

Cristina Petraş
Alexandru Ioan Cuza University of Iaşi

Corpus Linguistics: Epistemological and Methodological Issues in Language Studies

Joint Master in European Languages, Cultures and Societies in Contact

I. Introduction

This lecture proposes an introduction to the field linguistics, which involves a deep reflection about how linguists work. In linguistics, one of the main issues is the relationship between type of data and theorization, with its epistemological consequences. In *traditional grammar* introspection data correspond to a hypothetical-deductive approach, while in *field linguistics (or corpus linguistics)* observational data involves an empirical-inductive approach (empirical investigations).

II. Epistemological issues

III. Methods of data collection

- language survey (questionnaires, interviews), participant observation, recording groups in interaction (ecological data - gathered in their environment - dinner, consultation, market, hotel, shop, phone, so on)

IV. Corpus

IV.1. A corpus is a construct

IV.2. Written corpora /vs/ Spoken corpora

IV.3. Corpus transcription. Corpus annotation

IV.4. What the corpus is used for ("corpus based /vs/ "corpus driven")

IV.5. Monitor corpora /vs/ balanced corpora /vs/ opportunistic corpora

IV.6. How to work on corpora?

IV.6.1. What type of corpus for what linguistic features? What kind of corpus approach (corpus-based or corpus-driven)

IV.6.2. Working on comparable corpora