PITCHING IT DIFFERENTLY: A COMPARISON OF THE PITCH RANGES OF GERMAN AND ENGLISH SPEAKERS

Ineke Mennen, Felix Schaeffler* & Gerard Docherty†*

*Speech Science Research Centre, Queen Margaret University, Edinburgh, UK
†School of Education Communication and Language Sciences, Newcastle University, UK
imennen@qmu.ac.uk; fschaeffler@qmu.ac.uk; g.j.docherty@ncl.ac.uk

ABSTRACT

This paper presents preliminary findings of a large-scale systematic comparison of various measures of pitch range for female speakers of Southern Standard British English (SSBE) and Northern Standard German (NSG). The purpose of the study as a whole is to develop the methodology to allow comparisons of pitch range across languages and regional accents, and to determine how they correlate with listeners' perceptual sensitivity to cross-language/accent differences.

In this paper we report on how four measures of pitch range in read speech (text, sentences) compare across the two groups of female speakers. Preliminary results show that the measures of the difference between the 90th and 10th percentile (in semitones), and +/- 2 standard deviations around the mean in ST differentiate the groups of speakers in the direction predicted by the stereotypical beliefs described in the literature about German and English speakers. Furthermore, these differences are most obvious in the read text and longer sentences and the effect disappears in sentences of a short duration.

Keywords: pitch range, cross-language, German, British English.

1. INTRODUCTION

The importance of pitch variation in communication is well documented; it conveys linguistic information, but also provides the listener with information about the speaker, including aspects of their identity (gender, age), affective state, social status, and social roles (e.g. [18]), and it is also a differentiator of accents (e.g. [10]). While it is well known that languages may have different intonation, there is growing evidence that languages also differ in aspects of the realisation of intonational patterns, such as global pitch range, where – in the absence of anatomical/physiological differences – groups of speakers of one language have a significantly different pitch range than speakers of another (e.g. [4;5;18]).

Despite the fact that within the last decade many cross-language and cross-regional studies of intonation have appeared (e.g. [9;10;17]), there are surprisingly few systematic comparisons of pitch range between speakers of different languages. This study sets out to fill this void by systematically comparing the pitch ranges used by groups of speakers of German and English.

2. BACKGROUND

2.1. Pitch range in German and English

There is no doubt that people are sensitive to overall differences in pitch range characterising the collective performance of speakers of a range of languages. There is strong anecdotal evidence that people perceive differences between speakers of the languages under investigation in this study – English and German – with English sounding higher and having more pitch variation than German. British voices (especially female) are often perceived as "over-excited" [6] or even "aggressive" [8] by German listeners. Conversely, to British listeners, German low-pitched voices may sound "bored" or "unfriendly" [8]. This belief has even found its way into the German film industry, which – despite a need to match the voices of the dubbing actors to the original ones – goes out of its way to use German dubbing actresses with a lower pitch and narrower pitch range than those of original English actresses to avoid this stereotyping [6].

2.2. Measuring pitch range

Pitch range is methodologically difficult to quantify, and this might be why there are so few systematic cross-language comparisons. Pitch range

can be analysed as varying along two partially independent dimensions, *level* and *span*. Level (or register) reflects the overall pitch height and span reflects how much pitch varies within a given speech sample. There appears to be no consensus as to what constitutes pitch range in previous cross-language comparisons (and indeed in general pitch range research) with a wide variety of measures being used, and studies often fail to control for factors influencing f0 (e.g. age, regional accent, type of speech materials), making it impossible to tease out the influence of the language or culture itself.

The work presented here forms part of a larger study which aims to develop the methodology to investigate the nature of variability in pitch level and/or span across speakers of different languages. As a first step we are testing several long term distributional (LTD) measures of pitch range under strictly controlled conditions in a relatively small group of German and English speakers.

3. METHOD

3.1. Materials

Two types of materials were devised; a set of 25 sentences and a text. The sentences were statements which varied in length and number of accents. We had four types of sentences (loosely based on the corpus described in [12]), embedded in short dialogues (in the following examples, italics indicate the sentences under investigation; capitals indicate the accented words): short single-accented sentences (e.g. Where were you on holidays? *In MALAGA*); long single-accented sentences (e.g. Would you like to go to the Lake District this weekend? No, *but we could go to the LOWLANDS some time*); sentences containing two accented noun phrases (*We'll order YELLOW LILIES in WOODEN VASES*); and sentences containing three accented noun phrases (*The ORANGE BLINDS with the YELLOW BORDERS will go to the lower DINING ROOM*). Care was taken to construct sentences with a similar rhythmical structure across the two languages. For the English text, we used the 'Dog and Duck' story [2]. This text was translated and slightly adapted for German.

3.2. Speakers and procedure

The English and German sentences and text were read by eight speakers of SSBE and eight speakers of NSG, respectively. These varieties were chosen as they are most likely to be the varieties which are the focus of the stereotype held by English and German speakers. Speakers were all functionally monolingual female university students in their twenties and thirties.

Speakers were first asked to read a word list in their respective language, which was used to verify whether they indeed spoke SSBE or NSG. They were then instructed to read the sentences and text as naturally as possible and to repeat any misread sentence. The experimenter monitored this and occasionally speakers were asked to repeat a sentence.

The English material was recorded in a sound proofed room using a Marantz flash recorder and an AKG condenser microphone. The German recordings were performed under similar conditions with a Tascam DAT-recorder and an Audio-Technica condenser microphone. The test materials were digitised at 44.1 kHz sampling rate.

3.3. Measurements

The first repetition without dysfluencies, noise or inappropriate phrase boundaries was selected for further analysis.

F0 was measured with Praat [1], using the analysis settings for female voices as recommended in the Praat manual (pitch floor 100 Hz, pitch ceiling 500 Hz).

For span, we measured the difference between the 90^{th} and 10^{th} percentile range (80% range) in semitones (ST), interquartile range (IQR) in ST, and +/- 2 standard deviations around the mean (SD4) in ST. For level, mean f0 (Hz) was measured.

4. RESULTS

4.1. Text

For each of the dependent variables (80% range, interquartile range, and SD4) we ran a separate t-test. As we tested three dependent variables from the same sample for span, the α-level for each dependent variable was set at .016 to guarantee an overall Type I error rate below .05 (Bonferroni correction). As we hypothesised on the basis of previous research and reports [6;14], that span measures would be wider and level measures higher for the English speakers, we used one-tailed t-tests.

Results showed that all three span measures were significant, 80% range [t(7.9)=2.845;

p<.016], SD4 [t(8.4)=2.793, p<.016] and IQR [t(14)=2.425, p<.016], with wider ranges for the group of English speakers.

Table 1: Means and standard deviation for each dependent variable for speakers of English and German

Dependent variables	Language	N	Mean	SDEV
Mean F0 (Hz)	English	8	213.3	19.7
	German	8	208.6	9.3
80% range (ST)	English	8	8.4	2.1
	German	8	6.2	0.5
IQR (ST)	English	8	4.6	1.3
	German	8	3.4	0.4
SD4 (ST)	English	8	14.2	3.2
	German	8	10.8	1.0

For level, mean f0 was not significantly different between the two groups. The means and standard deviations for each of the dependent variables are presented in Table 1.

Figure 1 gives the information from Table 1 as a visual representation of the span (80% range ST) and level (mean f0 Hz) for all 16 speakers in a scattergraph. From this figure it can be seen that level and span measures seem to be largely independent with there clearly being speakers who have a wide span yet differing levels (e.g. speakers 6 and 3).

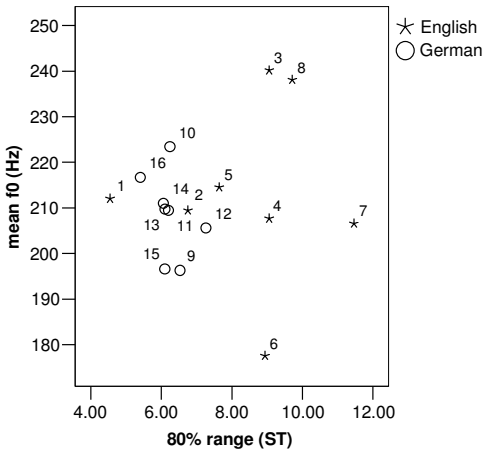
Likewise there are speakers who have very similar levels but differing spans (e.g. speakers 7 and 13). Nevertheless, there is a clustering of the native German speakers at the lower end of the x-axis (representing span) in the figure, while the English speakers cluster mostly at the higher end of the x-axis.

4.2. Sentences

Three mixed ANOVAs with sentence type as a within-subjects factor and language as a between-subjects factor were run on the span measures 80% range (ST), IQR (ST) and SD4 span (ST). To control for family-wise type I error rate, the significance level for main effects and interactions was set to .016 (Bonferroni correction).

No level measures were entered as the results for the text (see 4.1) did not reveal any differences between the groups of speakers and descriptive statistics showed that this was similar for the sentences (i.e. the difference for mean f0 between the two languages was just 6.6 Hz).

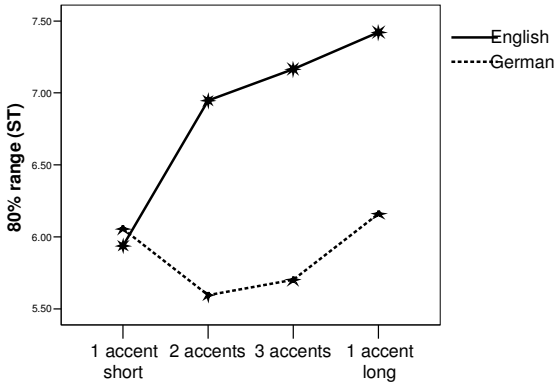
Figure 1: Scattergraph showing span and level for speakers of English and German.



For SD4 span, there was a significant effect of sentence type (F(3,42)=7.331, p<.016). The interaction between sentence type and language and the effect of language was not significant.

For 80% range, there was a significant effect of sentence type (F(3,42)=5.965, p<.016), and a significant interaction between sentence type and language (F(3,42)=7.511, p<.016), cf. Figure 2. The effect of language was not significant. To qualify the interaction between sentence type and language, separate within-subjects ANOVAs were calculated for sentence type for each language group. Sentence type was significant for the English group (F(3,21)=8.647, p<.016). Pair-wise post-hoc comparisons showed a significant difference between short single-accented sentences on the one hand and two- and three-accented sentences on the other hand. For the German group, sentence type was not significant. For IQR, neither sentence type nor language showed a significant main effect, and interaction between sentence type and language fell short off significance (F(3,42)=3.669, p=.020).

Figure 2: Interaction between the factors sentence type and language for the variable 80% range.



5. DISCUSSION

The present study found evidence for a wider pitch span but not pitch level in the group of female speakers of SSBE when compared to productions of similar materials by female speakers of NSG. The cross-language difference in span was most obvious in the text material, although a difference was also found in the sentence materials where the effect of sentence type was significant for English but not for German speakers.

Previous research has found similar effects of sentence length on pitch span – like that found in SSBE – in various languages, amongst others Southern Swedish [3], Danish (e.g. [16]), and British English [11]. These studies found that local f0 maxima (f0 peaks) were higher at the beginning of a sentence or text and f0 minima (final f0) were either constant across different lengths [3;11] or became lower as the length increased [16]. For other languages, such as Catalan [7], Mexican Spanish [15], and American English [13], no effect of sentence length on pitch span was observed. As the measures used for span differ between our study and previous research, it is difficult to interpret these findings without further research.

Although our findings of pitch span differences echo reported stereotypes, we cannot be sure that these beliefs are in fact based on the differences found. It is well possible that there are other factors contributing to the perception of cross-language differences in span. For example, the time spent near the top or bottom of the range may influence perception. Or it may turn out that local rather than global pitch range differences are shaping our auditory impression (as suggested in [17]). In future work we expect to refine our measures of pitch range, by including linguistically based measures (i.e. related to tonal targets; [12]) and different measurement scales (e.g. Hz, ST, ERB) in larger groups of speakers. Finally, we plan to correlate these measures and measurement scales with listeners' perceptual sensitivity to cross-language differences.

6. ACKNOWLEDGEMENTS

This study was funded by the UK Economic & Social Research Council (RES-000-22-1858). We thank Frank Kügler and his colleagues for the collection of the German data.

7. REFERENCES

[1] Boersma, P., Weenink, D. 2006. Praat (Version 4.5). http://www.praat.org.

[2] Brown, A., Docherty, G. J. 1995. Phonetic Variation in Dysarthric Speech As a Function of Sampling Task. *Eur. J. Disorder. Comm.* 30(1), 17-35.

[3] Bruce, G. 1982. Textual Aspects of Prosody in Swedish. *Phonetica* 39, 274-287.

[4] De Pijper, J. R. 1983. *Modelling British English Intonation.* Foris Publications.

[5] Dolson, M. 1994. The Pitch of Speech As a Function of Linguistic Community. *Music. Percept.* 11(3), 321-331.

[6] Eckert, H., Laver, J. 1994. *Menschen und ihre Stimmen: Aspekte der vokalen Kommunikation.* Weinheim: Psychologie Verlags Union.

[7] Estebas-Vilaplana, E. 2000. Peak F0 Downtrends in Central Catalan Neutral Declaratives. *Speech, Hearing and Language: work in progress.* London, 16-41.

[8] Gibbon, D. 1998. German Intonation. In: Hirst, D. J., Di Christo, A. (eds), *Intonation Systems: A Survey of Twenty Languages,* Cambridge, MA: Cambridge University Press, 78-95.

[9] Gilles, P., Peters, J. 2004. *Regional Variation in Intonation.* Tübingen: Niemeyer Verlag.

[10] Grabe, E., Post, B., Nolan, F., Farrar, K. 2000. Pitch Accent Realization in Four Varieties of British English. *J. Phonetics* 28(2), 161-185.

[11] Ladd, D. R. 1988. Declination "Reset" and the Hierarchical Organization of Utterances. *J. Acoust. Soc. Am.* 84(2), 530-544.

[12] Ladd, D. R., Terken, J. 1995. Modelling Intra- and Inter-Speaker Pitch Range Variation. *Proc.of ICPhS.* Stockholm, 386-389.

[13] Liberman, M., Pierrehumbert, J. 1984. Intonational Invariance Under Changes in Pitch Range and Length. In: Aronoff, M., Oehrle, R., Kelley, F., Stephens, B. W. (eds), *Language Sound Structure,* Cambridge, MA: MIT Press, 157-233.

[14] Mennen, I. 2007. Phonological and Phonetic Influences in Non-Native Intonation. In: Trouvain, J., Gut, U. (eds), *Non-Native Prosody: Phonetic Descriptions and Teaching Practice,* Berlin: Mouton de Gruyter.

[15] Prieto, P., Shih, C., Nibert, H. 2007. Pitch Downtrend in Spanish. *J. Phonetics* 24(4), 445-473.

[16] Thorsen, N. 1983. Standard Danish Sentence Intonation - Phonetic Data and Their Representation. *Folia Linguist.* 17, 187-220.

[17] Ulbrich, C. 2006. Pitch Range Is Not Pitch Range. *Proc.Speech Prosody 2006.* Dresden.

[18] van Bezooijen, R. 1995. Sociocultural Aspects of Pitch Differences Between Japanese and Dutch Women. *Lang. Speech* 38, 253-265.