOCENTRICITY class, interpreted as loose unbiased government, includes phenomena like co-predication or co-dependence, for instance, which are traditionally not explicitly distinguished. For representative examples cf. (Figure 11). The CORRESPONDENCE class, interpreted as loose biased juxtaposition, includes phenomena like co-marking or resumption, for instance, which are traditionally not explicitly distinguished. For representative examples cf. (Figure 12). Eventually, the PARATAXIS class, interpreted as loose unbiased juxtaposition, includes phenomena like the traditional coordination, for instance. For representative examples cf. (Figure 13).

Figure 6: Tight assembling: endocentric phenomena (strict selection)

sys-rel (α, β)	α *dominating* (functor ≡ head)	β *dependent*
subcategorisation	predicate	complement
relational-case	verb	referential object/subject
object-cliticisation	verb	object clitic
	noun	possessive clitic
cross-referencing	predicative pronominal clitic	replicated referential material

Figure 7: Tight assembling: concentric phenomena

sys-rel (α, β)	α *non-centric*	β *dependent*
modification	X	(form-compatible) adjunct
concordial-case	noun	adjective
ascriptive-predication • *attributive-predication* • *classificational-predication*	referential subject	ascriptive predicative • attributive adjective • classificational adjective

Figure 8: Tight assembling: paracentric phenomena

sys-rel (α, β)	α *centric* (head)	β *free* (functor)
marking	major category	minor (functional) category
morphosyntactic-marking	verb	auxiliary
inflectional-marking	predicative	copula

Figure 9: Tight assembling: attachment phenomena

sys-rel (α, β)	α *non-centric*	β *free*
adjunction • *case-adjunction* • *secondary-predication* • *predicative-case-adjunction*	X	(free) adjunct • case-predetermined-adjunct • predicative adjunct • case-predetermined-predicative-adjunct

Figure 10: Loose assembling: anchoring phenomena (strict selection)

sys-rel (α, β)	α *biased* (triggering expression)	β *dependent*
control	controller	controlled predicate
identificational-predication	nominal subject	nominal predicative
equative-predication	X	Y being set equal to X
specificational-predication		Y further specifying X

Figure 11: Loose assembling: exocentric phenomena

sys-rel (α, β)	α *unbiased*	β *dependent*
co-predication	primary predicate	secondary predicate
co-dependence	referential complement of X	referential complement of X
coupling	autonomous expression	autonomous expression
existential-coupling	existing entity	"location"
localisation-coupling	"location"	existing entity
possessive-coupling	entity	"possessor (as location)"

Figure 12: Loose assembling: correspondence phenomena

sys-rel (α, β)	α *biased*	β *free*
co-marking	marker of X	marker of X
resumption	expression	resuming pronominal form
relativisation	nominal expression	relative pronoun

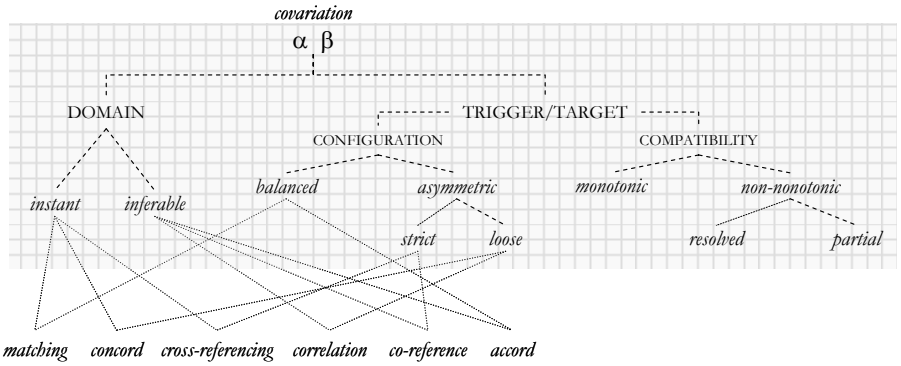
Figure 13: Loose assembling: paratactic phenomena

sys-rel (α, β)	α *unbiased*	β *free*
coordination	coordinand	coordinand
composite	part of compositum	part of compositum
detachment • *isolating* • *parenthetical* • *expository*	embedding expression	detached expression • isolated expression • parenthesized expression • expository expression

3 CLASSIFICATION OF COVARIATION RELATIONS

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. Most investigations typically concentrate on the linguistic items themselves (as agreement sources) and on the relevant properties of these items (in terms of agreement features and conditions), while it is the relational nature of agreement that may serve as the basis for a phenomena-driven modularisation. Co-variation affects the so-called 'phi-features' (prototypically: person, number, gender) in distinct morphosyntactic settings. In a multidimensional taxonomy, (Avgustinova and Uszkoreit 2006) derive the space of possible agreement relations from a small number of distinctions suitable for immediate formalization.

Figure 4: The co-variation dimension



In particular, the descriptive device of relational dependency is utilized to provide a formal framework for encoding these relationships in such a way that the descriptions can be linked to constraint-based grammar formalisms. The typology of agreement phenomena results from admissible cross-classifications in a multiple-inheritance hierarchy. As already mentioned above, the most general type *sys(tematic)-rel(ation)* is associated with two attributes ARG1 and ARG2 with values of type *sign*, borrowing terminology from HPSG, so that is would be possible to define a number of relationships among signs – cf. (Figure 4) and (Figure 5).

Figure 5: Types, features and constraints

Types	Associated features	Relational constraints
sys-rel	ARG1 *sign* [ ] ARG2 *sign* [ ]	
covariation	ARG1 [1] [… [3]] ARG1 [2] [… [4]]	covar-sources ([3], [4])
instant		local ([1], [2])
inferable		non-local ([1], [2])
balanced		x-center ([1], [2])
asymmetric		center ([1]) ∨ center ([2])
loose		unstipulated ([3], [4])
strict		unidirectional ([3], [4])

The *covariation* type, in particular, involves distinct linguistic entities ([1] corresponding to α, and [2] corresponding to β) with dedicated features ([3] and [4]) that are identified as *covar(iation)-sources* by an associated two-place predicate. The minimal distinction needed for adequate classification of Slavic agreement is the DOMAIN of co-variation, which can be either instant or inferable, and the TRIGER / TARGET configuration, which can be either balanced or asymmetric. The *instant* sub-type is associated with a two-place predicate *local*, while the *inferable* sub-type with a two-place predicate *non-local*. The *balanced* sub-type involves a two-place predicate *x-center* which establishes the situation where neither of the items can be identified as central to the co-variation, in contrast to the *asymmetric* sub-type that is associated with a disjunctive one-place predicate identifying that one of the related items plays a prominent role in the trigger-target configuration. In the latter case a further refinement of the type hierarchy is needed with regard to the dedicated features of the involved items in order to distinguish *loose* from *strict* asymmetry by means of the two-place predicates *unstipulated* and *unidirectional* respectively. Eventually, we are in a position to define six classes of co-variation phenomena:

- *Matching* as instant balanced co-variation is found between the auxiliaries and the main verb in periphrastic forms. As discussed in (Avgustinova 1997), the person-number-gender information in Bulgarian analytic (periphrastic) verb forms can be distributed among several components, namely, the main verb itself and a set of auxiliaries functioning as markers to it.

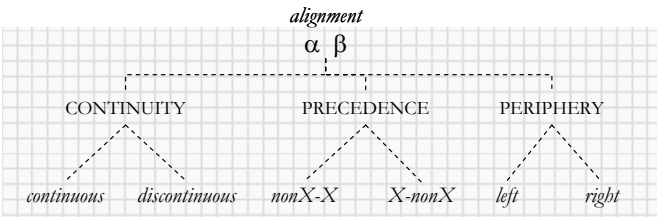
- *Concord* as instant loosely asymmetric o-variation is prototypically found within noun phrases, between the adjective and the noun, or possibly between adjectives that modify the same noun.
- *Cross-referencing* as instant strictly asymmetric co-variation holds between the verb and its subject or complement; the same type of co-variation can be assumed between the verbal clitic pronoun cliticized on the verb and the nominal object cross-referenced by this clitic.
- *Correlation* as inferable loosely asymmetric co-variation is typically observed in relative clause constructions between the relative pronoun and the noun modified by the relative clause.
- *Co-reference* as inferable strictly asymmetric co-variation holds between an object (or a verbal clitic cross-referencing this object) and the predicative adjective controlled by it; or more generally, between a referential expression, on the one hand, and a co-referent pronoun or a secondary predicative, on the other hand.
- *Accord* as inferable balanced co-variation holds between the subject and the complement which are co-dependents of the same verb.

An important aspect of covariation phenomena concerns the compatibility of the involved features, which is modelled by an additional dimension in (Figure 4). Agreement between α and β can be *monotonic* or *non-monotonic*, with the latter type further being *resolved* or *partial*. Eventually, all six phenomena specified up to now — i.e. with regard to DOMAN and TRIGER/TARGET configuration — can further be cross-classified along this TRIGER/TARGET compatibility dimension, resulting in even more fine-grained but yet modular system.

4 CLASSIFICATION OF ALIGNMENT RELATIONS

Alignment relations specify how two linearization-motivated units (α and β) are related to each other. At least the following minimal distinctions are needed to capture important linearization aspects (Figure 6) like CONTINUITY – *continuous* or *discontinuous*, PRECEDENCE – whether α precedes β (e.g. *nonX-X*) or whether β precedes α (e.g. *X-nonX*) and PERIPHERY – *left* or *right*.

Figure 6: The alignment dimension



Partial agreement with coordination is an interesting and non-trivial example of a complex phenomenon which allows us to illustrate how the dimensions of alignment and co-variation interact to provide straightforward modular interpretations. As (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.

This is reflected in (Figure 4) by introducing a COMPATIBILITY dimension distinguishing *monotonic* from *non-monotonic* co-variation. 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. The interested reader may consult (Corbett 1998; Corbett 2000) 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. Strategy B *(partial)* means that the one of the conjuncts is favored as decisive in establishing co-variation, mainly on alignment grounds – cf. (Corbett 1998), p. 9:

> When agreement is with one conjunct it is almost always with the nearest. (...some languages, exceptionally, allow agreement with the first conjunct when it is not the nearest.)

In particular, co-variation with the nearest conjunct, which is also the first, can be seen in (i) and (ii). A situation where the nearest and first are distinct is illustrated by the Cassubian sentence in (iii). The Slovene sentence in (iv) illustrates a rare possibility called by Corbett distant agreement, i.e. agreement with the first conjunct, which, with subject-verb word order, is not the nearest; co-variation with all (i.e. Strategy A) would require in the latter case the masculine plural.

(i)

Èta	*vzyskatel'nost',*	*samokritičnost'*	*tožę*	*raspolagali*	*k nemu.*
this.SG	exactingness.SG,	self-criticalness.SG	also	disposed.PL	to him

This exactingness and self-criticalness also disposed me favourably towards him.

(ii)

Byla	*u nego*	*ešče*	*gitara*	*i*	*samoučitel'*	*k*	*nej.*
was.SG.F	at him	also	guitar.SG.F	and	manual.SG.M	for	it

He also had a guitar and a manual for it.

(iii)

Odraza	*i*	*strach*	*czierowôl*	*jego*	*postępkama.*
refulsion.SG.F	and	fear.SG.M	directed.SG.M	his	actions

Revulsion and fear directed his actions.

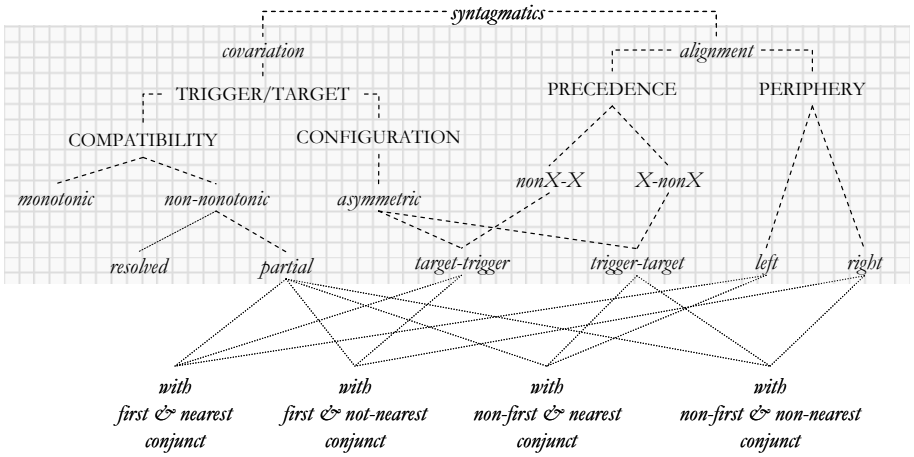
(iv)

Knjige	*in*	*peresa*	*so*	*se*	*podražile.*
book.PL.F	and	pen.PL.N	AUX.PL	RFL	got-dear.PL.F

Books and pens have become more expensive.

Such Strategy B realizations fall out as a result of admissible cross-classifications in (Figure 7). 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.

Figure 7: Partial co-variation with conjoined noun phrases



While the first three variants of Strategy B are fairly common across languages, the fourth variant is also attested – in particular, (Corbett 1998; Corbett 2000) mentions some interesting although limited evidence that in Serbo-Croatian Čakavian dialects of 16th-17th centuries agreement has been attested "with the most important conjunct, even if this was not the nearest or the first".

5 OUTLOOK

An approach to linguistic modelling as (meta)annotation of linguistic phenomena becomes essential at the interface between grammar engineering and grammar induction as well as for comparison and evaluation of linguistic resources. The inventory of surface syntactic relations employed in the Meaning-Text Theory (Mel'cuk 1974; Mel'cuk 1995), for instance, reveals a significant potential of informing the linguistic research in this direction. We have concentrated on a generalization of the notion of dependency in combination with a rich typology of systematic relations. The research strategy involves a pre-theoretical ontology of relational types that has been developed on the basis of Slavic morphosyntax. The overall perspective advocated here is vital for the development of grammar and treebank evaluation frameworks by enabling cross-linguistic and cross-formalism generalisations and comparisons.

By focusing on the relational aspect, the outlined approach allows us to specify more precisely the nature of the observable grammatical phenomena as well as to properly sub-classify them. The space of possible relationships is derived from a small number of distinctions, employing the power of multidimensional inheritance networks for a systematic and concise description. The resulting ontology of systematic relations is open enough to accommodate typologically diverse phenomena. This eventually leads to thorough "meta-annotation" of morphosyntactic phenomena based on minimal dependencies, which is by design compatible with theory-specific annotation schemes. As a consequence, subtasks in grammatical research can be defined more cleanly, which can ultimately provide linguistically informed modularity in technological applications.

REFERENCES

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

Avgustinova, T. (2000). Arguments, grammatical relations, and diathetic paradigm. Berkeley Formal Grammar, 7th International Conference on HPSG. Berkeley, University of California**:** 85-89.

Avgustinova, T. (2003). Metagrammar of systematic relations: a study with special reference to Slavic morphosyntax. Syntactic Structures and Morphological Information. U. Junghanns and L. Szucsich. Berlin / New York, Mouton de Gruyter**:** 1-24.

Avgustinova, T. (2004). Shared Grammatical Resources for Slavic Languages. Selected Topics in Multilingual Grammar Design with Special Reverence to Slavic Morphosyntax. Saarbruecken, University of Saarland.

Avgustinova, T. (2007). Generalized Dependency Theory. 19th European Systemic Functional Linguistics Conference and Workshop: Data interpretation in linguisitc analysis, Saarbrücken.

Avgustinova, T. (2007). Language Family Oriented Perspective in Multilingual Grammar Design. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien, Peter Lang - Europäischer Verlag der Wissenschaft.

Avgustinova, T. and H. Uszkoreit (2000). An ontology of systematic relations for a shared grammar of Slavic. 18th International Conference on Computational Linguistics COLING'2000. Saarbrücken. **1:** 28-34.

Avgustinova, T. and H. Uszkoreit (2003). Towards a typology of agreement phenomena. The Role of Agreement in Natural Language: TLS 5 Proceedings. W. E. Griffin. Austin, TX, Texas Linguistics Forum. **53**.

Avgustinova, T. and H. Uszkoreit (2006). Shared and Non-Shared Grammar in Modelling Slavic Morphosyntax. Readings in Multilinguality. Selected papers for young researchers in BIS-21++. G. Angelova, K. Simov and M. Slavcheva. Shoumen, Bulgaria, Incoma Ltd.

Corbett, G. G. (1998). Agreement in Slavic (position paper). Workshop on Comparative Slavic Morphosyntax, Indiana University.

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

Corbett, G. G. (2005). The canonical approach in typology. Linguistic Diversity and Language Theories. Z. Frajzyngier, A. Hodges and D. S. Rood. Amsterdam, John Benjamins**:** 25-49.

Farrar, S. and T. Langendoen (2003). "A linguistic ontology for the Semantic Web." GLOT International **7**(3): 97-100.

Mel'cuk, I. A. (1974). Opyt teorii lingvisticeskix modelej "Smysl <==> Tekst". Moskva, Nauka.

Mel'cuk, I. A. (1995). The Russian Language in the Meaning-Text Perspective. Moskau - Wien, Gesellschaft zur Förderung slawistischer Studien.

Pollard, C. and I. Sag (1994). Head-Driven Phrase Structure Grammar. Chicago, University of Chicago Press.

Sag, I. A., T. Wasow and E. M. Bender (2001). Syntactic Theory: A Formal Introduction. Stanford, CSLI.