

An investigation of cross-language differences in pitch range for speakers of English and German

Ineke Mennen*, Felix Schaeffler#, & Gerard Docherty^

*School of Linguistics, University of Bangor, UK; els803@bangor.ac.uk

#Speech Science Research Centre, Queen Margaret University, UK; fschaeffler@qmuc.ac.uk

^School of Education Communication & Language Sciences, Newcastle University, UK;
g.j.docherty@ncl.ac.uk

A key issue for the Laboratory Phonology community is to understand along which phonetic dimensions languages can systematically differ, and how such differences are actually generated in speech production, evaluated in speech perception, and acquired by those learning to speak a language for the first time.

While there has been a good deal of effort devoted to understanding the intonational structure of speech and its relationship to the form and timing of salient aspects of an f_0 contour across different languages (giving rise, of course, to many interesting discussions about the what those salient aspects might be), there has been much less focus on pitch range as a source of cross-language differences. For analysis purposes, pitch range can be divided into two sub-characteristics of a speaker's performance; pitch level and pitch span. While there are clearly organic and physiological factors which govern inter-individual variation in both of these quantities (e.g. greater vocal fold mass would, all else being equal, normally be associated with relatively low pitch levels), there is also some evidence that when groups of speakers of different languages are compared there can be a significant difference in aspects of pitch range notwithstanding a large degree of overlap in the range deployed by many speakers of both languages (i.e. a difference which is a characteristic of the collective but not *necessarily* of individual speakers).

Two languages where this may well be the case are Southern Standard British English (SSBE) and Northern Standard German (NSG). For example, there is strong anecdotal evidence that people perceive differences between speakers of English and German - with English sounding higher and having more pitch variation than German. British voices (especially female) are often perceived stereotypically as "over-excited" or even "aggressive" by German listeners. Conversely, to British listeners, German low-pitched voices may be evaluated as sounding "bored" or "unfriendly".

The aim of this study is to examine the extent to which this putative difference can be validated by a quantitative analysis of the pitch range of speakers of these two languages and, if it can be, to consider the implications for our understanding of language-specific phonetic specification.

There are two primary dimensions to our investigation. The first is to evaluate which measures are best suited to capture any cross-language differences in pitch range. The second is to use a range of these measures to attempt to identify the nature of the cross-language differences in performance which underpin the differences that people perceive. In pursuit of this second objective, we have carried out perceptual experimentation in order to verify which aspects of pitch range variation lead listeners to be more likely to identify a stimulus as being produced by a German speaker than by an English one.

Our experimental work is centred on the performance of a range of materials by 60 speakers, 30 for each of SSBE and NSG, all female university students in their twenties and thirties and functionally monolingual (i.e. no more than moderately proficient in another language). 30 participants took part in the perceptual experiments in which they were presented with

resynthesised pitch contours from the production study and asked to judge how German or English the speaker sounds.

In this paper we present results from both aspects of this study. We first focus on a comparison of the findings from various measures derived from the raw f0 time series extracted from a stretch of speech and corrected for tracking errors (e.g. measures of mean/median f0, long term distributional measures such as max-min f0, 90% range) and measures derived from specific linguistically-relevant turning points in the f0 contour (as suggested in previous work by Patterson (2000) which in turn built on work by Ladd & Terken (1995) and Shriberg et al. (1996). In presenting the findings of our perception study, we highlight the correlation between listener judgements and our different pitch range measures.

In discussing our results, we highlight the advantages and disadvantages of the two types of measures for cross-language comparisons., we consider a number of factors which might account for listeners' sensitivity to global pitch range differences but which have not yet been embedded within our experiments (such as the time that speakers spend near the top or bottom of their range), and we provide a preliminary account of how such cross-language differences might come to exist in the first place and how they may be propagated across generations.

Acknowledgements

Project funded by UK Economic & Social Research Council [R000221858]. We thank Frank Kügler and his colleagues for the collection of the German data.

References

- Ladd, D. R., Terken, J. 1995. Modelling Intra- and Inter-Speaker Pitch Range Variation. In *Proceedings of the 13th International Congress of Phonetic Sciences* (pp. 386-389). Stockholm, Sweden.
- Patterson, D. (2000). *A linguistic approach to pitch range modelling*. PhD dissertation, University of Edinburgh.
- Shriberg, E., Ladd, D.R., Terken, J., & Stolcke, A. (1996). Modeling pitch range variation within and across speakers: predicting F0 targets when "speaking up". In *Proceedings for the International Conference on Spoken Language Processing* (Addendum, 1-4). Philadelphia, PA.