

Prototypical Predicate Constructions: Evidence from Slavic

Patrick M. Murphy

University of North Carolina at Chapel Hill

<patrick_murphy@unc.edu>

36th Annual Meeting of the Southern Conference for Slavic Studies

March 20, 1998

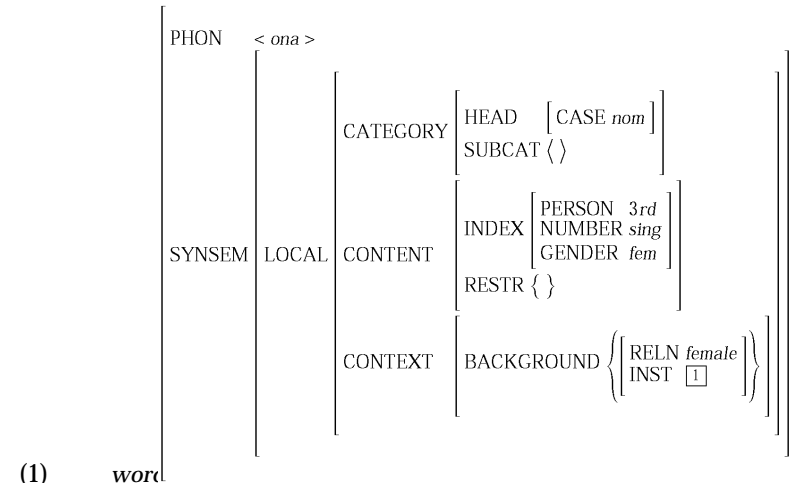
Mechanisms employed by generative grammar to capture the systematic relationship between linguistic constructions:

- Transformations – both the active and passive surface structure are derived from the same thematically relevant deep structure (D-S *rabóće strojat dom* S-S *domstrojatsjarabóám*); relates tree structures
- Lexical rules – passive verbs are derived from active verbs by means of a rule which affects their argument structure; relate one head to another
- Type theory/inheritance – capture relations through type partitioning

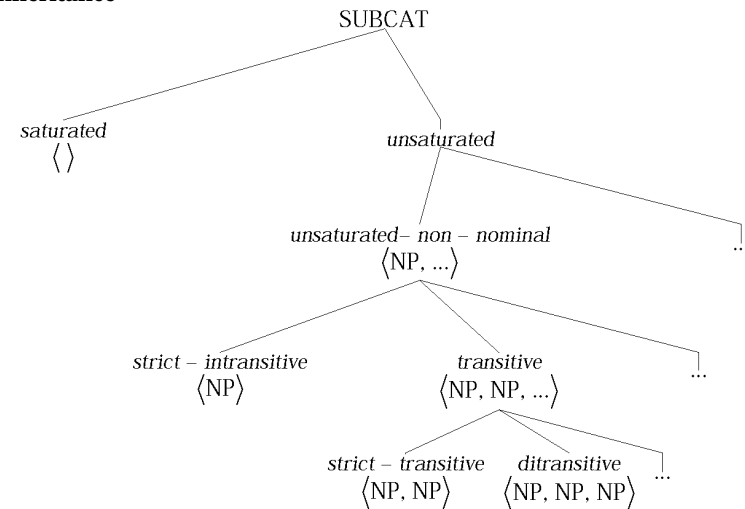
Abney's (1996b) arguments for weighted/stochastic models of grammar:

- Known facts of language acquisition: during language acquisition children do not abruptly change their grammars in the fashion suggested by parameters
- Diachrony: language change is also not as non-continuous as parameters would predict
- Dialect continua and the *Sprachbund* the slight changes that occur with distance is more easily modeled in terms of preference if we have a speech community with grammatical variance
- Non-universality of universals: even famed language universals (double passivization, Complex NP Constraint) are rarely truly universal, but rather it seems violating constructions are simply so unlikely or marked that we predict they appear with almost no frequency

Mechanisms of Head-Driven Phrase Structure Grammar:



(1) Type inheritance



Givón, Myhill and Topicality

One area that strongly suggests evaluation in terms of a weighted grammar is that of word order variation. A large amount of evidence seems to suggest that the topicality of subject and object seriously affects word order.

(3) Referential Distance (RD)

This measurement assesses the gap between the previous occurrence in the discourse of a referent/topic and its current occurrence in a clause, where it is marked by a particular grammatical coding device. The gap is thus expressed in terms of number of clauses to

the left. The minimal value that can be assigned is thus 1 clause, which is maximally continuous.

(4) Topic Persistence (TP)

We measure persistence in terms of the number of clauses to the right — i.e. in subsequent discourse from the measured clause — in which the topic/participant continues as an uninterrupted presence as a semantic argument of the clause, an argument of whatever role and marked by whatever grammatical means. The minimal value that can be assigned is thus zero, signifying an argument that decays immediately, i.e. of the lowest persistence.

(Givón 1983a: 15)

Givón's (1983a) study of Ute (59% SV and 78% OV) demonstrated that preverbal NPs have high RDs, whereas low RDs were associated with postverbal NPs. Topic Persistence also correlates to word order, but is different for subjects and objects: subjects have high TP if preverbal and low TP if postverbal, while objects were just the opposite. Thus Myhill's generalizations about Ute NPs:

(5) Function of preverbal and postverbal position in Ute

Preverbal subject — switch attention to a character not on the scene before who will become the center of attention

Postverbal subject — keep attention on a character who will not be continued for long as the center of attention

Preverbal object — refer to an entity not mentioned in the preceding or following discourse

Postverbal object — refer to an entity mentioned in the preceding and following discourse

(Myhill 1992: 171)

After looking at several languages with different basic word orders, Myhill suggested the following universal pragmatic word order pattern: marked or contrastive topics tend to appear in initial position. Estival and Myhill (1988) examined the role of topicality and choice of word order in English. While subjects are more topical than objects, passives and other inverted word orders are used when the logical subject is lower in topicality and the logical object is higher in topicality (they cite Svartvik's (1966) text count study which found inanimate logical subjects in 81% of English passives but only in 27% of actives). The following table shows some of the features affecting topicality:

(6)	Active	Passive	% Passive
Nominal A	350	18	5%
Pronominal A	1760	1	0%
Non-human A	100	8	7%
Human A	2010	11	1%
Indefinite A	180	8	4%
Definite A	1930	11	1%

(Here A is the subject of a transitive verb or the corresponding oblique of the passive, the "logical subject")

Siewierska's (1993) data

Siewierska's (1993) study of Polish word order in terms of both 'syntactic weight' (length and complexity of syntactic constituents) and information structure (measurement of topicality of subject and direct object) in transitive sentences found that while word order was fairly flexible for subjects and objects of "normal" topicality (around 1.00 for subjects, 0.63 for objects), unusual topicality is marked by two word orders: SOV and VSO (only 16.2% of the six word orders).

Word Order	Subject TP	Object TP
SOV	1.69	0.65
SVO	1.02	0.68
VSO	0.56	1.11
VOS	0.94	0.50
OVS	0.88	0.49
OSV	0.92	0.37

(Siewierska 1993: 257)

- SOV — topicality of the subject argument is considerably higher than average (1.69)
- VSO — topicality of the subject considerably below its average (0.56) and topicality of the object considerably above its own norm (1.11)

Modeling Prototypical Predicates

Topicality is taken to be a function of those features which seem to be indicative of topicality (person, number, humanness, animacy, definiteness), and represented here as being between 0.65 and 0.85 on a scale from 0 to 1, where 0 represents total lack of topicality and 1 represents a most topical sign.

Definitions of 'neutral' topicality for an agent or patient argument:

- neutral-top-agent ([] index) :-
topicality ([] , 0.65, 0.85).
- increased-top-agent (1 index) :-
topicality ([] , 0.85, 0.90).
- neutral-top-patient (1 index) :-
topicality ([] , 0.15, 0.35).
- etc.

Neutral topicality (*neutral-top-agent*) for an agent is true if the topicality of the agent is between 0.65 and 0.85. It is of type *increased-top-agent* if its topicality is between 0.85 and 0.90, etc.

Definition of what a prototypical verb looks like:

- central-trans ([] , word) :-

$$\left[\begin{array}{l} \boxed{1} \text{ SYNSEM|LOCAL } \left[\text{CONT} \left[\begin{array}{l} \text{AG } \boxed{2} \\ \text{PT } \boxed{3} \end{array} \right] \right] \end{array} \right],$$

neutral-top-agent ($\boxed{2}$),

neutral-top-patient ($\boxed{3}$).

marginal-trans ($\boxed{1}$, word) :-

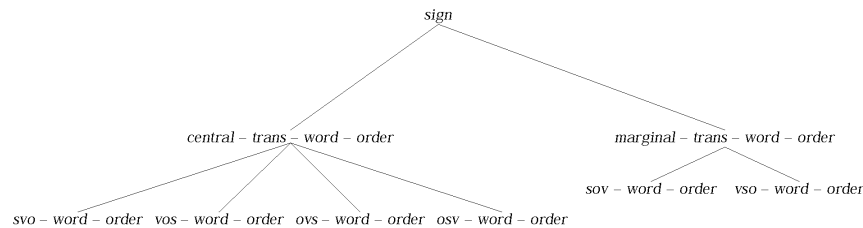
not (central-trans ($\boxed{1}$)).

A predicate is considered to be of central transitivity (type *central-trans*) if *neutral-top-agent* is true of its agent and *neutral-top-patient* is true of its patient (i.e., subject and object are within their normal topicality ranges).

Type Partitions:

central-trans-word-order (subtype of *sign*): *svo-word-order*, *vos-word-order*, *ovs-word-order*, *osv-word-order*

marginal-trans-word-order (subtype of *sign*): *sov-word-order*, *vso-word-order*



central-trans-word-order

predicate ($\boxed{1}$),

central-trans ($\boxed{1}$).

sov-word-order

predicate ($\boxed{1}$),

$$\left[\begin{array}{l} \boxed{1} \text{ SYNSEM|LOCAL } \left[\text{CONT} \left[\text{AG } \boxed{2} \right] \right] \end{array} \right],$$

increased-top-agent ($\boxed{2}$).

The metagrammar evaluates signs based on appropriateness conditions. Signs of type *central-trans-word-order* must be predicates which belong to the type *central-trans* which in turn requires its arguments to be of neutral topicality. The other two word orders in Polish belong to the type *marginal-trans*, which favors non-neutral topicality for its arguments. For SOV word

order, we require an additional constraint: that the predicate's agent be of unusually high topicality (and thus be of type *increased-top-agent*)

These types are subject to two constraints: one affecting the sign's PHON feature (a matter of observable form), and another affecting SYNSEM (the syntactic-semantic properties of the sign). We formalize this by placing an ordering constraint on the phonologies on signs of each word order type (in terms of linear precedence constraints), while SYNSEM is constrained by the metagrammar's appropriateness conditions.

(8) Principle of Construction Evaluation

Let C be a construction. Then C is valued highly by the evaluation metric if C is headed by a predicate which is either a central or a marginal member of the transitive prototype.

This paper has argued that linguistic generalizations such as those of relative transitivity and topicality cannot be captured by transformations, lexical rules, or other derivations, but should rather be modeled as signs weighed by an evaluation metric. Siewierska's (1993) quantitative analysis of word order in Polish suggests that Polish speakers encode unusual topicality for subjects and/or objects by means of two word order patterns. This is formalized by specifying constraints on a predicate such that the prototypical transitive construction's arguments must be of neutral topicality for their type (subject or object), and then describing these two Polish word order patterns as subtypes of a type that requires their predicates to not be of a more marginal type. While this particular choice may be specific to Polish, it is proposed that other languages would encode marginal topicality (or other cognitively plausible and appropriate information) in other features of the sign, including word order, phonology, and morphology, in a manner similar to that outlined above.

Works Cited

- Abney, Steven. (1996b) "Statistical Methods and Linguistics." In Klavans, Judith and Philip Resnik [eds.] *The Balancing Act* Cambridge: MIT Press.
- Estival, Dominique and John Myhill. (1988) "Formal and functional aspect of the development from passive to ergative systems." In Shibatani, Masayoshi [ed.] *Passive and Voice* Amsterdam: John Benjamins.
- Givón, Talmy. (1983) *Topic Continuity in Discourse: A Quantitative Cross-Language Study* Amsterdam: John Benjamins.
- Hopper, Paul J. and Sandra A. Thompson. (1980) "Transitivity in grammar and discourse." *Language* Volume 56, Number 2, 251-299.
- Keenan, Edward L. (1985) "Passive in the world's languages." In Shopen, Timothy [ed.] *Language Typology and Syntactic Description* Cambridge: Cambridge University Press.
- Murphy, Patrick M. (1995a) "The Theme Construction: Discourse in Information-Based Syntax and Semantics." Presentation to the Spring Linguistics Colloquium, University of North Carolina at Chapel Hill, 1 April 1995.

- Myhill, John. (1992) *Typological Discourse Analysis: Quantitative Approaches to the Study of Linguistic Function* Oxford: Blackwell.
- Pollard, Carl and Ivan A. Sag. (1987) *Information-Based Syntax and Semantics, Volume 1: Fundamentals* ESLI Lecture Notes no. 13. Stanford: Center for the Study of Language and Information (distributed by the University of Chicago Press).
- Pollard, Carl and Ivan A. Sag. (1994) *Head-Driven Phrase Structure Grammar* Stanford: Center for the Study of Language and Information (distributed by the University of Chicago Press).
- Siewierska, Anna. (1993) "Syntactic weight vs information structure and word order variation in Polish." *Journal of Linguistics* 29, 233-265.
- Svartvik, Jan. (1966) *On Voice in the English Verb* The Hague: Mouton.