Title: Word order in zero-marking languages: A typological study
Category: oral or poster

It has often been noted that word order type might correlate with type of morphological marking of core arguments cross-linguistically. Word order type here means the dominant arrangement of the main clausal elements (S, O, and V) in a language. Core argument marking is about discriminating the arguments of a prototypical two-place transitive predicate, one more agent-like (S) and the other more patient-like (O) (Comrie 2005). Two types of morphological marking, head marking (or cross-reference) and dependent marking (or case), interact in this domain. While earlier studies (e.g. Nichols 1992) have investigated the correlation of word order and type of overt morphological marking, correlation of word order and absence of overt morphological marking, here called zero marking, has not been fully studied. Zero-marking languages [1-2] are generally expected to have dominant SVO order, which is purportedly the most economic word order available for argument discrimination. Regardless of SVO being motivated in zero-marking languages, several counterexamples exist around the globe [3-5]. West Africa and Southeast Asia are further well-known hotbeds of zero-marking SVO languages, but it is unclear whether zero marking correlates with SVO outside these areas. If not, the proposed correlation could be a mere consequence of areal tendencies, not of universal principles.

In my presentation, I discuss the results of modeling the distribution of zero marking over word order, areas and language families by multiple logistic regression in the spirit of Bickel (2008). Logistic regression tests to what extent the independent variables (here word order and areas) can predict the outcome of the dependent variable (here zero vs. overt marking). Logistic regression enables modeling e.g. the effect of areas and test whether a hypothesized association between two structural variables (zero marking and word order) is independent of them. The effect of each variable of interest was tested with maximum likelihood estimation, comparing a model with the variable of interest to one without it; p-values were deduced by randomization. Different areal resolutions were used to test the effect of areas. The effect of genealogical affiliation was tested by assessing diachronic intra-family tendencies with the Skewed Family Method (Bickel 2008), which estimates whether languages in genealogical units have maintained/developed the same linguistic property to a statistically significant degree. Data comes from a sample of ca. 800 languages, which aims at being genealogically and areally maximally representative. Based on the tests, zero marking is argued to correlate with SVO independently of areas and genealogical affiliation despite the distribution of zero marking being areally and genealogically skewed. The method allows universals, such as the proposed correlation, to be interpreted dynamically as universal structural pressure on how languages change over time.

Examples:

(1) Thai (Kam-Tai; Thailand; Iwasaki and Ingkaphirom 2005: 110)

Lék	*tè*	*nɔ̀ɔy*
Lek	kick	Noy
S	V	O

'Lek kicks Noy.'

(2) Iquito (Zaparoan; Peru; Brown 2004: 13)

Ja	*Marcos*	*jihuɨɨta-qui*	*nu-canɨɨsi*
Already	Marcos	find-COMPL	3SG-bag
	S	V	O

'Marcos has already found his bag.'

(3) Kolokuma Ijo (Ijoid; Nigeria; Williamson 1965: 58)

ígeré	*eré-ma*	*kọ̀rị-mị*
crocodile	woman-DEF	catch-PST
S	O	V

'A crocodile caught the woman.'

(4) Arára Karó (Tupian; Brazil; Gabas 1999: 153)

Iyõm	*wat*	*awe*	*cape-t*
father	1SG.POSS	brother	beat-IND
S		O	V

'Father beat my brother.'

(5) Quiegolani Zapotec (Zapotecan; Mexico; Black 2000: 45)

W-eey	*Benit*	*mël*
COMPL-take	Benito	fish
V	S	O

'Benito took a fish.'

Abbreviations: 1 first person, 3 third person, COMPL completive, DEF definite, IND indicative, O object, POSS possessive, PST past tense, S subject, SG singular, V verb.

References

Bickel, B. 2008. *A general method for the statistical evaluation of empirical universals*. Available at <http://www.uni-leipzig.de/~bickel/>.

Black, C.A. 2000. *Quiegolani Zapotec Syntax: A Principles and Parameters Account*. Dallas: SIL and University of Texas at Arlington.

Brown, M. 2004. *Topics in Iquito Syntax: Word Order, Possession, and Nominal Discontinuity*. MA Thesis, University of Texas at Austin.

Comrie, B. 2005. Alignment of case marking. In *The World Atlas of Language Structures*, M. Haspelmath, M. Dryer, D. Gil, and B. Comrie (eds), 398-405. Oxford: OUP.

Gabas, N. 1999. *A Grammar of Karo, Tupi (Brazil)*. PhD Dissertation, University of California at Santa Barbara.

Iwasaki, S. and P. Ingkaphirom 2005. *A Reference Grammar of Thai*. Cambridge: CUP.

Nichols, J. 1992. *Linguistic Diversity in Space and Time*. Chicago: The University of Chicago Press.

Williamson, K. 1965. *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press in association with The West African Languages Survey and The Institute of African Studies, Ibadan.