

Filled Pauses in Writing: What can they Teach us about Speech?

Ralph ROSE <rose@waseda.jp>

Waseda University Faculty of Science and Engineering (Tokyo, Japan)

This presentation reports on a research effort to use filled pauses ('uh', 'um': hereafter, FPs) in blog writings to better understand how and why speakers use them in spontaneous speech. Blog FPs are written intentionally and cannot be the result of some linguistic processing shortcoming (i.e., speech-repair as in Levelt, 1983). Hence, if written FPs can be accurately characterized, then the spoken FPs that fit this characterization can be removed from consideration leaving a smaller, potentially more uniform set of other FPs for further study.

Samples of FPs in blog writings were gathered from 100 top blogs. Samples of FPs in spontaneous speech were taken from the Switchboard corpus. A balanced sample of 227 FPs were gathered of each type. Each FP was categorized according to its medium (written or spoken), its location (at clause boundary or clause-internal), the part-of-speech of the immediately following word (content or function, following Maclay and Osgood's 1959 classification), and the FP type (open 'uh' or closed 'um', after Rose, 1998). The data was analyzed under a generalized linear model with chi-square tests.

There was a main effect of FP Type (Chi-square=48.4, $p < 0.001$) with a ratio of open to closed FPs of approximately 2:1. This is comparable to previous studies (e.g., Rose, 1998). There were no other main effects. There was an interaction between medium and following word type (Chi-square=37.0, $p < 0.001$), as well as between medium and FP type (Chi-square=5.4, $p < 0.05$). In the spoken medium, the following word was 30% more likely to be a function word than a content word, while in the written medium, this trend reversed: the following word was 70% more likely to be a content word than a function word. Also, in the spoken medium, the ratio of open to closed FPs was almost 3:1, but in the written medium, this ratio dropped to 1.4:1.

Results from FPs in writing suggest a hybrid view of FPs in speech: Some FPs are used intentionally and with some selectional restrictions (i.e., before content words) in order to serve some pragmatic function (cf., filler-as-word hypothesis in Clark and Fox Tree, 2002), with open FPs being slightly preferred in this role. Other FPs in speech are the result of difficulties during linguistic processing and occur semi-automatically as part of speech repair (cf., Levelt, 1983).

References

- Levelt, W.J.M., 1983. Monitoring and self-repair in speech. *Cognition*, 14: 41-104.
- Maclay, H. and Osgood, C.E., 1959. Hesitation phenomena in spontaneous English speech. *Word*, 15: 19-44.
- Rose, R.L., 1998. The communicative value of filled pauses in spontaneous speech. Master's Dissertation, University of Birmingham, UK.
- Clark, H., & Fox Tree, J. E., 2002. Using uh and um in spontaneous speaking. *Cognition*, 84:73-111.