

# Mouse chases cat

## Exploring unexpected word order patterns in Jakarta Indonesian and Kupang Malay using quantitative methods

Tina Gregor

Isolating languages usually have a fairly fixed word order. This is reflected by a statistically significant correlation between the lack of case marking and rigid (only one word order is possible) or restricted word order (the basic word order and one variant are possible) as well as by a significant correlation between the lack of agreement marking on the verb and rigid or restricted word order, which was found by Siewierska (1998: 507-512) in her large-scale typological study. In this talk, I will explore unexpected word order variants in Jakarta Indonesian and Kupang Malay in a sample from the Jakarta Field Station corpus using quantitative methods.

Jakarta Indonesian (cf. Sneddon 2006) and Kupang Malay (cf. Paauw 2008) both have neither case nor agreement marking and are classified as SVO languages. However, SVO is not the only word order pattern that occurs in these languages. In the extensive corpus of spoken language collected by the Jakarta Field Station, four word order variants can be found for Jakarta Indonesian and five for Kupang Malay. Indeed, the word order variation found in the latter language exhausts all the possible permutations. Utterances with non-SVO order are not very frequent in the sample I drew from the corpus; only 17.16% for Jakarta Indonesian and 8.25% in Kupang Malay. Still, they show that both languages are more flexible when it comes to the placement of arguments relative to the verb than one would expect based on Siewierska's (1998) results, and they also reveal some surprising differences between Jakarta Indonesian and Kupang Malay. Statistical tests such as the Chi-squared method indicate that there is a significant difference between Kupang Malay and Jakarta Indonesian, when it comes to the presence or absence of arguments and the flexibility of their ordering. The odds that a verb in an utterance has overt arguments are 1.97 times higher in Kupang Malay than in Jakarta Indonesian. Moreover, the percentage of utterances with a verb and just an agent argument is higher in Kupang Malay than in Jakarta Indonesian, but on the other hand, the percentage of utterances with just a verb and its patient argument is higher in Jakarta Indonesian. Seemingly opposed to that, the agents' positioning relative to the verb is more flexible in Jakarta Indonesian than in Kupang Malay, while the patients' placement is more flexible in Kupang Malay. This shows that even though Jakarta Indonesian and Kupang Malay share crucial properties, for instance being highly isolating and pre-dominantly SVO, they still allow for some flexibility in their word order. Additionally, they differ in

how flexibly their arguments can be rearranged. This can be taken as example of how quantitative studies can yield insights about exceptions to typological generalisations and eventually into how languages work.

## References

Gregor, Tina (unpublished). *Word Order in Two Malayic Languages: A Corpus Study*. MA thesis. University of Leipzig.

Paauw, S. H. (2008). *The Malay Contact Varieties of Eastern Indonesia: A Typological Comparison*. PhD thesis. New York: State University of New York.

Siewierska, Anna (1998). Variation in major constituent order; a global and a European perspective. In: Siewierska, Anna (ed.). *Constituent Order in the Languages of Europe, Empirical Approaches to language typology*. EURO-TYP volume 20. Berlin and New York: Mouton de Gruyter. pp. 475-551.

Sneddon, J. N. (2006). *Colloquial Jakartan Indonesian*. Canberra: Australian National University.