The influence of L1 prosody on Bulgarian-accented German and English

Bistra Andreeva¹, Snezhina Dimitrova²

¹Language Science and Technology, Saarland University, Germany
²Department of English and American Studies, Sofia University “St. Kliment Ohridski”, Bulgaria

andreeva@lst.uni-saarland.de, snezhina@uni-sofia.bg

Abstract

The present study investigates L2 prosodic realizations in the readings of two groups of Bulgarian informants: (a) with L2 German, and (b) with L2 English. Each of the two groups consisted of ten female learners, who read the fable “The North Wind and the Sun” in their L1 and in the respective L2. We also recorded two groups of female native speakers of the target languages as controls. The following duration parameters were obtained: mean accented syllable duration, accented vs. unaccented syllable duration ratio, and speaking rate. With respect to F0 parameters, mean, median, minimum, maximum, span in semitones, and standard deviations per IP were measured. Additionally, we calculated the number of accented and unaccented syllables, IPs and pauses in each reading. Statistical analyses show that the two groups differ in their use of F0. Both groups use higher standard deviation and level in their L2, whereas the ‘German group’ use higher pitch span as well. The number of accented syllables, IPs and pauses is also higher in L2. Regarding duration, both groups use slower articulation rate. The ratio between accented and unaccented syllables is lower in L2 for the ‘English group’. We also provide original data on speaking rate in Bulgarian from an information theoretical perspective.

Index Terms: L2 German, L2 English, L1 Bulgarian, duration, fundamental frequency

1. Introduction

Research into second-language prosody during the last couple of decades has focused on various aspects of the suprasegmental characteristics of L2 speech ([1], [2], [3], [4], [5], [6], [7], among many others). Fundamental frequency-related features of L2 production which have been studied include the use of specific pitch patterns or contours, the placement of pitch accents, their phonetic realization, pitch range and pitch level variation. In particular, pitch range has been investigated in the talk of speakers from diverse L1 backgrounds, and of learners who speak different L2 target languages ([6], [8], [9], [10], [11], [12], [13], [14], [15], [16]). Most results seem to suggest that L2 learners commonly use a lower pitch level and a compressed pitch span in the target language. One possible explanation that is put forward in order to account for this finding is that learners are less confident when they speak the L2 than when they talk in their mother tongue ([16]). As a consequence of using a narrower pitch range, they can be perceived as sounding dull and monotonous ([6], [11]). However, there are also findings showing that some speakers may use a higher F0 level or a wider span in the L2 (e.g., [6] for German vs. English). It is thus questionable whether the tendency to compress the pitch range when speaking a foreign language can be considered universal, and the roles played by the language-appropriate L1 and L2 pitch range norms also remain unclear.

The investigation of the temporal characteristics of L2 speech has focused on speaking rates, the duration of segments, syllables, intonational phrases and pauses. More pauses of longer duration ([17]) and slower and more variable speaking rates ([18], [19]) have been found in second-language speech compared with first-language speech. These durational characteristics have also been related to L1 and L2 fluency ([20], [21], [22]). Recent work on speaking rate involving L1 English monolinguals and L2 English bilinguals from diverse language backgrounds ([23], [24]) has shown that L1 speaking rate can significantly predict L2 speaking rate, and that the more variable second-language speaking rates which have been found can also be accounted for if L1 interspeaker variation is taken into account.

Recent research has also shown that speaking rate variations found across languages have an inverse relation to information density (ID): high density languages that require relatively few syllables to transmit information during speech communication are produced at relatively slow speaking rates compared to low density languages ([25], [23], [24] and references therein). In L2 speech, on the other hand, the inverse relationship between ID and speaking rate was not observed ([24]). This finding suggests less systematicity in L2 temporal structure.

The present study investigates the amount of interference of L1 F0- and duration-related characteristics in the speech of advanced Bulgarian learners of two related Germanic languages – German and English. Andreeva et al. [15] have found that Bulgarian male and female speakers use wider pitch range and are more variable compared to German and English speakers. Assuming that there is transfer of F0-related characteristics from the L1, we expect expansion of the L2 (German and English) target norms for pitch range. Alternatively, there may be adaptation of the native language pitch range to that of the target language. To the best of our knowledge, no systematic investigations of speaking rate in Bulgarian exist to date. Our aim is to fill this gap by providing data on speech and articulation rate in L1 Bulgarian and compare it with similar data for L1 and L2 German and English. Taking an information theoretic approach, we also want to study the relationship between ID factors and speaking rate in the three languages under investigation.

2. Method

2.1. Corpus

To test our predictions, we recorded two data sets (a) a ‘German data set’ and (b) an ‘English data set’. The German data set consisted of ten Bulgarian speakers of German and ten German
native speakers as controls. The English data set consisted of ten Bulgarian speakers of English and six English native speakers as controls. All speakers were female university students of comparable age (average 20.7 years) and spoke the respective standard language varieties. The Bulgarian participants had some knowledge of the phonetics and phonology of the respective L2.

The material recorded was Aesop’s fable “The North Wind and the Sun”, with the Bulgarians reading the text in Bulgarian as well as in their respective L2 (German or English).

2.2. Measurements

First, syllable and Intonation Phrase (IP) boundaries as well as pauses were segmented and lexically stressed syllables were labelled manually in Praat [26]. Second, all accented syllables were marked and counted, including those in lexical words with double prominence and in prominent function words. We also counted the numbers of pauses and IPs per reading.

2.2.1. Pitch analysis

Pitch analysis was performed as follows. First, F0 was extracted automatically from all recordings by means of the ESPS algorithm (“get_f0” [27]) with time steps of 5 ms. Secondly, a manual inspection and correction (i.e., removal of data points) of the extracted pitch contours was performed in Praat. The corrections included the removal of octave jumps as well as other artefacts (e.g., due to creaky voice). From the cleaned data the following F0 long-term distributional (LTD) measures per IP were calculated using Praat scripts: mean, median, minimum, maximum, standard deviation (all in Hz), and span in semitones (ST). The conversion from Hz was performed with the following formula [28]:

\[
\text{Pitch Span (ST)} = 39.863 \times \log_{10}\left(\frac{\text{MaximumF0}}{\text{MinimumF0}}\right)
\]

2.2.2. Temporal features

The durations of the IPs, pauses and prominent syllables were extracted per reading, speaker and native/target language using Praat scripts. Mean duration of accented syllables as well as accented/unaccented syllable duration ratios were computed. In addition, we calculated two measurements of speaking rate: (a) speech rate (SR, the number of canonical syllables divided by the duration of the respective recording) and (b) articulation rate (AR, the number of canonical syllables divided by the sum of IP durations per reading).

2.2.3. Information density (ID)

Considering that the semantic content (S) of the fable is the same in Bulgarian (BG), German (DE) and English (GB), we calculated normalized ID for each text by each language, using Bulgarian as a benchmark and following the methodology proposed in [25]:

\[
1. \text{ID}_L = \frac{1}{L} \times \frac{1}{\text{BG}} = \frac{S}{\sigma_{BG}} \times \frac{\sigma_{BG}}{S} = \frac{\sigma_{BG}}{\sigma_L}
\]

For each text I and language L, the normalized ID(I) was calculated through a pairwise comparison of the text lengths (in terms of canonical syllables) in L and BG respectively. ID values smaller than one indicate lower ID than the Bulgarian benchmark (identical semantic content transmitted with more syllables). ID values greater than one indicate higher ID than the Bulgarian benchmark (identical semantic content transmitted with fewer syllables).

2.3. Statistical analyses

For statistical validation, we used the software JMP 16 [29]. For both the F0- and the duration-related parameters we calculated two different models: one for the German data set and one for the English data set, comparing native language and target language. Linear mixed models (LMM) were fitted for the duration-related parameters, with the respective log-transformed measure as dependent variable, SPEAKER as random factor, and LANGUAGE (native language/target language) as fixed factor. Separate Tukey post-hoc tests were carried out per variable, if appropriate. The confidence level was set at \( \alpha = .05 \).

For the analyses of the F0-related parameters we used a nonparametric Kruskal-Wallis test because the data were not equally distributed. To determine differences between speaker groups we performed post-hoc Dunn’s pairwise tests with Bonferroni adjustment.

Since the two groups of learners differ in their L2 language proficiency, the analyses will be carried out for each group separately.

3. Results

3.1. F0-related parameters

Means and standard deviations for each of the F0-related parameters are presented in Table 1 for the German data set and in Table 2 for the English data set. Since we use nonparametric tests we also report the median values.

Following Ladd [30], we consider the measures for mean and median (related to pitch level) and span to be attributes of pitch range, and the standard deviation (SD) – an attribute of pitch variation.

3.1.1. German data set

<table>
<thead>
<tr>
<th>parameter</th>
<th>BG_L1</th>
<th>DE_L2</th>
<th>DE_L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>243 (38.5)</td>
<td>230 (31.1)</td>
<td>214 (21.5)</td>
</tr>
<tr>
<td>median</td>
<td>238 (39.0)</td>
<td>225 (31.7)</td>
<td>211 (22.6)</td>
</tr>
<tr>
<td>SD</td>
<td>36.7 (36.7)</td>
<td>30.7 (19.7)</td>
<td>20.6 (7.8)</td>
</tr>
<tr>
<td>span (ST)</td>
<td>9.1 (3.7)</td>
<td>8.1 (4.1)</td>
<td>6.2 (1.9)</td>
</tr>
</tbody>
</table>

For the German data set, we found a main effect of LANGUAGE on all LTD measurements (see Table 1 and Figure 1). Post-hoc tests revealed significant differences between the three groups for pitch level: mean \( (\chi^2(2) = 75.91, p < .0001) \), median \( (\chi^2(2) = 63.03, p < .0001) \), pitch span \( (\chi^2(2) = 90.63, p < .0001) \) and pitch variation \( (\chi^2(2) = 78.80, p < .0001) \) with the highest values for German spoken as a native language (BG_L1), the lowest values for German spoken as a native language (DE_L1) and intermediate values for the target language - German (DE_L2). These findings indicate that the Bulgarian speakers use a narrower pitch range and are less variable in the target language which confirms results in earlier studies (e.g. [10], [18]).
3.1.2. English data set

For the English data set, we also found a main effect of LANGUAGE (see Table 2 and Figure 2) on all measurements: for mean ($\chi^2(2) = 39.86, p < .0001$), for median ($\chi^2(2) = 39.29, p < .0001$), for SD ($\chi^2(2) = 36.09, p < .0001$), for span ($\chi^2(2) = 30.99, p < .0001$). Post-hoc tests revealed that (a) for pitch level, BG_L1 and GB_L2 have significantly higher values than GB_L1, (b) for pitch span, GB_L1 has significantly higher values than GB_L2 and GB_L1, and (c) for pitch variation, the three groups have significantly different values with BG_L1 having higher values than GB_L2, which in turn has higher values than GB_L1.

Table 2: F0-related parameters for Bulgarian and English.

<table>
<thead>
<tr>
<th>parameter</th>
<th>BG_L1</th>
<th>GB_L1</th>
<th>BG_L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>223 (32.0)</td>
<td>220 (33.4)</td>
<td>200 (21.9)</td>
</tr>
<tr>
<td>median</td>
<td>219 (32.1)</td>
<td>217 (35.5)</td>
<td>198 (21.7)</td>
</tr>
<tr>
<td>SD</td>
<td>27.79 (27.8)</td>
<td>23.7 (14.2)</td>
<td>18.1 (8.1)</td>
</tr>
<tr>
<td>span (ST)</td>
<td>7.78 (3.2)</td>
<td>6.6 (3.0)</td>
<td>6.1 (2.8)</td>
</tr>
<tr>
<td>median values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>217 (32.0)</td>
<td>214 (33.4)</td>
<td>198 (21.9)</td>
</tr>
<tr>
<td>median</td>
<td>216 (32.1)</td>
<td>213 (35.5)</td>
<td>197 (21.7)</td>
</tr>
<tr>
<td>SD</td>
<td>24.1 (15.3)</td>
<td>20.1 (14.2)</td>
<td>17.7 (8.1)</td>
</tr>
<tr>
<td>span (ST)</td>
<td>7.4 (3.2)</td>
<td>6.6 (3.0)</td>
<td>5.5 (2.8)</td>
</tr>
</tbody>
</table>

The comparison between BG_L1 and DE_L1 vs. BG_L1 and GB_L1 corroborates the results in [15] who report the use of wider pitch range and higher variability in two Slavic languages (Bulgarian and Polish) compared to two Germanic languages (German and English). The latter findings can also account for the higher F0-related parameter values which we found in DE_L2 and GB_L2 compared to DE_L1 and GB_L1, respectively.

3.2. Duration-related parameters

Means and standard deviations for each of the duration-related parameters are presented in Table 3 for the German data set and in Table 4 for the English data set.

We compare the duration-related parameters in native German and English readings with those in the respective target language readings by the Bulgarian speakers. At this point we do not include BG_L1 in the analyses because of the differences between the respective texts in terms of number of words and syllables, syllable complexity, etc.

3.2.1. German data set

The Bulgarian speakers of German produced considerably more IPs and silent pauses than the native German speakers (27.2 vs. 18.9 IPs, and 20.6 vs. 12.1 pauses, respectively). They also produced more accented syllables than the natives (74.5 vs. 50.8). Both of these findings are in line with previous research (e.g. [17]).

Table 3: Duration-related parameters for Bulgarian and German.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BG_L1</th>
<th>DE_L2</th>
<th>DE_L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean acc. dur.</td>
<td>218.3 (31.4)</td>
<td>300.2 (63.5)</td>
<td>235.5 (17.6)</td>
</tr>
<tr>
<td>acc./unacc. ratio</td>
<td>1.6 (0.1)</td>
<td>1.5 (0.2)</td>
<td>1.6 (0.1)</td>
</tr>
<tr>
<td>articulation rate</td>
<td>6.2 (1.1)</td>
<td>4.3 (0.8)</td>
<td>5.8 (0.5)</td>
</tr>
<tr>
<td>speech rate</td>
<td>5.1 (0.8)</td>
<td>3.5 (0.7)</td>
<td>5.0 (0.5)</td>
</tr>
</tbody>
</table>

There is no main effect of language with respect to the accented/unaccented syllable duration ratio in DE_L2 and DE_L1. This is not surprising given the similar ratio between accented and unaccented syllables in BG_L1. However, we found significant differences between accented syllable duration in DE_L1 and DE_L2 with longer durations in Bulgarian-accented German ($F[1, 18] = 9.62, p<0.0062$). As for speaking rate, the Bulgarian speakers of German were significantly slower than the German native speakers: speech rate ($F[1, 18] = 32.03, p<0.0001$), articulation rate ($F[1, 18] = 23.21, p<0.0001$). This also accounts for the longer accented syllable duration in their German productions.

3.2.2. English data set

The Bulgarian speakers of English also produced a higher number of IPs and silent pauses than the native English speakers (26.0 vs. 16.8 IPs, and 13.3 vs. 9.3 pauses, respectively). They also produced more accented syllables than the natives (58.5 vs. 49.2). Our analyses show a main effect of language with respect to the accented/unaccented syllable duration ratio ($F[1, 7] = 8.61, p<0.0109$), which is significantly lower in GB_L2 than in GB_L1. On the other hand, the mean accented syllable durations in GB_L2 and GB_L1 do not differ significantly. This suggests that the Bulgarian speakers of English do not reduce sufficiently the duration of unaccented syllables in the target language. As for articulation rate, the Bulgarian speakers of English were significantly slower than the English native speakers ($F[1, 7] = 27.12, p<0.0001$), whereas the speech rates of the two groups did not differ significantly. We attribute this to the longer pause durations between utterances produced by the native English speakers.

Analyses were also carried out by collapsing the speaking rate measurements in the L2 realizations of all the Bulgarian speakers in one group and comparing them to the speaking rate measurements in their L1 realizations. Pearson correlations...
were used to determine whether L2 speaking rate can be predicted on the basis of L1 speaking rate. We found a strong positive correlation between L1 and L2 speaking rates (speech rate: \( r(24) = .783, p < .0001 \); articulation rate: \( r(24) = .791, p < .0001 \)) which is in line with the findings in [(20), (21), (23)].

Table 4: Duration-related parameters for Bulgarian and English.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BG_L1</th>
<th>GB_L2</th>
<th>GB_L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean acc. dur.</td>
<td>209.7(16.3)</td>
<td>338.1(34.8)</td>
<td>305.0(25.7)</td>
</tr>
<tr>
<td>acc./unacc. ratio</td>
<td>1.5(0.3)</td>
<td>2.2(0.2)</td>
<td>2.5(0.9)</td>
</tr>
<tr>
<td>articulation rate</td>
<td>6.2(0.8)</td>
<td>4.4(0.4)</td>
<td>5.5(0.3)</td>
</tr>
<tr>
<td>speech rate</td>
<td>5.3(0.4)</td>
<td>3.9(0.6)</td>
<td>4.5(0.4)</td>
</tr>
</tbody>
</table>

3.3. Information density

The ID values for Bulgarian, German and English are given in Table 5. The values for L1 and L2 English and L1 and L2 German are identical, because they were calculated on the basis of the canonical syllables in the text. ID values higher than one indicate a “denser” language than the Bulgarian benchmark (identical semantic content transmitted with fewer syllables). The higher values for English compared to German correspond to those reported in [25]. On the other hand, both languages have higher values than Bulgarian.

Table 5: Cross-language comparison of information density. Bulgarian is used as a benchmark.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BG_L1</th>
<th>DE_L1</th>
<th>GB_L1</th>
<th>DE_L2</th>
<th>GB_L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID(reBG)</td>
<td>1</td>
<td>1.1</td>
<td>1.4</td>
<td>1.1</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Since denser languages transmit information at slower speaking rates and vice versa, we expected to find significant negative correlation between ID and speaking rate in the realizations of all L1 speakers. Spearman rho correlation analysis confirmed this expectation for SR (\( r(24) = -.41, p < .0133 \)) as well as for AR (\( r(24) = -.34, p < .0386 \)). This result corroborates the findings in [24] and [25] with respect to a new language – Bulgarian.

4. Discussion and conclusions.

In this study we investigated the prosodic characteristics of Bulgarian-accented L2 German and English compared to (a) L1 German and English, and (b) L1 Bulgarian. All F0-related parameters in the speech of the Bulgarian learners of German were lower than in their L1 but higher than those of the native German speakers. This was also true for the variability of the Bulgarian learners of English. However, the latter group of learners had higher values for level compared to the English native speakers, and higher values for span compared to their L1. Thus, our first assumption that the L2 target norms for pitch range will be expanded due to L1 influence was confirmed for the Bulgarian learners of German and for the pitch level and variability of the Bulgarian learners of English. With regard to the pitch span used by the latter group, we observed adaptation of the native language norms towards those of the target language.

With regard to the duration-related parameters, we found that the Bulgarian speakers used slower articulation rate, more IPs and pauses in their L2 than the native speakers. They also failed to de-accentuate: we found more accented syllables in L2. In the absence of statistically significant difference between accented syllable duration in L1 and L2 English, we interpret the lower ratio between accented and unaccented syllable durations in L2 English as an indication of the smaller amount of reduction of unaccented syllables. The strong correlation found between L1 and L2 speaking rates of the Bulgarian speakers show evidence that L1 speaking rate can indeed predict the speaking rate in L2. Our results suggest that the so-called L2 speaking style is influenced by L1 prosody with respect to F0-related features. As for duration-related characteristics of L1, they can explain some of the variability found in L2 speech. However, durational variation due to different L2 proficiency levels is also worth further investigation.

Our analysis of the duration-related parameters also provides first insights into the speaking rate characteristics in Bulgarian from an information theoretical perspective. The comparison with L1 German and English reveals that Bulgarian is a ‘fast language’.

5. Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 232722074 – SFB 1102 ‘Information Density and Linguistic Encoding’ at Saarland University and by the Bulgarian National Science Fund, project No. Kt-06-H40/11 from 12.12.2019 ‘Prosodic aspects of Bulgarian in comparison with other languages with lexical stress’.

6. References


