

## **Fundamental Frequency and the Alternative Baseline in Forensic Speaker Identification**

*Jonas Lindh*

*Department of Linguistics, Göteborg University, Gothenburg, Sweden*

*jonas.lindh@ling.gu.se*

In forensic speaker recognition it is desirable to both decide on a reliable long time measure for fundamental frequency as well as a suitable way to model intonation. One measure tested here is based on a well-known observation in various types of motor activity. In human gestures and movements there seem to be a point of departure, a resting position or baseline. It seems reasonable to assume that something similar should be true for vocal fold vibrations resulting in a neutral mode and frequency of vibration to which they return after prosodic or other types of excursions in frequency. There is currently no solid physiological evidence to support this assumption but indirect evidence has been found in studies of vocal behavior within the modulation theory of speech (MTS) where empirical evidence supporting the concept of a base line for fundamental frequency excursions has been found (Traunmüller, 1994, 2000). If the above assumptions are correct, then the fundamental frequency base line (Fb) should have a fair chance of being more robust than other measures since it better represents the neutral articulation of a given individual. In order to test this hypothesis three experiments were performed. In one experiment the speech style was varied to simulate different emotional states. In a second experiment the same message was recorded from different channels (cassette tape, cell phone, etc.) and in the third experiment the source of variation was the speaker's vocal effort. In all three experimental conditions various long-term fundamental frequency measures were tested and compared. The material in each of the experimental conditions contained one sample regarded as neutral with respect to the variation in question (neutral emotional expression, best quality channel, and neutral vocal effort respectively). The fundamental frequency measures obtained from these neutral samples were considered to be the most representative for the speakers in question. Robustness was then defined as the degree to which a given measure would yield the same results as those obtained for the neutral samples when conditions were varied. An ideally robust method should produce identical results regardless of variations in emotional expression, channel of transmission or vocal effort used by the speaker. Based on the results from these tests we would like to suggest a measure we call the alternative fundamental frequency baseline as the measure that is most robust with respect to the above-mentioned sources of variation. The alternative baseline is defined as the level below which 7.64% of the F0 values fall. In theory the baseline and alternative baseline should be identical and for high quality recordings the difference is negligible. But for poorer recordings the difference may be quite substantial as outliers affect the mean and standard deviation substantially.

### **Inter- and Intraspeaker Variation of the F0 Alternative Baseline**

To become a valuable parameter in forensic speaker identification, the parameter should show low intraspeaker and high interspeaker variation. The previously suggested measures of mean and standard deviation (Nolan, 1983, Rose, 2002) are unstable measures with high intraspeaker variation. To investigate the variation of the baseline the Swedia database data used in Lindh (2006) for young Swedish males were used. The tests mentioned above together with intraspeaker comparisons show low variation for the alternative baseline and slightly lower interspeaker variation than the median.

### **Conclusions**

The mean and the standard deviation of the fundamental frequency are not sufficiently robust measures for forensic phonetic case work. The baseline, representing a neutral mode of articulation for an individual speaker, is a much more robust measure. The alternative baseline should be preferred since it is less sensitive to outliers. However, the baseline is affected by variation in vocal

effort and is therefore not a reliable measure when the questioned and known samples differ in this respect. If a given baseline value is close to the mean of a representative population, however, the comparisons will not yield high likelihood ratios.

## **References**

- Lindh, J. (2006). Preliminary Descriptive F0-statistics for Young Male Speakers, *Papers from FONETIK 2006*, Working Papers, 52, Department of Linguistics and Phonetics, Lund University, 89-92.
- Nolan, F. (1983) *The Phonetic Bases of Speaker Recognition*, Cambridge: Cambridge University Press.
- Rose, P. (2002) *Forensic Speaker Identification*. New York, Taylor & Francis.
- Traunmüller, H. Conventional, biological, and environmental factors in speech communication: A modulation theory, *Phonetica*, 51, 170–183, 1994.
- Traunmüller, H. The Modulation Theory of Speech. <http://www.ling.su.se/staff/hartmut/module.htm>, Last modified 2000.