

SOUTHEAST ASIA IS A LINGUISTIC AREA - BUT HOW TYPICAL ARE SOUTHEAST ASIAN LANGUAGES?

Bernard Comrie

*Max Planck Institute for Evolutionary Anthropology and
University of California Santa Barbara
comrie@eva.mpg.de*

SEALS20 - Abstract

Earlier work making use of the *World Atlas of Language Structures*, e.g. Dahl (2008) and Comrie (2007), has strikingly confirmed the validity of Southeast Asia as a linguistic area. In particular, Comrie (2007) identifies a set of 21 features that clearly set Southeast Asia off from northern Asia, often Eurasia more generally, with a buffer zone of languages (including Mandarin and other Sinitic varieties) that sometimes go with Southeast Asia, sometimes with northern Asia.

However, it now becomes how relevant how typical Southeast Asian languages are against the background of the world's overall typological diversity: Clearly, an area that is defined in terms of feature values that are rare cross-linguistically is more striking than one defined by feature values that are common, even a majority. Ongoing research aims to address this issue, or rather to answer two more specific questions.

The first is the more general one: How typical are Southeast Asian languages? A measure is developed that enables one to calculate the relative degree of typicality of any pair of languages for which WALS provides data on a large enough number of features. (In principle, the method can be generalized to triples etc. of languages, though the number of shared WALS features drops as one increases the number of languages.) Preliminary results comparing Southeast Asian languages, East Asian languages, English, and Wari' (the last-named a known typological outlier) suggest that Southeast Asian languages are overall rather typical, at around the upper end of the typicality range of East Asian languages, with English falling at the bottom of this range and Wari' well below it. This is potentially worrying, given the higher validity of unusual feature values in defining a linguistic area.

A methodology is therefore developed that restricts itself to the 21 WALS feature values identified as characteristic of Southeast Asia, and the question is posed how typical these values are against the background of the languages, genera, and families of the world, with the genealogical classificatory level of genera probably giving the most relevant results. While some of the 21 features are indeed majority patterns across the genera of the world, 16 are not, and for more than half the feature values identified as characteristic of Southeast Asia, that feature value is found in fewer than 35% of the world's genera. Southeast Asia is thus defined as a linguistic area not only in terms of shared feature values, but also in terms of the relative rarity of those feature values.

References

- Comrie, B. 2007. Areal typology of mainland Southeast Asia: what we learn from the WALS maps. In Pranee Kullavanijaya (ed.): *Trends in Thai Linguistics* (=Manusya, Special Issue 13), 18–47. Bangkok: Chulalongkorn University.
- Dahl, Ö. 2008. An exercise in a posteriori language sampling. *STUF* 61: 208–220.