

Paweł Rutkowski

**Sign language as a challenge to electronic lexicography:
The Corpus-based Dictionary of Polish Sign Language and beyond**

The main aim of this paper is to discuss to what extent corpus data can and should form the empirical basis for sign language lexicography. In this context, special attention will be paid to the *Corpus-based Dictionary of Polish Sign Language* (CDPSL), an open-access electronic publication, available freely at the following address:

www.slownikpjm.uw.edu.pl/en/

Polish Sign Language (*polski język migowy*, usually abbreviated as PJM) is a full-fledged three-dimensional language used by the Polish Deaf community. It started to evolve in the second decade of the nineteenth century, with the foundation of the first school for the deaf in Poland. Until recently, PJM attracted very little attention from the hearing linguistic community in Poland. The situation has changed in the last decade, thanks to a research project based at the University of Warsaw, aimed at creating the first-ever corpus of PJM and producing a corpus-based dictionary of that language.

Internationally, sign language lexicography has been developing quite dynamically in recent years, both in technological and theoretical terms. Sign language dictionaries are no longer glossaries based on the simplistic assumption that each and every sign will have a one-to-one lexical equivalent in the surrounding spoken language. Instead, more and more lexicographers feel that modern sign language dictionaries should describe individual signs and their different senses on the basis of documented (corpus-attested) usage. However, the new subfield of sign language lexicography has found a challenge in numerous structural differences between spoken and sign languages (the visual modality being the most obvious but not the only one). A selection of problems faced by a sign language lexicographer (and possible solutions thereof) will be discussed in this talk on the basis of how CDPSL was compiled. An overview of procedures leading to the creation of the dictionary will be presented (along with sample entries and video clips) to illustrate the most important advantages and disadvantages of the methodological choices that the CDPSL team made.