

Cluster analysis in a typological study of participles

In typological studies participle is usually defined as a verb form that can be used for nominal modification and that demonstrates at least some degree of deranking by losing certain verbal morphological categories and/or by changing the verbal argument encoding, cf. (Haspelmath, 1994, Kalinina, 2001; Ylikoski, 2003). However, the range of forms that comply with this definition is remarkably broad.

For example, in Lezgian, cf. (1), the participle can relativize a variety of syntactic positions. All the arguments in the relative clause are coded in the same way as in the corresponding independent clause. The participle does not demonstrate any agreement with the modified noun, and the participial marker can combine with different markers of tense and aspect.

(1) Lezgian (North Caucasian) (Haspelmath, 1994: 154)

[<i>ruš-a</i>	<i>čar</i>	<i>kxi-zwa-j</i>]	<i>stol</i>
girl-ERG	letter.ABS	write-IMPERF-PTCP	table

‘the table on which the girl is writing a letter’

On the other hand, in Finnish (Uralic), cf. (2), the *-ma-* participial form can only be used to relativize the direct object of a transitive clause. The agent of the relative clause is coded as a possessor. The participle agrees in case and number with the modified noun and, unlike finite verbs, it does not show any distinctions in tense.

(2) Finnish (Uralic)

[<i>tytö-n</i>	<i>kirjoitta-ma-ssa</i>]	<i>kirjee-ssä</i>
girl.SG-GEN	write-PTCP.PASS.SG-INE	letter.SG-INE

‘in the letter written by the girl’

Therefore, it is clear that there is no single core in the vast participial zone, but rather several clusters that should be studied separately. The goal of this paper is to reveal such typologically relevant clusters in a data-driven way. In order to do that, participles in 70 genetically and geographically diverse languages are characterized with respect to several morphological and syntactic criteria, and the resulting data is then analyzed with statistical methods dealing with distance between participial forms in different languages.

The analysis reveals a number of different-sized clusters within the participial zone and the importance of some specific criteria for clustering, for example the presence of orientation (whether the form can relativize one position or more). It is especially important that one of the most distinctive clusters comprises the forms that are referred to as relative participles (*Relativpartizipien*) in (Lehmann, 1984), and, thus, are already known to constitute a cross-

linguistically relevant type. This fact is one of the signs that the introduced method is effective and can be used in the typological study of participles and beyond.

References

Haspelmath, M. (1994). Passive participles across languages. In B. Fox & P. J. Hopper (Eds.). *Voice: Form and Function*. (Typological Studies in Language, 27, pp. 151–177) Amsterdam: Benjamins.

Kalinina, E. Yu. (2001). *Nefinitnye skazuemye v nezavisimom predložanii*. Moscow: IMLI RAN.

Lehmann, C. (1984). *Der Relativsatz*. Tübingen: Gunther Narr.

Ylikoski, J. (2003) Defining non-finites: action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16, 185–237.