

The Effect of the ‘Telephone Situation’ on Formant Frequencies

Julia Fischer-Weppler¹, Michael Jessen², and Florian Schiel¹

¹Institute of Phonetics and Speech Processing, Ludwig-Maximilians-Universität, München, Germany

julia.fw@gmx.de

schiel@phonetik.uni-muenchen.de

²Department of Speaker Identification and Audio Analysis, Bundeskriminalamt, Germany

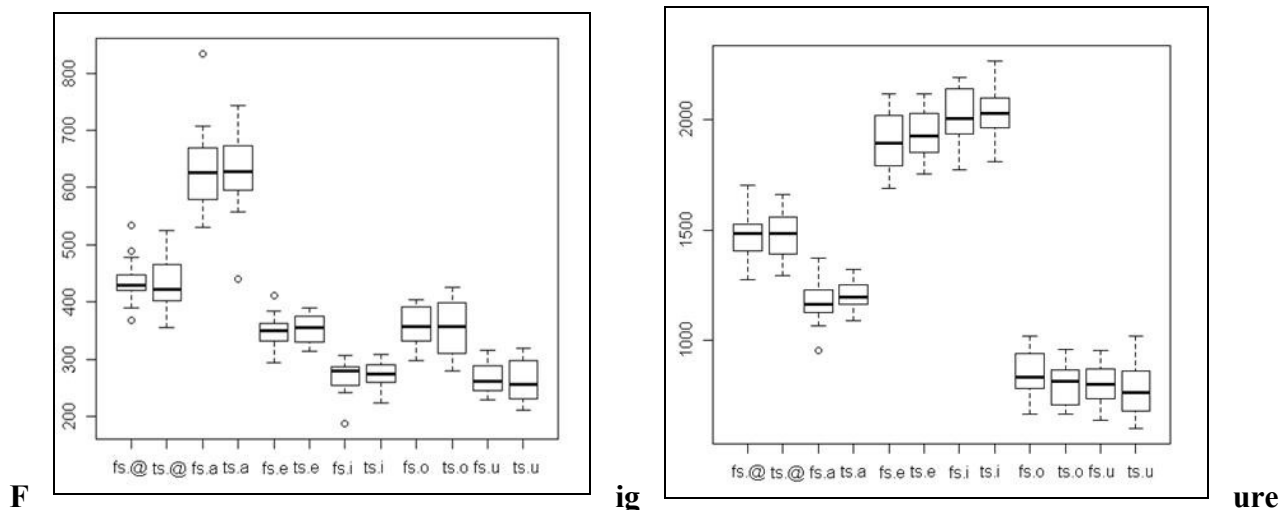
michael.jessen@bka.bund.de

In the past years forensic research on the technical effects of telephone transmission on the speech signal has been conducted (see e.g. Künzel 2001). What has not been taken into account so far is the contextual effect of the telephone situation on the manner of speaking. More specifically: do speakers adapt to the perceived band limitation of the telephone transmission by changing their own average formant frequencies in order to enhance the intelligibility? If this would be the case, it would be interesting for the forensic practise to know about the extend and direction of such adaptive behaviour and possible dependencies to the average format level of the speaker.

To determine the effect of a telephone communication channel on the speaker’s formant frequencies an experiment using data from the ‘Pool 2010’ corpus was conducted. The data were recorded by the Department of Speaker Identification and Audio Analysis (Jessen et al. 2005).

The aim of the experiment was to investigate if there are differences in the formant structure of free speech vs. telephone speech (in this case: speech recorded directly from the speaker while talking on the phone) which indicate an adjustment of the formant frequencies to the lower or higher cut off region of the telephone bandpass (300 - 3400Hz).

18 speakers (all male) were chosen from the corpus with respect to their average formant height so that 9 speakers show distinctive low and the remaining 9 high average formant frequencies. The selected recordings of these speakers were all semi-spontaneous speech in the conditions ‘free speech’ (fs) and ‘telephone speech’ (ts). Measurements of formants F1 and F2 were conducted on six vowel categories (/ə/, /a/, /e/, /i/, /o/ and /u/) using the SNACK formant tracker in WaveSurfer.



1. Mean F1 (left) and mean F2 (right) values of all speakers for the analyzed vowel categories in both recording conditions (fs and ts).

Repeated-measures ANOVA did not yield any significant changes for F1 or F2 regarding the communication situation. Insofar our hypothesis has to be rejected.

However the pattern of the differences in formant frequencies of free and telephone speech indicate that formant values tend to move into more extreme positions, with other words that the speaker hyper-articulates his vowels (see figure). Although these tendencies are not significant on our data set (which consists only of 18 speakers) it might nevertheless prove to be of relevance in a larger dataset and therefore to be considered for forensics.

References

- Jessen, M., Köster, O. and S. Gfroerer. (2005). Influence of vocal effort on average and variability of fundamental frequency. *Speech, Language and the Law*, **12**, 174-213.
- Künzel, H. J.. (2001). Beware of the 'telephone effect': The influence of telephone transmission on the measurement of formants frequencies. *Forensic Linguistics*, **8(1)**, 80-99.