

Problems in selecting transcription system for spoken discourse: a case of stand-up comedy.

Key words: oral discourse, transcription, stand-up comedy.

Textual analysis of oral data presents a certain degree of difficulty for researchers due to its technical aspect, namely, the process of transcribing the data. But for the fact that the process itself is known to be quite exhausting and monotonous, there is no ready-made system of transcription to suit the needs of every scientific case. In fact, transcription systems for spoken discourse are presented in a wide variety. In this paper, I intend to discuss some of the difficulties that may occur in the process of selecting a model for transcription of texts. The reference of Gibbon (1998), who gives particular emphasis to the purpose of the study while determining the type of transcription to use, is taken as a starting point of the paper. In particular, this paper looks at several commonly recognized transcription systems for conducting research in social and human sciences. One of them is Discourse Transcription (DT) – a system described by John W. Du Bois (University of California, Santa Barbara). Later developed into DT2 (2006), it provides guidance in the transcription process. Brian MacWhinney (Carnegie Mellon University, Pittsburgh) is known for his CHAT Transcription Format (2000), which is used as a tool for analyzing baby talk within the CHILDES Project (Child Data Language Exchange System) and which was further adopted as a transcription model for the multilingual C-ORAL-ROM Project presented in partial collaboration with CLUL, Faculdade de Letras, UL. The above-mentioned models are to be compared in order to reveal their applicability to the transcription of oral data taken from stand-up comedy (henceforth SUC) shows in the USA and in Portugal. The oral data in conjunction present the corpus of the current PhD project. On the basis of the results of the analysis of various transcription systems, a special transcription model will be designed taking into considerations the following delimiting parameters: 1) the new transcription must satisfy the needs of the current study, i.e. bear in mind the objectives of the analysis; 2) also, it must take into account the linguistic units which serve as central focus. Considering the current stage of the investigation, adaptation of the existing above-mentioned systems is expected due to several factors: (a) the outlined systems are not oriented towards analyzing humoristic discourse of SUC with its peculiarities (a lone comedian standing on the stage as the interlocutor, extensive use of audible pauses, audience's laughter in response, etc.); (b) the systems are overdetailed, i.e. they give too much importance to various aspects of language at the same time, for instance to the phonetic aspect (dialectal variations, aspiration, shortenings), including pauses and timing (e.g. detailed description of tones, millisecond/quarter second/half second pauses in speech). All these extra features do not present any value for the current study. Thus, the systems need to be adapted. The result of this work will be presented in the paper.