What the conventions of grammars tell us about communicative efficiency: Some current issues and prospects

John A. Hawkins
University of California Davis & Cambridge University

Keywords: Grammatical conventions, left-right asymmetries, Maximize On-line Processing, Uniform Information Distribution, Prediction

Many early ideas about communicative efficiency were inspired by grammatical patterns across languages. Greenberg (1966) proposed morphological hierarchies such as Singular > Plural > Dual > Paucal for number marking on nouns, and Nominative > Accusative > Dative > Other for case marking, based on the cross-linguistic distribution of these morphemes and on patterns of allomorphy (with richer morphological variants and more zero expression at the top end of these hierarchies). In order to explain these patterns in grammars he pointed to correlations with (declining) frequencies of usage within languages (e.g. Singular nouns in a corpus of Sanskrit = 70.3%, Plural = 25.1%, Dual = 4.6%), and these performance-grammar correspondences then led to efficiency principles such as Minimize Forms in Hawkins (2004, 2014). Greenberg's (1963) word order universals and his proposed correlations between head orderings in grammars (e.g. VO languages prefer Prepositions before NP, OV languages prefer NP before Postpositions, cf. Dryer 1992) inspired the processing principles of Minimize Domains (Hawkins 2004, 2014) and Dependency Length Minimization (Futrell et al. 2015), both of which were formulated in terms of the relative efficiencies of competing word order selections within languages as measured in usage corpora. Keenan & Comrie's (1977) Accessibility Hierarchy for relative clauses across grammars also contributed to, and has been explained in terms of, Minimize Domains.

The present paper points to some lessons that cross-linguistic grammatical patterns can teach us today for issues in efficiency research. In particular, there are numerous "left-right asymmetries" that compete with Greenberg's largely symmetrical (VO vs OV) patterns, and these have been relatively neglected in the literature but were captured in the Maximize On-line Processing principle of Hawkins (2004, 2014). Grammars provide evidence for a preferred "front loading" of information in a sentence (i.e. get as much information to the hearer as quickly as possible). This is seen particularly clearly whenever one linguistic category is strongly dependent on another, like a gap on a moved filler, or an anaphor on its antecedent, or a narrow-scope quantifier on a wide-scope one, or a predication on a topic. The informationally richer category either invariably or at least generally precedes the dependent one in these cases. Similarly, rich verb agreement patterns (whereby the verb has inflectional agreement with two or more arguments) is particularly favored in languages in which the verb occurs early in the clause, and rich case marking (for inflectional marking on nouns that differentiates between several cases) is favored in languages in which nouns occur early and the verb is final (SOV). These grammatical data suggest that early positioning in language usage is favored not so much for categories that predict what lies ahead (so reducing surprisal, cf. Levy 2008), but for those that are rich in explicit content and to which subsequent categories need access for full and immediate processing. These data also provide a large class of exceptions to the claim that information density is systematically uniform throughout a clause (Jaeger 2010). These exceptions involving the front loading of information in grammatical conventions are exemplified and quantified using cross-linguistic data from typological samples including the World Atlas of Language Structures (Dryer & Haspelmath 2013). They should now be heeded and their lessons for efficiency in language usage should be tested on usage corpora and in experiments.

References

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81-138.
Dryer, Matthew S. & Martin Haspelmath, (eds.). 2013. *WALS Online* (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.738553 (Available online at https://wals.info)
Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112(33). 10336-41
Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 73-113. Cambridge, MA: MIT Press.
Greenberg, Joseph H. 1966. *Language Universals with Special Reference to Feature Hierarchies.* The Hague: Mouton.
Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
Hawkins, John A. 2014. *Cross-linguistic Variation and Efficiency*. Oxford: Oxford University Press.
Jaeger, T. Florian. 2010. Redundancy and reduction: Speakers manage syntactic information Density. *Cognitive Psychology* 61. 23-62.
Keenan, Edward L. & Bernard Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63-99.
Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106. 1126-77.