

# The effects of voice disguise on f0 and on the formants

*Ingrid Hove and Volker Dellwo*

*Phonetics Laboratory, Department of Comparative Linguistics, University of Zurich, Switzerland*

ingrid.hove@uzh.ch; volker.dellwo@uzh.ch

Voice disguise is a serious problem for forensic speaker identification. In order to help provide solutions to deal with disguised (or possibly disguised) voices we aim at finding out which acoustic characteristics change and which remain consistent in different disguise conditions. For the characteristics which are affected by voice disguise, the aim is to find out whether these changes are systematic, i.e. whether they always go in the same direction for a specific disguise condition or not.

Previous research in the temporal domain has shown that certain durational characteristics are idiosyncratic throughout different disguise conditions (Hove & Dellwo 2012). In the present study, we focus on effects of different kinds of voice disguise in typical frequency-domain based speaker specific characteristics like average fundamental frequency and formants.

The corpus we recorded contains read speech of 12 speakers of Zurich German. Every speaker reads 24 translated sentences from the Bamford-Kowal-Bench corpus (Bench/Kowal/Bamford 1979), 12 sequences of nonsense words of the type CV<sup>1</sup>CV<sup>2</sup>CV<sup>3</sup> with each word pronounced three times, plus a well-known Swiss German nursery rhyme. The disguise conditions are two types of prosodic modification, namely high-pitched voice and low-pitched voice, as well as four types of articulatory obstruction: speaking with a pencil in the mouth, speaking with a lollipop in the mouth, speaking with a pinched nose and speaking with a hand in front of the mouth. For all sentences we compare the fundamental frequency (mean pitch), the standard deviation and the minimum and maximum of the fundamental frequency.

For the prosodic modification conditions, first results show that the speakers differ in their ability to modify their pitch: when speaking in a high- or low-pitched voice certain speakers succeed well in raising or lowering their voice whereas others only show small differences to their normal speaking voice. We are also looking at how consistently the speakers can keep up this modified pitch. Furthermore, the effect of the disguise conditions on the formants will be examined. For this part of the analysis, the focus will lie on the nonsense words of the type [pa:pi:pu:] or [xi:xy:xu:].

The comparison of our results to the findings of other studies on voice disguise such as Künzel (2000), Perrot et al. (2007), Moosmüller (2001), Eriksson and Wretling (1997), or Masthoff (1996) should expand our understanding of the effects of voice disguise on the speech.

## References

- Bench J., A. Kowal and J. Bamford. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Brit J Audiol.* 13, 108–112.
- Boersma, P. and D. Weenink. Praat: Doing phonetics with computers, [www.praat.org](http://www.praat.org), accessed 31 Mar 2014.
- Eriksson, A. and P. Wretling. (1997). "How flexible is the human voice? – A case study of mimicry", *Eurospeech 97*(2), 1043-1046.
- Hove, I. and V. Dellwo. (2012). *The effect of articulatory obstruction on temporal characteristics of speech*. Abstract presented at IAFPA 2012, Santander/Spain.
- Künzel, H. (2000). "Effects of voice disguise on speaking fundamental frequency", *Forensic Linguistics*,

7(2):150- 179, 2000

Masthoff, H. (1996). "A report on a voice disguise experiment", *Forensic Linguistics*, 3(1), 160-167.

Moosmüller, S. (2001). "The influence of creaky voice on formant frequency changes", *The International Journal of Speech, Language and the Law*, 8/1, 100-112.

Perrot, P., G. Aversano & G. Chollet (2007). "Voice Disguise and Automatic Detection: Review and Perspectives", Y. Stylianou, M. Faundez-Zanuy, A. Esposito (Eds.): *WNSP 2005, LNCS 4391*, Berlin: Springer, pp. 101 – 117.