

The size of syllables, words and clauses in 16 Austronesian languages: quantitative data and cross-linguistic comparisons

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

In this paper we present quantitative data on syllable complexity as well as on word and clause length in the following 16 Austronesian languages: Bahasa Indonesia, Bahasa Melayu, Cham, Chuukese, Hawaiian, Javanese, Kadazan, Kemak, Malagasy, Mambac, Minangkabau, Karo Batak, Nias, Rinconada, Roviana, Tagalog.

Syllable complexity and word length are considered as important variables in language typology. But how to compare languages with respect to syllable complexity? Should one count the maximally elaborate syllable type or the predominant syllable type in a language [1]? And how to compare word length in different languages - on the basis of lexicon entries or on the basis of textual material? And clause length? In our ongoing project [e.g.2] we are using parallel textual material for large-scale cross-linguistic comparisons concerning these variables:

Method: Native speakers were asked to give an interlinear translation of 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.* Furthermore, they were asked to count the number of syllables in normal speech. The written translations allowed, moreover, counting the number of words per clause. The number of phonemes was determined with the help of the native speakers and of grammars of the respective languages.

Some Results: In a sample of 51 languages, the mean number of syllables per clause is 7.02, ranging from 4.64 in Thai up to 10.96 in Telugu. The mean number of phonemes per syllable is 2.24, ranging from 2.79 in German to 1.76 in Hawaiian.

The Austronesian languages in our sample show a relatively large number of 8.72 syllables per clause and a relatively low number of 2.03 phonemes per syllable. Some well-known or assumed (areal-)typological differences within the group of Austronesian languages could be corroborated on quantitative grounds. For instance:

- The Oceanic languages (Chuukese, Hawaiian, Roviana) exhibit the highest average number of syllables per clause (9.59) and the lowest mean syllable complexity (1.88 phonemes/syllable).
- Cham that is said to be highly influenced by Vietnamese [3], is indeed the only language in our sample that shows a mean of only 6.32 syllables per clause (for comparison: Vietnamese 4.91) and as much as 2.37 (Vietnamese 2.24) phonemes per syllable. In the other 15 languages, syllable complexity is restricted to a small range from 1.765 in Hawaiian to 2.191 in Karo Batak (standard deviation: 0.13). In our matched textual material Cham shows a relatively high number of monosyllables (30), much more than e.g. Malya/Indonesian (3) or Nias (11).

The findings will also be related to typological differences in speech rate and speech rhythm and will be discussed, more generally, within the framework of systemic typology.

- [1] Maddieson, I. (2009). Calculating phonological complexity. In: F.Pellegrino et al. (eds.) *Approaches to complexity*. Berlin: de Gruyter, 85-109
- [2] 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
- [3] Thurgood, G. (2005). Phan Rang Cham. In: K. A. Adelaar & N. Himmelmann (eds.) *The Austronesian languages of Asia and Madagascar*. London/New York: Routledge, 489-512