

The constant flow of linguistic information: subject first preference in word order and in ambiguity resolution

I here start with an outline of the *constant flow of linguistic information hypothesis* (Fenk & Fenk 1980), and propose it as a framework for understanding the preponderance of the subject initial word orders SOV and SVO across languages. I will then discuss a probabilistic-based explanation for the subject preference in ambiguity resolution across languages.

- The *constant flow of linguistic information hypothesis*:

In an effective and economical communication system, the information transmitted should be distributed as uniformly as possible across time, and the average level of information transmitted per time should be adapted to our capacity limits (Fenk & Fenk 1980:402; see also Jäger 2010).

We propose two main mechanisms that contribute to a rather constant flow of linguistic information: (i) The regularity “the more frequent, the shorter”. According to information theory high relative frequency (\approx high probability) is related to low informational content. An element carrying a small amount of information can be processed within a shorter time. (ii) The regularity “the more predictable first”. This regularity seems to compensate for the successive reduction of information over time. In general, as a sentence continues, the number of possible continuations and thus subjective information decreases. To place elements carrying a high amount of information at the beginning of a sequence, i.e. a position which is *per se* characterized by high information (Shannon 1951), would produce peaks of cognitive overload (Fenk-Oczlon 1983, 1989, 2001).

- Explaining the preponderance of SOV and SVO word orders across languages

In Fenk-Oczlon (1983) I argued that a strategy favoring highly predictable units at the beginning of sentences could help ensure a constant flow of information and potentially account for the cross-linguistic preponderance of the subject first word orders SOV and SVO. The rationale: Subjects are highly discourse prominent and most frequently topic (old information) and therefore more predictable than verbs or objects which are prototypically comments and new information. The power of the processing strategy “the more predictable first” shows also in the ordering of binominals (Fenk-Oczlon 1989).

- Subject preference in ambiguity resolution

Recent ERP studies (e.g. Bickel et al. 2015, Wang et al. 2009) provide convincing evidence for a S/A preference in ambiguity resolution of initial arguments. This processing strategy seems to be universal and can even be observed in languages like Hindi or Mandarin in which the grammatical category subject is not as clear-cut. The *principle of simplicity* has been put forward as a potential explanation for the observed S/A preference (Bornkessel & Schlesewsky 2006): initial ambiguous arguments are first analyzed as the sole argument (subject) of an intransitive verb.

Here, I will suggest a probabilistic approach to ambiguity resolution, favoring the most probable units. I will argue that subjects have the highest probability as initial arguments even in languages such as Mandarin – frequently considered to be topic prominent – or Hindi, which has split ergativity. Further interrelationships between predictability of NPs and morphological case marking (nominative, ergative, absolutive) will be discussed.

- Bornkessel, I., & Schlesewsky, M. (2006). The Extended Argument Dependency Model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113, 787-821.
- Fenk, A. & Fenk, G. (1980) Konstanz im Kurzzeitgedächtnis – Konstanz im sprachlichen Informationsfluß? *Zeitschrift für experimentelle und angewandte Psychologie* 27, 400-414
- Fenk-Oczlon, G. (1983) Ist die SVO-Wortfolge die „natürlichste“? *Papiere zur Linguistik* 29, 23-32
- Fenk-Oczlon, G. (1989) Word frequency and word order in freezes. *Linguistics* 27, 517-556
- Fenk-Oczlon, G. (2001) Familiarity, information flow, and linguistic form. In J. Bybee, & P. Hopper (eds.) *Frequency and the emergence of linguistic structure*, 431-448. Amsterdam/Philadelphia: Benjamins
- Jaeger, T.F. (2010) Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61, 23-62
- Shannon, C.E. (1951) Prediction and entropy of printed English. *The Bell System Technical Journal* 30, 50-54
- Wang, L., Matthias Schlesewsky, M, Bickel, B. & Bornkessel- Schlesewsky I. (2009) Exploring the nature of the ‘subject’-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes* 24 (7/8), 1180-1226