

From VSO to SVO in Austroasiatic? Towards a study of pragmatic pressures in the loss of verb-initial order

Wei-Wei Lee
Institute of Comparative Linguistics (UZH)

Verb-initial word orders (VSO/VOS) seem to be widely disfavored in the world's languages. Only around 9% of the languages exhibits a basic order that is verb-initial (Dryer 2013), and this is never the exclusive order (Greenberg 1966:79; Nichols 2003:305). Verb-initial orders have moreover been shown to conflict with cognitive biases in language processing and communication. Yet, VSO and VOS orders are found to a considerable degree within the dominantly SVO-ordered Austroasiatic language family (Jenny forthcoming). The Austroasiatic languages are spoken today in an area stretching from eastern India to Vietnam. Interestingly, the V1 structures occur in geographically scattered languages across very diverse language groups. While the appearance of VOS order in the Nicobarese languages can be attributed to areal influence, the unexpected VSO structures attested in other language groups are still in need of an explanation.

The first theoretical paper of my cumulative dissertation will tentatively posit a scenario in which the VSO order already existed in the protolanguage, explaining the current VSO structures as remnants. This scenario assumes that this order was lost in the present-day SVO languages. A shift from VSO to SVO is deemed more likely than the reverse process, based on the (intra- and interlinguistic) locus of the structures and evidence of similar developments in other language phyla. In many of these cases, pragmatic factors have been suggested to play a significant role in the process (e.g. Donohue 2005; Dimmendaal 2006). I suggest that this may also be the case in Austroasiatic languages. Languages in Mainland Southeast Asia are characterized by a highly pragmatically oriented grammar (Bisang 2006). Furthermore, word order variation triggered by information-structural aspects is common in languages in the area, including the Austroasiatic languages (Jenny forthcoming). For most of the languages in question (i.e. with attested VSO structures), however, the little data currently available are only vaguely indicative at most. At this stage, it remains unclear what factors the alternation of word order is conditioned by. Collection of additional narratives and conversations and the resulting corpus will enable the first systematic investigation of the role of pragmatics in these languages. This also includes the interaction with other linguistic domains, particularly morphosyntax and semantics.

The follow-up, data-driven study will focus on the Palaungic branch. Fieldwork in Myanmar will start in July, the output of which is expected to enable an exploratory pilot study. This is an essential step prior to corpus annotation, not only for practical reasons, but perhaps more importantly because of the lack of notional and terminological consensus in the field of information structure (Jacobs 2001; Song 2012:4; Matic' & Wedgwood 2013). This problem, along with other issues, has shaped an incongruous picture of research in this area with often incompatible results. Several of these issues will be briefly addressed during the talk. For this pragmatic study, I aim to create a broad cross-linguistic overview of the documented findings so far with respect to VS/SV order variation and change. The challenge is to reconcile a range

of diverging approaches and employed concepts. Only in this manner will cross-linguistic comparison become possible.

The approach taken here is both fieldwork- and corpus-based. After a brief scan of the first texts and selection of the pragmatic factors to examine, these will be coded in the corpus on separate tiers, in addition to morphosyntactic and semantic tiers. Detailed annotation guidelines will be set up beforehand to maximize homogeneity across annotators. Due to the fuzzy nature of notions like 'topic' and 'focus' (e.g. Vallduví 1990; Chafe 1994; Lambrecht 1994), it is most likely that more basic features will be used instead into which these concepts may be deconstructed, such as 'given' and 'new'. In addition to theoretical reasons, it has also been shown that the inter-annotator agreement is higher for the latter set of features in comparison to the former (Ritz et al. 2008). I will point out some issues that are involved in coding information-structural features. Finally, the annotated corpus will enable statistical analysis of frequencies and mixed effects.

In my subsequent research, I may look at argument coding in the Austroasiatic languages from a typological perspective. This is an area in which interactions with word order are to be expected, especially in the isolating and largely zero-marking Austroasiatic languages in Southeast Asia. The exact focus of my following papers will however to a large extent depend on what I find in the new data.

References

- Bisang, Walter. 2006. Southeast Asia as a Linguistic Area. *Encyclopedia of Language & Linguistics*. 587–595.
- Chafe, Wallace L. 1994. *Discourse, Consciousness, and Time: the Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Dimmendaal, Gerrit Jan. 2006. Head marking, dependent marking and constituent order in the Nilotic area. In F. K. Erhard Voeltz (ed.), *Studies in African Linguistic Typology*, 71–92. Amsterdam, Philadelphia: John Benjamins.
- Donohue, Mark. 2005. Word order in New Guinea: Dispelling a myth. *Oceanic Linguistics* 44(2). 527–536.
- Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Greenberg, Joseph H. 1966. *Universals of Language*. 2nd ed. Cambridge, London: MIT Press.
- Jacobs, Joachim. 2001. The dimensions of topic–comment. *Linguistics* 39(4). 641–681.
- Jenny, Mathias. Forthcoming. *Verb-initial structures in Austroasiatic languages*.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.
- Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49. Cambridge University Press. 127–163.
- Nichols, Johanna. 2003. Diversity and stability in language. In Brian D. Joseph & Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 283–310. Malden/Oxford/Melbourne/Berlin: Blackwell.
- Ritz, Julia, Stefanie Dipper & Michael Götze. 2008. Annotation of information structure: An

evaluation across different types of texts. *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008, 26 May – 1 June 2008*, 2137–2142.

Song, Jae Jung. 2012. *Word Order*. Cambridge: Cambridge University Press.

Vallduví, Enric. 1990. The Informational Component. Philadelphia: University of Pennsylvania PhD dissertation. http://repository.upenn.edu/ircs_reports (20 March, 2018).