

Real language: A Corpus study of Hip Hop language.

For a corpus to be of any real value to corpus linguistic analysis, it should consist of machine-readable text. Thanks to widespread use of computers and word-processors, most written texts nowadays are machine-readable. Unfortunately, the collection of spoken language in corpus format remains problematic. With the advent of the Internet, however, a variety of machine-readable texts are available which constitute a new genre of discourse, arguably similar to spoken language, but in written form, such as chats or e-mails. A particular example of this genre, what Crystal (2001) calls the "third medium", is the text of message boards. Message board entries reflect a tendency among contributors to write as they speak. It can therefore be considered a variant of spoken language, but written by the 'speakers' themselves as they choose to represent their 'speech.' The advantages to using this type of language in corpus studies are two-fold: not only are corpus linguists freed from the task of transcription and annotation of recorded speech, but they also have access to speech-like language in machine-readable format.

In this poster presentation, I present an analysis of a 100,000 word corpus consisting of message board entries on hip-hop music and culture. This sample corpus clearly shows that the contributors write in a manner characteristic of spoken language, but that they also exploit the medium of the message board to further identify themselves linguistically in ways which are not possible in speech. This corpus serves to exemplify the new wave of "real language" which corpus linguistics has yet to account for. Furthermore, the analysis addresses the problem of knowing "real speakers" by considering speaker identity not as a combination of sociolinguistic variables, but as a linguistic construction of self.

Select bibliography:

- Crystal, D. 2001. *Language and the Internet*. Cambridge: CUP.
George, N. 1999. *Hip Hop America*. Penguin.
Suler, J. 2004. *CyberPsychology and Behavior*, 7. 321-326.