

# **The tension between language documentation and descriptive linguistics - the case of deictic motion verbs in Hocąk**

**Johannes Helmbrecht, Iren Hartmann**

## **Abstract**

It is the main goal of language documentation to create a representative corpus of texts of an individual language, which are processed in a way that they are easily accessible to non-speakers and non-specialists of the respective language. "Representative" in this context means that a selection of texts is included in the corpus covering the uses of language in the speech community. The hope which is silently implied is that such a text corpus would be representative also with respect to the language system. This, however, cannot be guaranteed. Often, the descriptive linguist has to broaden the range of linguistic data he/she has at hand from the corpus by various types of elicitation techniques applying his/her knowledge on linguistic notions and their typological variation and elaborating on categories, paradigms and constructions already found in the target language.

It is the goal of the proposed paper to present a study on deictic motion verbs in Hocąk (Sioux) that illustrates at the same time this type of tension between language documentation and descriptive linguistics. Some practical ways out of this dilemma will be presented as conclusions.

1) The proposed paper starts with a survey of the occurrences of the Hocąk deictic motion verbs in the corpus and their contexts (application of corpus linguistic methods). It will be demonstrated that the systematic semantic nature of these lexical items shows up only after a typologically informed and thorough analysis of the contexts and additional elicitation of data (substitution test, psycholinguistic tests). The reasons for the lack of data in the corpus will be specified.

2) Then, the system of the Hocąk deictic motion verbs will be presented. The 12 Hocąk deictic motion verbs are comparable to and are often translated simply as *come* and *go* in English, but they have a much more complex inherent semantics. The first distinction is between motion toward or not toward the deictic center; the second distinction encoded is between motion toward a base/home versus not toward base/home; and the third distinction is between the aktionsarten departure, on the way/ in progress, and arrival.

3) It will be shown that the inherent semantics explains in part the different frequencies of their respective occurrence in the corpus and the ability to combine with other verbs in complex verbal predicates.

The main conclusion to be drawn from the study of deictic motion verbs in Hocąk is that there is no such thing as a pure principled documentation of an individual language, but that language documentation and language description go hand in hand to some degree. Good language documentation needs reference to descriptive linguistics.