



The Usage-Based Model (II): Performance-Grammar Correspondences

Why seek 'converging evidence'?

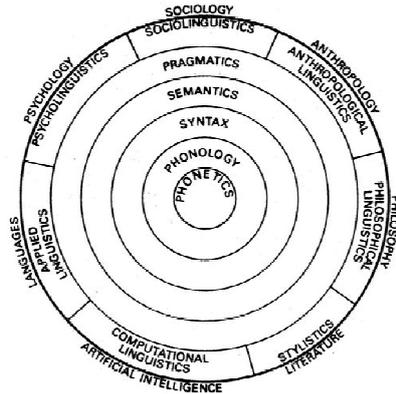
Assumptions and predictions of the PGCH

Some methodological issues

1. Converging evidence

important theoretical arguments: compartmentalisation of linguistics in the 20th century and an ensuing lack of dialogue between people working in different branches of linguistics

- ▶ linguistics in the 21st century: studying 'language structure in its human context' ('integrative functionalism', cf. CROFT 1995, to appear)
- ▶ implication for typologists: recognition that there are "close correlations between universal preferences in structure with universal preferences in cognition and communication" (BICKEL 2007: 240)



plus: converging evidence can give more strength to the plausibility of functional motivations: a usage-based corpus-linguistic approach anchors typological distributions in usage-patterns of individual languages, where functional motivations actually operate.

methodological argument: the type of evidence for the motivation of language structures consists in "answers based not on philosophical speculation but on the growing body of empirical data from grammars and from performance" (HAWKINS 2004: xi)

2. Assumptions and predictions of the PGCH

2a. Symmetrical relationship between performance and grammar Grammars are not just 'accessed' in performance, but language use actively shapes individual grammars.

- exemplar-based representations of linguistic knowledge (BYBEE 2001)
- large-scale structural regularities (e.g. 'the grammar of English/Italian/Mapudungun') are emergent phenomena (HOPPER 1987):

"When human beings use symbols to communicate with one another, stringing them together into sequences, patterns of use emerge and become consolidated into grammatical constructions." (TOMASELLO 2003: 5)

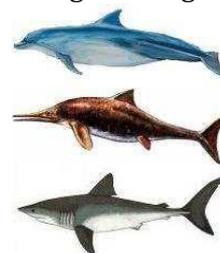
2b. Variation as a fundamental property of language, both intra- and cross-linguistic. Individual utterances in communicative interaction replicate conventionalised form-function pairings. Break with conventions on purpose (expressivity, economy) or unintentionally (reanalysis/ misinterpretations) [cf. Croft 2006]. Altered replication introduces variation. Norm: co-existence of several variants of a functionally similar construction, with **quantifiable preferences (i.e. absolute token frequencies) in language use (preference rankings A>B>C>D)**

► **Performance-grammar correspondence hypothesis (PGCH) I:**

Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (HAWKINS 2004: 3)

Cross-linguistic variation exists because of the fundamental diversity of structural options for dealing with the same communicative problem. But despite the principal arbitrariness of the sign, grammars do not vary without limits and in entirely unpredictable ways, BECAUSE:

2c. Grammars are shaped by various ‘external’ forces which act as selection pressures, as it were, on the nature and the combinatorial possibilities of linguistic signs. Hawkins (2004) emphasises processing pressures (which we will encounter multiple times during this course), but functionalists have recognised the influence of other factors as well. Grammars have come to be seen as functional adaptations to such pressures. In Deacon’s (1997) terms, structural similarities between unrelated languages are analogous features that have evolved by **convergent evolution** (just as in unrelated biological species subjected to the same environmental pressures).



“[..E]xternal functional factors determine both individual linguistic behaviour (innovations) and crosslinguistic/diachronic patterns that are obviously not part of an individual’s grammar. Since these are **independent sources of evidence**, one can develop hypotheses of functional factors on the basis of one source of evidence and test them in the other.” (CROFT 1995: 525)

► **2d. Performance-grammar correspondence II: Language-specific preference patterns match cross-linguistic structural distributions.** In other words, there is

“a correspondence between preferences in performance in languages with choices and variants, and preferences in the grammatical conventions of languages with fewer choices and variants.” (HAWKINS 2004: 255)

Three more specific predictions of the PGCH

Grammars are ‘frozen’ or ‘fixed’ performance preferences. If a structure A over A’ of the same structural type in performance, then A will be more productively grammaticalized, in proportion to its degree of preference; if A and A’ are more equally preferred, then A and A’ will both be productive in grammars. (Hawkins 2004: 6)

- (1) variation in word order (SVO and OVS) in a single language corresponds to universally preferred orders (subjects before objects)
YAMASHITA 2002: *SO > OS in Japanese (1.43% scrambled sentences)*
- (2) Greenbergian word order correlations (GREENBERG 1966a):
 - VO & PrepN *frequent and productive across grammars (A)*
 - OV & NPostp *frequent and productive across grammars (B)*
 - VO & NPostp *infrequent and non-productive across grammars (A’)*
 - OV & PrepN *infrequent and non-productive across grammars (B’)*

A special case of this prediction: ***hierarchical preferences and conventionalisations***
“Markedness hierarchies are conventionalizations of performance frequency rankings.”
 (Hawkins 2004: 64)

- (3) Greenberg’s markedness hierarchies (Greenberg 1966b):
- a. Nom > Acc > Dat > other (case marking)
 - b. Sing > Plur > Dual > Trial/Paucal (number)
 - c. Masc, Fem > Neut (gender)
 - d. Positive > Compar > Superl (adjectival forms)

morphological inventories: the presence of a distinctly grammaticalized feature on each hierarchy implies the presence of all higher features (i.e. those to the left).

systematic increase in formal marking down the hierarchies (e.g. by non-zero v. zero forms) as well as a decrease in the inflectional potential (i.e. distinct feature combinations) of each category.

Crucially, the hierarchies correlate with frequency rankings in languages that have productive morphemes for each category:

(4)	Number in Sanskrit nouns	(5)	Case in German nouns (Primus 1999)
	SINGULAR 70.3%		NOM 17,500 tokens
	PLURAL 25.1%		ACC 9,700 tokens
	DUAL 4.6%		DAT 5,450 tokens

Processing effects: economical coding (Zipf 1935, CROFT 2003, HASPELMATH 2008) and entrenchment (BYBEE 1985) of unmarked (higher-ranking) category

► inflectional complexity in marked paradigms is disfavoured psychologically and levelled out diachronically ► the systematic loss of entire cases or of inflectional distinctions within cases proceeds in keeping with the synchronic typological hierarchy:

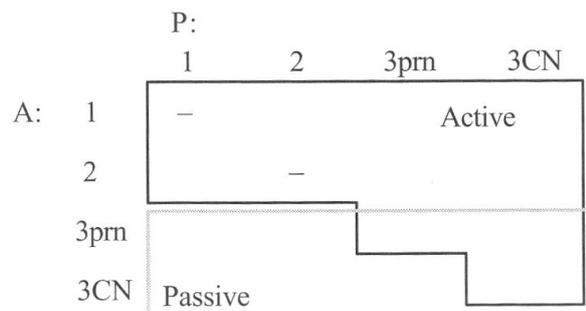
- (6) Nom/Acc/Dat (Old English) > Nom/Acc-Dat (merge/collapse in Modern English)
 (7) progressive loss of inflectional distinctions in marked categories (PL, DAT, NEUT) > increase in case syncretism (Old High German 11 distinctions, Middle High German 10, New High German 8; cf. HAWKINS 2004: 77)

“The progressive loss of feature combinations in [...] morphological systems is highly correlated with frequency of use. [...] It is therefore plausible to assume that performance preferences operate at each historical stage and exert a selective pressure in favour of the retention or loss of certain distinctions.” (HAWKINS 2004: 79)

Categorical and preferential choices of voice constructions

“The same categorical phenomena which are attributed to hard grammatical constraints in some languages continue to show up as statistical preferences in other languages” (Bresnan et al. 2001).

► prototypical combination of animacy and participant role in that As are typically high on the person scale (1,2), while Ps are comparatively low (3rd person), and in many languages, the distribution of active and passive constructions is sensitive to this typological prototype (CROFT 2003: 175):



This principle carries over to many kinds of overt/covert choice in performance, e.g. explicit relativizers and complementizers, adjectives with and without agreement marking, case copying, control structures versus finite complementation, use of resumptive pronouns etc.

Performance and grammatical conventions are shaped by competing motivations.

If two preferences P and P' are in (partial) opposition, then there will be variation in performance and in grammars, with both P and P' being realised, each in proportion to its degree of motivation in a given language structure. (Hawkins 2004: 6)

► principled patterns of symmetry and asymmetry in constituent orders

“Asymmetric ordering preferences in performance and grammars [are] those in which different preferences converge on a single order, whereas symmetries (VO and OV) have complementary preferences and dispreferences in each order.” (HAWKINS 2004: 259)

Examples of asymmetries:

- (8) distributional skewing of *wh*-elements to the left of their sister clauses
- (9) subjects preceding direct objects, even in languages that have a choice (see above)
- (10) NRel in VO languages (unity, hardly any exceptions)

Examples of symmetries:

- (11) RelN and NRel in OV languages (conflicting demands of minimal processing domains and ambiguity avoidance)
- (12) [V PP PP] constructions in which one PP is longer (and would thus be shifted due to minimal domain preference), but contains a lexical dependency with the verb (e.g. *count on* – preferred adjacency of lexical dependent) = two preferences P and P' are in partial opposition, hence there is variation and both patterns exist (symmetrically)

Table 2: Weight and lexical dependency in English prepositional phrase orderings
(HAWKINS 2000: 247)

n=206	Pd>Pi by			Pd=Pi	Pi>Pd by		
	5+	2-4	1		1	2-4	5+
[V Pd Pi]	2	6	17	24	23	49	30
[V Pi Pd]	28	12	6	5	2	2	0

Example: The man counted on his son [in his old age].
lexical dependence (Pd) independent PP (Pi)

- (13) productive grammaticalization of both OV and VO orders in the world's languages: V and its object-NP have a mutual dependency relationship (V depends on the NP for its specific meaning assignment (*run (in the park)* versus *run a shop*), and the NP depends on the verb for its thematic role assignment and for providing the mother node VP to which it is attached) – therefore, both orderings (VO and OV) are well-attested. But it is clear that “whichever order is selected in a grammar conventionalises one preference at the expense of another, and compensating syntactic and morphosyntactic properties are predicted as a consequence.” (HAWKINS 2004: 259)

- ▶ each order typically comes with a specific set of morphosyntactic marking that maximises on-line property assignment:

“the benefits of the suppressed order w[ill] assert themselves [...] in transformations of the basic word order, morphosyntactic devices, or other structural responses. [...] One productive way of doing this is through rich verb agreement and rich case marking, and the result is a predicted set of asymmetries in which morphosyntactic marking is skewed to the left of the clause.” (HAWKINS 2004: 245)

Implicit in the above-mentioned assumptions is, of course, the view that diachronic change is the key mediating mechanism which links the conventionalisation of performance preferences in single languages with large-scale typological distributions, cf. HASPELMATH 2008.

3. Methodological issues

In typology: moving away from ‘holistic typology’ to coding, classifying and correlating small individual parameters (BICKEL 2007); construction-specific work (e.g. recognition of the local nature of syntactic categories, VAN VALIN and LAPOLLA 1997, CROFT 2001) and auto-typological methods of developing variables (BICKEL and NICHOLS 2007); probabilistic nature of universals (statistical implications).

In language-specific research: ‘more rigorous corpus linguistics’ (GRIES 2006): importance of sophisticated statistical analyses, e.g. recognition of competing motivations needs to be translated into multivariate statistical methods (cf. DIESEL to appear).
Therefore:

“While in some cases cross-linguistic evidence can in fact be indicative of such tendencies (i.e. language-specific correlations, K.S.), the methods by which such claims are to be supported definitely need to be more advanced. [...] In the context of variation phenomena, such techniques are especially rewarding since they enable us to subject our theories to the most rigorous test conceivable, namely to actually predict what native speakers will do in a particular situation.” (GRIES 2003: 171-73)

In both disciplines, therefore: advances in quantitative analyses (e.g. permutation/Monte-Carlo statistics reflecting the peculiarities of corpora and typological sampling, cf. JANSSEN et al. 2006).

References

- Bickel, Balthasar (2007). Typology in the 21st Century: Major Current Developments. *Linguistic Typology* 11: 239-251.
- Bickel, Balthasar and Johanna Nichols (2007). Inflectional Morphology. In: *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon* 2nd ed. Ed. Timothy Shopen. Cambridge: Cambridge University Press. 169-240.
- Bresnan, Joan, Shipra Dingare and Christopher D. Manning (2001). Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. *Proceedings of the LFG 01 Conference*. Eds. Miriam Butt and Tracy Holloway King. CSLI Publications. On-line.
- Bybee, Joan L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Bybee, Joan L. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Croft, William (1995). Autonomy and Functionalist Linguistics. *Language* 71.3: 490-532.
- Croft, William (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Croft, William (2003). *Typology and Universals*. 2nd ed. Cambridge: Cambridge University Press.
- Croft, William (2006). Evolutionary Models and Functional-Typological Theories of Language Change. In: *The Handbook of the History of English*. Eds. Ans van Kemenade and Bettelou Los. Oxford: Blackwell. 68-91.
- Croft, William (to appear). Language Structure in its Human Context: New Directions for the Language Sciences in the 21st Century. In: *The Cambridge Encyclopaedia of the Language Sciences*. Ed. Patrick Colm Hogan. Cambridge: Cambridge University Press.
- Deacon, Terence (1997). *The Symbolic Species: The Co-Evolution of Language and the Human Brain*. London: Penguin.
- Diessel, Holger (to appear). Iconicity of Sequence: A Corpus-Based Analysis of the Positioning of Temporal Adverbial Clauses in English. *Cognitive Linguistics*.
- Greenberg, Joseph H. (1966a). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: *Universals of Language*. 2nd ed. Ed. Joseph H. Greenberg. Cambridge, Mass.: MIT Press. 73-113. [1st ed. 1963]
- Greenberg, Joseph H. (1966b). *Language Universals, with Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Gries, Stefan Th. (2003). Grammatical Variation in English: A Question of 'Structure vs. Function'? In: *Determinants of Grammatical Variation in English*. Eds. Günter Rohdenburg and Britta Mondorf. Berlin, New York: Mouton de Gruyter (TiEL 43). 155-173.
- Gries, Stefan Th. (2006). Some Proposals towards a More Rigorous Corpus Linguistics. *Zeitschrift für Anglistik und Amerikanistik* 54.2: 191-202.
- Haspelmath, Martin (2008). Creating Economical Morphosyntactic Patterns in Language Change. In: *Language Universals and Language Change*. Ed. Jeff Good. Oxford: Oxford University Press.
- Hawkins, John A. (2000). The Relative Order of Prepositional Phrases in English: Going beyond Manner-Place-Time. *Language Variation and Change* 11: 231-266.
- Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Hopper, Paul J. (1987). Emergent Grammar. In: *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society*. Eds. Jon Aske, Natasha Beery, Laura Michaelis and Hana Filip. Berkeley: University of California Press. 139-157.
- Janssen, Dirk P., Balthasar Bickel and Fernando Zúñiga (2006). Randomization Tests in Language Typology. *Linguistic Typology* 10.3: 419-440.
- Jelinek, Eloise and Richard Demers (1983). The Agent Hierarchy and Voice in some Coast Salish Languages. *International Journal of American Linguistics* 49: 167-185.
- Primus, Beatrice (1999). *Cases and Thematic Roles: Ergative, Absolutive and Active*. Tübingen: Niemeyer.
- Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Van Valin, Robert D. Jr. and Randy LaPolla (1997). *Syntax: Structure, Meaning, Function*. Cambridge: Cambridge University Press.
- Zipf, George (1935). *The Psychobiology of Language: an Introduction to Dynamic Philology*. Cambridge, Mass.: MIT Press.