

## The reliability of formant measurements in high quality audio data: The effect of agreeing measurement procedures

Martin Duckworth<sup>1</sup>, Kirsty McDougall<sup>2</sup>, Gea de Jong<sup>2</sup> and Linda Shockey<sup>3</sup>

<sup>1</sup>The College of St Mark & St John, Plymouth, UK.

mduckworth@marjon.ac.uk

<sup>2</sup>Phonetic Laboratory, Department of Linguistics, University of Cambridge, UK.

{kem37|gd288}@cam.ac.uk

<sup>3</sup>Department of Applied Linguistics, University of Reading, UK.

l.shockey@reading.ac.uk

This research examines the inter-analyst robustness of formant frequency (F) measurements and presents some proposals for agreeing a procedure for the extraction of F values. Harrison (2004) outlines a number of the issues which need to be taken into account in extracting F values using currently available software tools. These include the settings used to adjust the algorithms within the program. Harrison concludes that, when deriving F values from the software program, analysts should be guided by the visible relationship between the on-screen trace of the extracted frequency and the spectrogram. In these programs there are also manual options for extracting F values such as using the cursor to make on-screen measurements of mid formant frequency from the spectrogram or a spectral slice. In addition, for both software and manually extracted F values there are decisions to be made about where in a vowel the measurements will be made. There are therefore a number of potential sources of intra- and inter-analyst variability.

In the present study three analysts used the latest available version of *Praat* to extract F1, F2 and F3 values from the same sets of high quality audio recordings of six repetitions of read utterances containing six target monophthongs, /i: u: a: ɔ: ʊ/. The subjects were two sets of twenty British males with a similar accent profile. The recordings were from the Dynamic Variability in Speech (DyViS) corpus collected by the University of Cambridge [UK ESRC RES-000-23-1248].

The first phase of the study examined the formant values extracted by three independent analysts from the same set of English monophthongs. In this phase decisions about how to extract these were left to the individuals. The second phase examined the effect of agreeing the general procedure for identifying the location and method of extracting the values. There still remained some choices for the individual analysts in relation to software settings and the exact location of the sample selected: for each speaker and vowel: setting the correct number of poles within a certain formant frequency range is crucial to ensure reliable formant tracking (Vallabha & Tuller 2002).

Results are presented which show the effect of analyst, methodology, vowel, and speaker upon the similarity of the values extracted.

We will also comment upon some of the issues encountered such as the difficulties of interpreting some spectrographic data and the problems in using some of the options in *Praat*.

With acknowledgements to Francis Nolan, Toby Hudson and Geoff Potter Department of Linguistics, University of Cambridge UK

This research was supported by an IAFPA research grant.

### References

- Harrison, P. (2004). *Variability of Formant Measurements*. Unpublished MA Thesis, Department of Language and Linguistic Science, University of York, UK.
- Vallabha, G. K. & Tuller, B. (2002) 'Systematic errors in the formant analysis of steady-state vowels', *Speech Communication* 38: 141-160.