

《International Workshop 1》

Word Order

National Museum of Ethnology, Osaka, Japan

September 27, 2013

Abstract

Fieldwork Method: How to Elicit Spoken Language Sentence Data

INAGAKI Kazuya

Kyoto University

This presentation introduces some selected topics in sentence data collection carried out in the course of fieldwork on the spoken language. It addresses four issues: language use, aims and methods of data collection, speakers, and word order. When sentences are uttered, the researcher is flooded with a lot of different elements such as pronunciation simplification, special intonation and pauses, or honorific expressions, all of them co-occurring simultaneously. Effective data collection requires a conscious effort to exclude as many of the above items as possible. In sentence data collection, it is important to refer to a corpus as well as to ask the speaker to judge artificially construed sentences. We must recognize the advantages and disadvantages of each method, and work to collect balanced data, compensating for each of the disadvantages. In order to elicit sentence data, we will have to train the speaker helping her to understand our intention sufficiently. What we should keep in mind is that the speaker does not always provide an optimal answer which the researcher wants to obtain by showing a translation or an artificial sentence. Elicited sentence data alone will not be sufficient for describing a language. We need to further enhance its reliability by other means, such as utilizing corpus data.

《国際ワークショップ1》

語順について考える

国立民族学博物館, 大阪

2013年9月27日

講演要旨

「フィールドワークの方法—音声言語の文のデータをどのように収集するのか」

稲垣和也

京都大学

本発表では、音声言語のフィールドワークでおこなわれる文データ収集についていくつかのトピックを選んで紹介する。言語使用、データ収集の目的と方法、話者、語順について言及する。文が産出される時、発音の簡略化、特別な抑揚や休止、敬語といった多くの様々な要素が同時にあらわれる。効率的にデータを収集するにはこれらの要素をなるべく除外する意識的な努力が必要になる。文データの収集では、コーパスを参照することも、人工的な文を話者に判断させることも重要である。それぞれの方法の利点と欠点を正しく認識し、それぞれの欠点を補いながら、バランスの良いデータの収集に努めなければならない。また、文データを引き出すためには、話者が調査者の意図を十分にくみ取れるように訓練しなければならない。念頭に置いておくべきことは、話者は、翻訳文や人工文を提示すれば調査者が求める最適な答を常に提示してくれるわけではないということである。言語を記述するには、話者から引き出す文データだけでは十分ではない。コーパスデータを利用するなど、他の手段によってデータの信頼性をさらに高める必要がある。