

The role of systematic sampling procedures in language documentation

Frank Seifart
Ruhr-Universität Bochum,
“People of the Center” DobeS project
Frank.Seifart@berlin.de

This paper explores the issue of the selection (or “sampling”) of the communicative events that are recorded and annotated in order to be included in a language documentation. It is intended as a contribution to the further development of the theoretical framework of documentary linguistics, which has so far paid much more attention to the format, rather than to the contents of documentations. Sampling is understood in this context as a technique to ensure the representativeness and completeness of the contents of a documentation. The principles and criteria that may underlie such sampling are discussed, using examples from the work of the DobeS team with the so-called People of the Center, a multilingual cultural complex in the North West Amazon. I explore the possibilities and limitations of “systematic sampling” by criteria that derive from properties of the communicative event types themselves, e.g. participant structure, language code, or frequency. I also discuss the relation of such sampling methods to others, such as “externally informed sampling”, e.g. according to the scientific interest in data from particular speech situations, as well as a number of practical and ethical issues.