



# Linguistic Measures of Pitch Range in Slavic and Germanic Languages

Bistra Andreeva<sup>1</sup>, Bernd Möbius<sup>1</sup>, Grazyna Demenko<sup>2</sup>, Frank Zimmerer<sup>1</sup>, Jeanin Jügler<sup>1</sup>

<sup>1</sup> Computational Linguistics & Phonetics, Saarland University, Germany

<sup>2</sup> Department of Linguistics, Adam Mickiewicz University, Poland

[andreeva, moebius, zimmerer, juegler]@coli.uni-saarland.de, lin@amu.edu.pl

## Abstract

Based on specific linguistic landmarks in the speech signal, this study investigates pitch level and pitch span differences in English, German, Bulgarian and Polish. The analysis is based on 22 speakers per language (11 males and 11 females). Linear mixed models were computed that include various linguistic measures of pitch level and span, revealing characteristic differences across languages and between language groups. Pitch level appeared to have significantly higher values for the female speakers in the Slavic than the Germanic group. The male speakers showed slightly different results, with only the Polish speakers displaying significantly higher mean values for pitch level than the German males. Overall, the results show that the Slavic speakers tend to have a wider pitch span than the German speakers. But for the linguistic measure, namely for span between the initial peaks and the non-prominent valleys, we only find the difference between Polish and German speakers. We found a flatter intonation contour in German than in Polish, Bulgarian and English male and female speakers and differences in the frequency of the landmarks between languages. Concerning “speaker liveliness” we found that the speakers from the Slavic group are significantly livelier than the speakers from the Germanic group.

**Index Terms:** pitch range, linguistically based measures, cross-language differences, Bulgarian, Polish, German, British English

## 1. Introduction

In human communication pitch variation is used for a range of functions such as the disambiguation of different syntactic structures, signaling the difference between statements and questions, and between different types of question, indicating the emotional state and attitudes of the speaker, highlighting important elements of the spoken message and regulating conversational interaction. Fundamental frequency ( $f_0$ ) can be attributed to two distinct dimensions of a speaker's performance: pitch level and pitch span [17]. Pitch level incorporates the overall height of the speaker's voice whereas pitch span displays the range of frequencies covered by the speaker. Level and span of fundamental frequency are key ingredients of pitch profiles that have been shown to be characteristic for specific linguistic communities (see [18] for different social groups, [9, 32] for different dialects, [13, 14, 21, 7, 15, 22, 23, 24] for different languages, [11, 32, 36, 37] for bilingual speakers).

Language specific profiles have also been found in the perceptual discrimination of languages. A number of studies have shown that listeners can identify their own language

based solely on prosodic cues, such as  $f_0$ , amplitude, and timing [9, 19, 20, 27].

Two Slavic (Bulgarian and Polish) and two Germanic (German and English) languages were the focus of our previous studies [1, 2, 8]. A systematic comparison of various “long-term distributional” (LTD) measures of  $f_0$  showed that male and female speakers of the Slavic group use considerably higher mean and median  $f_0$ , interquartile range, span (in semitones) and maximum  $f_0$ . Furthermore, they revealed larger standard deviations than the speakers in the Germanic group. We found no statistically significant correlation between body size and LTD measures. Taken together, these findings demonstrate that the general pattern of higher and more variable  $f_0$  values for the Slavic speakers compared to Germanic speakers is not necessarily due only to possible physiological differences between speakers of the different languages. Classification with Multi-Layer Perceptrons based on span, kurtosis and skewness as input variables also showed a clear separation of the Germanic from the Slavic group. Systematic differences between tasks were observed in [1] which appears to be attributable to differing strategies that speakers employ when reading short stories versus lists of numbers. Inter-speaker variability was considerably greater for the lists of numbers. The syntactic-semantic structure of the story seems to constrain the speakers' prosodic options.

A possible caveat of these results is that they are based on LTD measures. For instance, Patterson suggests that these LTD measures are not reliable since there are only weak correlations between the LTD and perceptual measures of pitch range [25]. Building on investigations by [16] and [31], Patterson proposes ‘linguistic measures’ as an alternative to measuring pitch range. According to the Bruce & Gårding model of intonational analysis in [5] Patterson assumes that turning points in an  $f_0$  contour (which LTD measures fail to capture) are linked directly to phonological specified tonal targets and are therefore linguistic in nature. The idea to link pitch span and pitch level to specific tonal targets within the contour, such as peaks and valleys is premised on the basic assumptions of the autosegmental metrical approach (cf. [26, 3, 17] among many others).

Patterson's approach was applied by Mennen and her colleagues [22, 23] in a study investigating the pitch range of female speakers of Southern Standard British English (SSBE) and Northern Standard German (NSG). Stereotypically, speakers of NSG are assumed to have a smaller pitch range than SSBE speakers. Therefore, Mennen and her colleagues recorded SSBE and NSG speakers and analysed their data with LTD measures and ‘linguistic’ measures based on Patterson's approach. The results showed that the ‘linguistic’ measures were superior to the LTD measures. Furthermore, they were

able to explain where the source for the stereotypical belief of monotonous German speakers was rooted.

The aim of this study is to investigate whether by using linguistically based pitch range measures, i.e. measures based on specific turning points in the signal that are linguistic in nature such as those proposed by Mennen et al. [23], we are able to characterise differences in pitch range across the four languages (Bulgarian, Polish, English, German) and between language groups (Slavic and Germanic).

## 2. Material and Methods

### 2.1. Materials and subjects

The material analyzed is continuous read speech taken from two comparable multi-lingual speech databases, EUROM-1 (for German and English) [6] and BABEL (for Bulgarian and Polish) [29, 30]. The BABEL database has been designed and recorded using the standards and procedures established in the European Union ESPRIT SAM project. BABEL follows the format of the EUROM-1 database. The passages were originally collected in the late 1980's for German and English and in the late 1990's for Bulgarian and Polish. We used a subset of the data, consisting of one cognitively linked short passage, containing 5 thematically connected sentences, read by 22 speakers per language (11 male and 11 female). The passages were based on identical, real-life topics for the different languages, freely translated and adapted for Bulgarian, German and Polish from the original English texts. The overall length of the analyzed material per language is about 27 minutes.

### 2.2. Measurements

To calculate linguistic measures for the comparative analysis of pitch level and span in Bulgarian, Polish, English and German, pitch contours were manipulated, re-synthesized and labelled manually in Praat following the method proposed by Mennen et al. [23]. This approach distinguishes between tonal landmarks (local maxima and minima) associated with prominent or non-prominent syllables and between initial and non-initial peaks. Every tonal landmark was identified auditorily and visually. Local maxima and minima were labelled  $H^*$  and  $L^*$ , if they aligned with stressed syllable. They were labeled with  $H$  and  $L$  if they aligned with an unstressed syllable. The first peak of a phrase was separately marked as  $H^*i$  or  $Hi$ . The beginnings and the final landmarks were labelled separately: phrase initial  $f_0$  value was labeled as  $I$ , final lows as  $FL$  and final highs as  $FH$ . Figure 1 shows an example of the  $f_0$  stylization process.

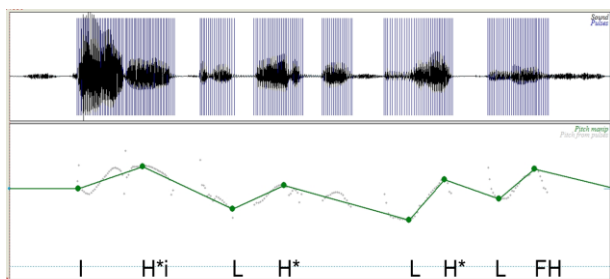


Figure 1: Stylized pitch contour and tonal targets for the Bulgarian Intonation Phrase 'Starijat ribar beše jak maž' (The old fisherman was a big man).

#### 2.2.1. Measures of pitch level

After assigning all landmarks, a Praat [4] script was used to calculate the  $f_0$  value of each landmark. Then, values were averaged across speakers to investigate differences between male and female speakers with different native languages (Bulgarian, Polish, English and German). The following level measures were calculated in Hz: prominent phrase-initial peaks ( $H^*i$ ), prominent non-initial peaks ( $H^*$ ), initial prominent and non-prominent peaks combined (FirstPeak, i.e., the combined measures of  $H^*i$  and  $Hi$ ), non-prominent initial peaks ( $Hi$ ), non-initial non-prominent peaks ( $H$ ), prominent valleys ( $L^*$ ), non-prominent valleys ( $L$ ), and phrase-final lows ( $FL$ ) and phrase-final highs ( $FH$ ).

#### 2.2.2. Measures of pitch span

The following span measures describing the pitch movements along the contours were calculated in semi tones:  $H^*i-L$ ,  $H^*i-FL$ ,  $H^*-L$ ,  $H^*-FL$ , FirstPeak-L, and FirstPeak-FL. The conversion from Hz was performed with the following formula (cf. [28]):

$$(1) \text{Span} = 39.863 * \log_{10}(\text{Max}f_0/\text{Min}f_0)$$

## 3. Results

As a first step towards determining the differences, linear mixed models with the respective measure as dependent variable, SPEAKER and ITEM as random factors, LANGUAGE (Bulgarian/Polish/English/German) and GENDER (male/female) as fixed factors, as well as all their possible interactions, were computed for each dependent variable in separate analyses. Separate Tukey post-hoc tests were carried out per variable, if appropriate. The confidence level was set at  $\alpha=0.05$ .

### 3.1. Distribution of tonal landmarks

The average values in Hz for each landmark obtained from the linguistic measures are plotted in figure 2 for female speakers and figure 3 for male speakers. Visual inspection of figure 2 and figure 3 shows that while the height of initial and non-initial peaks is fairly similar for the German speakers, there is a clear difference in peak heights for the other languages, with higher initial peaks. The patterns for German and English were also found for female speakers in [23].

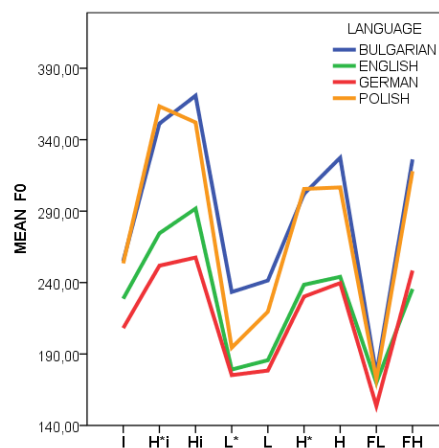


Figure 2: Average values of linguistic measures (female speakers).

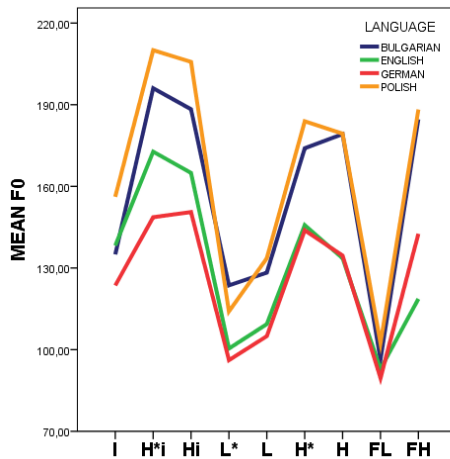


Figure 3: Average values of linguistic measures (male speakers).

We first investigated the distribution of the different targets in our corpus. Table 1 gives an overview of the distribution of the tonal landmarks separately for male and female speakers of the different languages. Bulgarian speakers produced more landmarks than speakers of other languages (female: 872, male 755), followed by Polish speakers (female 698, male 647) and German speakers (female 557, male 521). The speakers with the fewest landmarks were the English speakers (female 439, male 478). This finding is also reflected in the number of intonation phrases that were produced by speakers of the different languages (see below). There were striking difference in the patterns that emerged for the languages, e.g. the high number of H\* and H\*i landmarks for Bulgarian speakers and the markedly lower number of final highs for English speakers compared to those of the other languages. This is also true for the number of H landmarks, and L\*. Male and female speakers also show some differences in the number of landmarks that they produced. Apart from English speakers, where we find the opposite trend, female speakers generally produce more landmarks than their male compatriots. Within the group of Polish speakers, men produce H\* more frequently than women, but men have fewer H and fewer L\*. For German speakers the most striking difference between male and female speakers can be found in the number of H landmarks. Bulgarian female speakers show considerably more FH, H\*, and Hi than Bulgarian male speakers. With respect to Polish speakers, we find the most striking differences in the number of H\* landmarks, male speakers producing more than female speakers, whereas female speakers of Polish produced H and L\* more often than Polish men.

Table 1. Distribution of tonal landmarks by language and speaker sex

	DE		BG		EN		PL	
	f	m	f	m	f	m	f	m
FL	58	55	76	67	71	64	63	65
FH	47	47	62	50	12	11	47	42
H*	99	101	217	176	149	171	129	156
H*i	53	58	87	87	58	64	55	62
H	55	35	75	73	18	13	93	60
Hi	48	43	47	29	25	12	55	43
L*	100	92	84	72	22	21	108	75
L	87	90	224	201	84	122	146	144

### 3.2. Level

Predictably, GENDER had a significant main effect on all level measurements, with females having significantly higher *f0* values. There was also a significant main effect of LANGUAGE on all linguistic measurements for level. The statistical analysis revealed a significant interaction between GENDER and LANGUAGE for prominent initial peaks H\*i (F [3, 79.28] = 4.7644, p<0.01), non-prominent initial peaks Hi (F [3, 74.91] = 4.3380, p<0.05), H\*i and Hi combined (First Peak) (F [3, 80.34] = 5.1156, p<0.01), non-initial peaks H\* (F [3, 84.87] = 3.6293, p<0.05), prominent valleys L\* (F [3, 84.48] = 3.0779, p<0.05) and non-prominent valleys L (F [3, 81.96] = 4.9023, p<0.05). Nearly all pitch level measures were significantly higher for the female speakers in the Slavic than the Germanic group. The only exceptions to this finding are the Polish female speakers with respect to the prominent valleys. The male speakers showed different results, with only the Polish speakers displaying significantly higher mean values for pitch level than the German males. No significant interaction was found for non-prominent non-initial peaks, phrase-final lows and phrase-final highs (see table 2).

Table 2. Language-group differences per gender for the *f0* level measures on the basis of Tukey post-hoc comparisons.

	FEMALES	MALES
H*i	BG = PL > EN = DE	PL = BG = EN > EN = DE
Hi	BG = PL > EN = DE	PL = BG = EN > BG = EN = DE
FIRSTPEAK	BG = PL > EN = DE	PL = BG = EN > EN = DE
L*	BG > PL, EN, DE	N.S.
L	BG = PL > EN = DE	PL = BG = EN > BG = EN = DE
H*	BG = PL > EN = DE	PL = BG = EN > BG = EN = DE
H	BG = PL > EN = DE	
FL	PL = BG = EN > EN = DE	
FH	BG = PL > EN = DE	

### 3.3. Span

GENDER did not differ in *f0* span measured in semitones. Our results for span showed a significant effect of LANGUAGE for the measure H\*–FL (F [3, 80] = 9.7060, p<0.001), H\*i – FL (F [3, 80] = 9.0548, p<0.001), H\*i – L (F [3, 80] = 3.7574, p<0.001), FirstPeak – FL (F [3, 72] = 8.3100, p<0.001) and FirstPeak – L (F [3, 72] = 4.4718, p<0.01). However, this effect was not found for the H\*–L measure. Separate post-hoc tests revealed that the speakers of both Slavic languages tend to have a wider pitch span than the German speakers for the H\*–FL, H\*i–FL and FirstPeak–FL measures. But for the linguistic measures, i.e. for the span between the initial peaks and the non-prominent valleys (H\*i–L and FirstPeak–L), we only find a difference between Polish and German speakers, with a wider pitch span for the Polish speakers.

### 3.4. Liveliness

In a next step, we investigated the extent to which the speakers of the four language groups differed in the liveliness of their productions. Liveliness has several aspects, which we tried to include in our investigation. Not all these aspects have been incorporated into the measures presented in this article, however, this would go beyond the scope of this study, but we focused on aspects that have been identified as being very important for liveliness.

For instance, Traunmüller & Erikson [35] have shown that speech rate influences liveliness. Therefore, we calculated measures per speaker that at least indirectly included speech rate in the liveliness measures: Slopes (i.e. rise or fall per unit time) were calculated between neighbouring target points. Their absolute values were summed (sum of slopes) for the first liveliness measure. A second measure was calculated by normalizing for the number of targets that were realized. This was done by dividing the first liveliness measure by the number of targets. A final value was calculated by dividing the first measure by the duration of the intonation phrases (IPs) that were realized by the speaker. Therefore, a kind of time-normalization was included in all three measures, especially in the third measure, where the duration of the IP was included; linguistic aspects were also part of the measures. In addition, the average syllable duration (without pauses) and the numbers of IPs per passage were measured.

We know that, concerning the perception of liveliness, male and female speakers are not treated the same. Evidence from experiments using manipulated male and female speech with the same average or base-value of *f0* suggests that if the syllables are to be heard as equally prominent or lively, the *f0* excursions need to be larger in the female speech [12, 35]. As [34] pointed out, listeners arguably evaluate the *f0* excursions with respect to the spectral space below the neutral F1, which is higher in female speech. For this reason, we calculated statistical linear mixed models for each of the three liveliness measures with the following factors: ITEM and SPEAKER were entered as random factors, GENDER and LANGUAGE as well as their interaction was also included as fixed factors.

We did not control for other factors, such as differences in rhythm and intensity or (perceived) age of the speaker which can also contribute to differences in liveliness. Very dysfluent speech has also been found to have an effect on measures of liveliness, such as the Pitch Dynamism Quotient (PDQ – [35]). This possible confound is not problematic for the measures reported here. Due to the hand-labelling of the items, and corrections for wrong measurements, we concentrated only on linguistically important targets.

The results of the analyses can be summarized as follows. For the first liveliness measure (sum of slopes), we find that both GENDER ( $F(1, 86.65) = 28.42, p < 0.001$ ) and LANGUAGE ( $F(3, 86.06) = 19.95, p < 0.001$ ) are significant factors, the interaction is not. Post-hoc tests show that the female speakers are livelier than the male speakers. Concerning the effect of language, these tests indicate that Polish and Bulgarian speakers are significantly livelier than English and German speakers.

The results for the second analysis (normalized by the number of targets) show the same trends: GENDER ( $F(1, 85.36) = 33.34, p < 0.001$ ) and LANGUAGE ( $F(3, 85.11) = 14.38, p < 0.001$ ) are significant factors, the interaction is not. Again, female speakers were livelier than male speakers, and Bulgarian and Polish speakers were livelier than German and English speakers, but no difference was found between the Bulgarian and Polish speakers, or between the German and English speakers.

The third analysis confirmed the pattern from the first two analyses: GENDER ( $F(1, 96.31) = 12.7, p < 0.001$ ) and LANGUAGE ( $F(3, 95.36) = 7.87, p < 0.001$ ) are significant factors, and the interaction turned out to be non-significant, as before. The same pattern emerged for male and female speakers, as well as for speakers of the four languages.

With respect to the average syllable duration we found significant effect of GENDER ( $F(1, 74) = 13.5608, p < 0.001$ ) and LANGUAGE ( $F(3, 75) = 8.1922, p < 0.001$ ), with females having significantly longer average syllable duration (186.01 ms) than males (165.15 ms) and Bulgarian speakers speaking significantly faster (153.29 ms) than the Polish (176.07 ms), English (186.33 ms) and German (186.64 ms) speakers. As far as the number of IPs per passage is concerned, we found a main effect of LANGUAGE ( $F(3, 101.3) = 9.3516, p < 0.001$ ), with English having significantly fewer IPs per passage (4.2) than German (5.3), Polish (5.6) and Bulgarian (6.3).

## 4. Discussion and Conclusions

The question that was investigated in this study is whether by using linguistically based pitch range measures we are able to characterise the differences in pitch range across Bulgarian, Polish, English and German and between language groups (Slavic and Germanic) found in our previous studies [1, 2, 8].

Overall, the results of this study confirm the general pattern found in our previous studies that the speakers of the Slavic group used considerably higher level and wider span. As far as *f0* level is concerned, the results showed that pitch level appeared to have significantly higher values for the female speakers in the Slavic group. The male speakers showed different results, with the Polish speakers displaying significantly higher mean values for pitch level than the German males. As far as *f0* span is concerned, the results show that the Slavic speakers tend to have a wider pitch span than the German speakers. But for the linguistic measure of the span between the initial peaks and the non-prominent valleys we only find difference between Polish and German speakers, with a wider pitch span for the Polish speakers. In terms of declination over the peaks of the intonation contours, however, the German male and female speakers have a smaller span than all three other languages, making the German contours flatter.

The liveliness measures are still exploratory in nature, since they have not been tested with respect to perceived liveliness. However, all of them show the same tendencies: Bulgarian and Polish speakers are closer to each other than to German and English speakers which, in turn, are very similar. The results reported here are intended to present a starting point from which we can base perceptual studies as well as refinements of what aspects to evaluate in the future.

Our findings support the hypothesis that linguistic communities tend to be characterized by particular pitch profiles.

## 5. Acknowledgements

We wish to thank Ryszard Gubrynowicz (Institute of Fundamental Technology Research, Polish Academy of Science) and Snezhina Dimitrova (English Dept., Sofia University “St. Kliment Ohridski”), for kindly providing the Babel databases for Polish and Bulgarian, respectively.

## 6. References

- [1] Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., Jastrzebska, M. and Trouvain, J. (2014). Comparison of pitch range and pitch variation in Slavic and Germanic languages. *Proc. Speech Prosody 7*, Dublin, 776–780.
- [2] Andreeva, B., Demenko, G., Möbius, B., Zimmerer, F., Jügler, J. and Jastrzebska, M. (2014). Differences of Pitch Profiles in Germanic and Slavic Languages. *Proc. Interspeech 2014*, 14–18 September, Singapore.
- [3] Beckman, M. and Pierrehumbert, J. (1986). Intonational Structure in Japanese and English, *Phonology Yearbook*, 3, 15–70.
- [4] Boersma, P. and Weenink, D. (2015). *Praat: doing phonetics by computer* [computer software, version 5.4.01]. Amsterdam: University of Amsterdam. Available at <http://www.praat.org>
- [5] Bruce, G. and Gårding, E. (1978). A Prosodic Typology for Swedish Dialects. Gårding, E., Bruce, G., Bannert, R. (eds), *Nordic Prosody Gleeurp*, 219–228.
- [6] Chan, D., Fourcin, A., Gibbon, D., Granstrom, B., Huckvale, M., Kokkinakis, G., Kvale, K., Lamel, L., Lindberg, B., Moreno, A., Mouropoulos, J., Senia, F., Trancoso, I., Veld, C. and Zeiliger, J. (1995). Eurocom - a spoken language resource for the EU. In *Eurospeech' 95. Proc. of the 4th European Conference on Speech Communication and Speech Technology*, 1, Madrid, 18–21 September 1995, 867–870.
- [7] Chen, G. T. (1972). *A comparative study of pitch range of native speakers of Midwestern English and Mandarin Chinese: An acoustic study*, doctoral dissertation, University of Wisconsin-Madison, Madison.
- [8] Demenko, G., Möbius, B., Andreeva, B. (2014). Analysis of pitch profiles in Germanic and Slavic languages. *Forum Acusticum 2014*, 7–12 September, Krakow. ISSN 2221-3767 (for CD), ISBN 987-83-61402-28-2 (for Book of Abstracts).
- [9] de Pijper, J. R. (1983). *Modelling British English intonation*, Dordrecht - Holland: Foris.
- [10] Deutsch, D., Le, J., Shen, J., and Henthorn, T. (2009). The pitch levels of female speech in two Chinese villages. *Journal of the Acoustical Society of America*, April, 125, EL208.
- [11] Graham, C., (2013). Revisiting f0 Range Production in Japanese-English Simultaneous Bilinguals. *Annual Report of UC Berkeley Phonology Lab*, 110–125.
- [12] Gussenhoven, C., and Rietveld, T. (1998). On the speaker-dependence of the perceived prominence of F0 peaks, *Journal of Phonetics*, Vol. 26: 371–380.
- [13] Hanley, T.D., Snidecor, J.C., and Ringel, R. (1966). Some acoustic differences among languages, *Phonetica* 14, 97–107.
- [14] Hanley, .D. and Snidecor, J.C. (1967). Some acoustic similarities among languages, *Phonetica* 17, 141–148.
- [15] Keating, P. & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin, *Journal of the Acoustical Society of America* 132, 1050–1060.
- [16] Ladd, D. R. and Terken, J. 1995. Modelling Intra- and Inter-Speaker Pitch Range Variation. *Proc. of ICPhS*. Stockholm, 386–389.
- [17] Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- [18] Luchsinger, R. and Arnold, G. (1965). *Voice-Speech-Language*. Constable&Co Ltd., London.
- [19] Maidment, J. A. (1976). Voice fundamental frequency characteristics as language differentiators. *Speech and hearing: Work in progress*, University College London, 74–93.
- [20] Maidment, J. A. (1983). Language recognition and prosody: further evidence, *Speech, hearing and language: Work in progress*, University College London 1, 133–141.
- [21] Majewski, W., Hollien, H. and Zalewski, J. (1972). Speaking fundamental frequency of Polish adult males, *Phonetica* 25, 119–125.
- [22] Mennen, I., Schaeffler, F. and Docherty, G. (2007). Pitching it differently: A comparison of the pitch ranges of German and English speakers. *16th International Congress of Phonetic Sciences (ICPhS XVI)*, Saarbrücken, 1769–1972.
- [23] Mennen, I., Schaeffler, F. and Docherty, G. (2012). Cross-language differences in fundamental frequency range: a comparison of English and German. *Journal of the Acoustical Society of America* 131(3), 2249–2260.
- [24] Nebert, A. U. (2013). *Der Tonhöhenumfang der deutschen und russischen Sprechstimme. Vergleichende Untersuchung zur Sprechstimmlage*. Halesche Schriften zur Sprechwissenschaft und Phonetik, Band 46. Frankfurt/M.
- [25] Patterson, D. (2000). *A Linguistic Approach to Pitch Range Modelling*. Ph.D. dissertation, University of Edinburgh.
- [26] Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*, Ph.D. thesis, MIT, Cambridge, MA.
- [27] Ramus, F. and Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America* 105 (1), 512–521.
- [28] Reetz, H. (1999): *Artikulatorische und akustische Phonetik*. Wissenschaftlicher Verlag, Trier.
- [29] Roach, P., Arnfield, S. and Hallum, E., (1996). BABEL: A multi-language speech database. *Proc. of SST-96: Speech and Science Technology Conference*, Adelaide, 351–354.
- [30] Roach, P., Arnfield, S., Barry, W.J., Dimitrova, S., Boldea, M., Fourcin, A., Gonet, W., Gubrynowicz, R., Hallum, E., Lamel, L., Marasek, K., Marchal, A., Meister, E. and Vicsi, K. (1998). Babel: a database of Central and Eastern European languages, *Proc. of the First International Conference on Language Resources and Evaluation*, vol. 1, 28-30 May 1998, Granada, Spain, 371–374.
- [31] Shriberg, E., Ladd, R. D., Terken, J. and Stolcke, A. (1996). Modeling Pitch Range Variation Within and Across Speakers: Predicting F0 Targets When "Speaking Up". *Proc. of the International Conference on Spoken Language Processing (ICSLP)*, Philadelphia, PA.
- [32] Todaka, Y. (1993). *A cross-language study of voice quality: bilingual Japanese and American speakers*, doctoral dissertation, University of California, Los Angeles.
- [33] Torgerson, R. C. (2005). *A comparison of Beijing and Taiwan Mandarin tone register: An acoustic analysis of three native speech styles*, master's thesis, Brigham Young University, 73–82.
- [34] Traunmüller, H. (2000). Evidence for demodulation in speech perception, *Proc. of the 6th ICSLP*, vol. III: 790–793.
- [35] Traunmüller, H. and Eriksson, A. (1995). The perceptual evaluation of F0 excursions in speech as evidenced in liveliness estimations, *Journal of the Acoustical Society of America*, Vol. 97: 1905–1915.
- [36] Yamazawa, H., and Hollien, H. (1992). Speaking fundamental frequency pattern of Japanese women, *Phonetica* 49, 128–140.
- [37] Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B. and Trouvain, J. (2014). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. *Proc. Speech Prosody 7*, Dublin, Ireland, 1037–1041.