Quantitative typology based on a matched set of simple declarative sentences

Gertraud Fenk-Oczlon & August Fenk, Universität Klagenfurt,
gertraud.fenk@uni-klu.ac.at

Our approach to quantitative typology makes use of a "simple" but massively parallel textual material obtained by an elicitation experiment.

The rationale for using simple textual material for cross-linguistic comparison is the following: To make large-scale cross-linguistic comparisons we need first of all textual material that is easily translatable into a large number of languages. Moreover, the number of possible translations should be kept at a minimum. Simple declarative clauses encoding one proposition and using a rather basic vocabulary meet these conditions; they also seem to be universal from a syntactic perspective (quite a number of languages use almost exclusivly a series of minimal-predications instead of more complex sentences. Cf. Sasse 1991, Heeschen 1994). Thus we constructed a set of 22 simple declarative sentences as source text. Examples for the test sentences are: *The sun is shining. I thank the teacher. Grandfather is sleeping. My father is a fisherman. Blood is red*.

Procedure: Native speakers of meanwhile 51 languages from all continents (19 European, 32 Non-Indo-European) were asked to translate the 22 sentences into their mother tongue. Furthermore, they were asked to count the number of syllables in normal speech. The written translations allowed, moreover, to count the number of words per clause. The number of phonemes was determined by the authors, assisted by the native speakers and by grammers of the respective languages.

Results: In our 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.
Further statistical analyses of our textual material revealed
significant cross-linguistic correlations between *syllable complexity* and

- number of syllables per clause and per word (negative correlations)
- number of monosyllabic words (positive correlation)
- number of cases (negative correlation)

interactions between metric and non-metric properties of language:

- associations between *syllable complexity* and morphological type (isolating,
- fusional, agglutinative)
- associations between syllable complexity and word order (OV vs. VO) as well as adposition order
- associations between intrinsic tempo (mean number of syllables per clause) and rhythm classes (stress-timed, syllable-timed, mora-timed)

Our method and material also provided quantitative data concerning some more or less hypothetical (areal-) typological differences. For instance, the 16 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, whereas the Oceanic languages (Chukeese, Hawaiian, Roviana) exhibit the highest average number of syllables per clause (9.59) and the lowest mean syllable complexity (1.88 phon/syll). Cham, which is said to be highly influenced by Vietnamese **(**Thurgood, 2005), steps out of line: It is the only Austronesian 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 1.77 in Hawaiian to 2.19 in Karo Batak (standard deviation: 0.13). In our matched textual material, Cham also exhibits a relatively high number of monosyllables (30), much more than e.g. Malay/Indonesian (3) or Nias (11).

REFERENCES

Fenk-Oczlon, G. & Fenk, A. (1999). Cognition, quantitative linguistics, and systemic typology. Linguistic Typology, 3, 151-177

Fenk-Oczlon, G. & Fenk, A. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In G. Fenk-Oczlon & Ch. Winkler (eds.) Sprache und Natürlichkeit. Gedenkband für Willi Mayerthaler, 75-86. Tübingen: Narr

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

Heeschen, V. (1994). How long are clauses and sentences in a papuan language like Eipo? Semaian, 10, 50-70

Sasse, H.-J. (1991). Predication and sentence constitution in universal perspective. In: D. Zaefferer (ed.) Semantic universals and universal semantics, 75-95. Berlin:Foris

Thurgood, G. (2005). Phan Rang Cham. In: K.A. Adelaar & N. Himmelmann (eds.) The Austronesian languages of Asia and Madagascar, 489-512. London/New York: Routledge