

Malay/Indonesian languages in cross-linguistic comparison: Using parallel texts and quantitative data

Gertraud Fenk-Oczlon & August Fenk, Alps-Adriatic University of Klagenfurt, Austria

In this paper we compare the Malay/Indonesian languages Bahasa Melayu and Bahasa Indonesia as well as Minangkabou and colloquial Indonesian as spoken in Jakarta with other languages of Indonesia (Javanese, Nias, Karo Batak) and with three other Malayo-Polynesian languages: Mambae and Kemak (East Timor) as well as Kadazan (Malaysia-Sabah).

For this comparison we used parallel textual material, i.e., a set of 22 simple declarative sentences encoding one proposition and using a rather basic vocabulary: e.g. *The sun is shining. I thank the teacher. Grandfather is sleeping. My father is a fisherman. The dog is outside.* Native speakers were asked to give an intralinear translation of these sentences into their mother tongue. Furthermore, they were instructed to count the number of syllables in normal speech (which is, apart from determining the borders of the syllables, no problem for the informants). The written translations allowed, moreover, counting the number of words per clause. The number of phonemes was determined by the first author, with the help of the native speakers and of grammars of the respective languages.

Using this method we not only obtained interesting qualitative data concerning lexical or grammatical information, but also quantitative data such as the average number of syllables per sentence and per word, the number of phonemes per syllable or the proportion of monosyllables. For instance: In our textual material the Malayic languages show far less monosyllables (only the prepositions *di and ka*) than the non-Malayic languages, with the exception of Javanese. The average word length as measured in syllables ranges from rather high values in Nias and Bahasa Indonesia (2.77 and 2.60 respectively) to rather low values in Kadazan (1.96) and Mambae (1.73).

The quantitative data are also discussed in the typological context of a sample of more than 50 languages (Fenk-Oczlon & Fenk, 2010). One of the relevant findings: Malay/Indonesian languages show, as compared with the majority of the languages in our sample, a relatively high number of syllables per sentence and a relatively low number of phonemes per syllable.

Fenk-Oczlon, G. & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 2010), Makuhari, Japan, 1537-1540.