Suprasegmental Features of Bulgarian English Speech

Snezhina Dimitrova
Sofia University "St. Kliment Ohridski" (Bulgaria)

Snezhina Dimitrova. SUPRASEGMENTAL FEATURES OF BULGARIAN ENGLISH SPEECH

Abstract. The suprasegmental characteristics of English speech constitute a well-known area of difficulty even for fairly advanced Bulgarian learners of the language. However, no systematic research has been carried out into these problems, not least because of the lack of an established theoretical and methodological framework for such investigation. The present study makes a first step towards filling this gap. It presents results from a contrastive study of stress and intonation in the speech of tertiary-level Bulgarian students of English. Six students read and recorded the Bulgarian and the English version of Aesop's fable "The North Wind and the Sun" – a standard text used in phonetic research. The recordings were analyzed acoustically and labelled intonationally in Praat using ToBI. A number of long-term distributional measures were obtained, namely, mean and median fundamental frequency (F0), pitch minima and maxima, pitch span, temporal characteristics such as mean syllable, intonation phrase (IP) and pause duration, as well as number of IPs, pauses, stressed and unstressed syllables. The results were next compared with data from recordings of the fable by native English Received Pronunciation (RP) speakers. Differences between Bulgarian, Bulgarian English and British English RP were found in terms of F0 measures, IP duration and number, as well as in the number of pauses, stressed and unstressed syllables. These results are discussed within the framework of a newly proposed model of L2 intonation learning (Mennen, 2015).

Keywords: suprasegmentals, English, Bulgarian, L2, Bulgarian English

Снежина Димитрова. СУПРАСЕГМЕНТНИ ХАРАКТЕРИСТИКИ НА АНГЛИЙСКАТА РЕЧ НА БЪЛГАРИ

Резюме. Супрасегментните характеристики на английската реч създават затруднения дори на българи, които изучават езика на ниво напреднали. Липсват обаче систематични изследвания на тези проблеми, като една от причините за това се корени в липсата на адекватна теоретико-методологическа рамка за провеждането им. Настоящото изследване представлява опит да бъде запълнена тази празнина. То представя резултати от анализа на основните акустични характеристиките на ударението и интонацията в речта на български студенти англицисти. Шестима студенти прочетоха

и записаха баснята „Северният вятър и слънцето“ (текст, който често се използва във фонетични изследвания) на български и на английски език. Записите бяха подложени на акустичен анализ и бяха маркирани интонационно с помощта на програмата за акустичен анализ *Praat* и на системата *ToBI*. Получените данни за честотата на основния тон (средна стойност и медиана, минимум и максимум, честотен обхват), темпорални характеристики като средно времетраене на срички, интонационни фрази и паузи, както и данните за общ брой интонационни фрази, паузи, ударени и неударени срички бяха сравнени с данни за тези характеристики в стандартния британски английски акцент, известен с името *RP*. Установени бяха разлики между български, английски като роден и английски като чужд език по отношение на стойностите на честотата на основния тон, времетраенето и броя на интонационните фрази, паузите, ударените и неударените срички. Данните са представени в контекста на новия модел за изучаване на интонацията на чужд език, предложен от Менен (2015).

Ключови думи: супрасегментни характеристики, английски, български, български английски език

Research/Научно изследване

We investigated some suprasegmental characteristics typical of the speech of advanced Bulgarian learners of English as part of an ongoing research project entitled “Prosodic aspects of Bulgarian in comparison with other languages with lexical stress” carried out by researchers from the Faculty of Slavic Studies and the Faculty of Classical and Modern Philology at Sofia University, and the Department of Language Science and Technology at Saarland University, Germany. We were interested in studying the prosodic features of what we will henceforth call “Bulgarian English” (BE) speech – the spoken language of Bulgarian learners of English. Suprasegmental features of speech, or just “suprasegmentals” (Lehiste, 1970), also known as “prosodic features” (Crystal and Quirk, 1964; Crystal, 1969), refer to the variations in fundamental frequency (F0, perceived as pitch), intensity (perceived as loudness), and timing or duration (perceived as length) in a spoken utterance. These variations usually (though not necessarily) extend over parts of the utterance which are longer than a single segment, hence the term “supra-segmental” features, that is, features which stretch above, or beyond, individual sounds.

The suprasegmental characteristics of L2 speech have for a long time been ignored by educators and researchers alike. The former have tended to focus on the segmental system (the vowels and the consonants) of the foreign language, on the assumption that mastering the individual sounds is crucial, if not sufficient, for efficient communication in the L2. The latter have for a long time ignored investigation into L2 prosody, not least because of the lack of sound and consistent methodology for the contrastive study of suprasegmental features in speech. Even some of the most popular L2 learning models, such as the Speech Learning Model

(Flege, 1995, 2007), the Native Language Magnet model (Kuhl, 1991, 1992, 2000), and the Perceptual Assimilation Model (Best, 1995; Best and Tyler, 2007) focus almost exclusively on the segmental level. Most of the early predictions made by the original models have been based on research carried out with learners who acquire the foreign language in a predominantly L2 environment, while later studies (e.g., Piske, 2007; Tyler, 2019) have tried to explore the models' implications for L2 students learning the language through formal instruction in the foreign language classroom in a predominantly L1 environment.

The Speech Learning Model (SLM) focuses on the accuracy with which L2 segments are perceived and claims that this is a major determinant of the accuracy with which those segments will subsequently be produced by the foreign learner. It also claims that cross-language phonetic interference tends to be bi-directional, affecting L2 speech production, but also exerting influence on certain aspects of L1 speech, especially that of learners who have been exposed to the foreign language from an early age. In other words, the elements of the phonetic systems of L1 and L2 co-exist in a "common phonological space", and thus constantly affect and mutually influence each other. Unlike the "Critical Period Hypothesis" (Lenneberg, 1969), SLM maintains that the ability to form new phonetic categories remains active throughout an individual's life span, but in L2 learning it takes time, and correlates with the amount and nature of the input received by the learner. All of the above assumptions and hypotheses of the original SLM model crucially concern the learning of individual speech sounds (segments).

The Native Language Magnet model (NLM) (Kuhl, 1991, 1992, 2000) focuses on early speech perception, suggesting that infants categorize the speech sounds of the mother tongue by creating in their brains a "sound map" – a complex network, or filter, which may then interfere with the acquisition of the phonemes of an L2. The prototype sounds in the sound map then act like magnets and tend to attract similar sounds, so that "initial learning" (of the L1 sounds) "can alter future learning" (that of the L2 sounds) (Kuhl, 2000, p. 11855).

As the name suggests, the Perceptual Assimilation Model of L2 speech learning (PAM / PAM –L2 (Best 1995; Best and Tyler, 2007) aims to explain the way(s) in which the learner's L1 shapes the perception of the sounds of the L2. A learner's success at acquiring phonological contrasts which exist in the L2 but are absent in the L1 ultimately depends on the way in which the L2 phonemes have assimilated to the phonological system of the learner's L1. In other words, the system of the mother tongue influences the perception of the foreign language, and unless phonetic differences which signal phonological contrast in the L2 have been assimilated so as to preserve the dissimilarity, perceptual learning is required in order for the learner to acquire the new L2 phonological category.

Although the main assumptions of the models briefly reviewed above target almost exclusively the acquisition of the segmental system of the respective L2,

they were reviewed because at least some of the predictions which they make could eventually be applied to the learning of suprasegmental features as well. However, as a result of the preoccupation with segmental acquisition, both L2 teaching and L2 research have suffered considerably from the lack of a comprehensive theory and model of L2 suprasegmental, or prosodic, learning.

A quick look at the materials for teaching suprasegmentals as part of foreign language instruction reveals that, if a broad definition of prosody is adopted, then some suprasegmental features do find their place in the curriculum. These include lexical stress and some basic intonation patterns, such as question intonation, the intonation of lists, etc. However, research based on in-depth comparisons between learners' L1 prosodic systems and the prosodic system of English has been relatively sparse (for an overview, see Mennen 2007).

The Autosegmental – Metrical (AM) framework of intonation analysis which was first put forward by J. Pierrehumbert (Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988) seems to have offered a new perspective on the cross-language investigation of suprasegmental features, and in particular on comparative intonation research. As a consequence of the application of the AM approach to the analysis of the intonation in a range of languages (Jun, 2005, 2014), it also exerted its influence on the teaching of (English) intonation to foreign learners. To give just one example, research by Estebas – Vilaplana (2015, 2018) compared the suprasegmental features in the English speech of two groups of Spanish students of English phonetics who were trained using different methodology: one group was taught in the tradition of the British School of intonational analysis, whereas the other was trained using what the author calls TL_ToBI – a version of Pierrehumbert's (1980) system adapted to the needs of teaching English intonation in a distance learning setting. Estebas-Vilaplana found that "students instructed with TL_ToBI produced more native-like intonation patterns" than those instructed with the British School model, "suggesting that a system based on tonal targets and their association to the metrical structure is more helpful", especially for self-tuition purposes (Estebas – Vilaplana, 2015, p. 42).

An important step towards the development of a comprehensive model of L2 prosody acquisition is the L2 Intonation Learning Theory (LILt) (Mennen, 2007, 2015). The theory attempts to offer an extensive account of the major suprasegmental problems experienced by L2 learners, especially those in the area of intonation. Mennen draws an important distinction between phonological representation and phonetic implementation. She hypothesizes that L2 learners first acquire the phonological patterns in the foreign language, and only afterwards try to master the phonetic implementations of those patterns. She therefore insists that, due to this, "a perceptually similar error may in fact have different underlying causes, which can be either difficulties with the phonological structure of the L2 or with its phonetic realization … it is important for teaching purposes to distinguish between phono-

logical and phonetic errors, so that the source of the problem can be addressed in teaching (Mennen, 2007, p. 71).

Mennen distinguishes four major dimensions along which L2 intonation may deviate. The first of these – the systemic dimension – deals with the inventory of structural prosodic elements and their distribution. The categorical elements can be pitch accents, accentual units of different size, or boundary phenomena, as outlined in Pierrehumbert's (1980) original version of the Autosegmental-Metrical theory. This dimension also involves the ways in which structural elements such as pitch accents combine with one another – for example, what combinations of High (H) and Low (L) pitch targets are admissible in a given language. In addition, it also looks at tune – text association (Ladd, 1996, p. 119), that is, the way the tune is mapped onto the segmental string. The second dimension of the LILt model – the realizational, or phonetic, dimension – is concerned with the phonetic implementation of the categorical elements of the system: this may involve the actual alignment of pitch accents, their scaling (i.e., their relative height), and their shape, or slope, e.g., shallow vs. steep rises or falls. The third dimension in Mennen's LILt model is the semantic one: it deals with the ways in which the systemic elements are used to signal intonation functions. The fourth and final dimension of LILt – the frequency dimension – takes into account how often the structural elements are used.

Mennen (2015) presents an overview and many examples, mostly from research on the major spoken varieties of English and a range of (mostly European) languages. In her review of L2 intonation studies, she given ample evidence for the existence of deviations attested along all four dimensions, though those in the second one – the realizational dimension – seem to be the most numerous of all, mostly concerning alignment / timing and scaling. She also admits that the LILt model is far from being unproblematic. For example, the dimensions can interact with one another, or it may be difficult to uniquely ascribe a certain instance of L2 intonational deviance to one of the four dimensions. But in spite of its limitations, LILt constitutes the first model aiming to account for L2 prosodic learning. It will be thus extremely useful to be able to utilize for the first time a model specially developed for characterizing L2 intonation in order to analyze the prosody in the speech of Bulgarian learners of English, albeit with the important caveat that LILt should be treated "as an evolving or 'working' model, which is subject to change when more data are published" (Mennen, 2015, p. 17).

The pilot study reported in the present paper presents a comparative analysis of the speech data obtained from six Bulgarian and six English female speakers. At the time of the recording, the six Bulgarian speakers were aged 19-23 and were undergraduates at Sofia University. They were born and were all living in Sofia. The six English speakers were all born and living in England at the time of recording, they were of comparable age, and were all university undergraduates. Their accent

was judged by two lecturers in English Phonetics to be representative of the kind of pronunciation that has been dubbed "Modern Received Pronunciation", "Southern Standard Pronunciation", etc. – an educated accent virtually devoid of any salient regional pronunciation features. Their accent will be referred to in the rest of the paper as just "Received Pronunciation", or RP. All speakers read and recorded Aesop's fable "The North Wind and the Sun" – a standard text routinely used in phonetic research. The Bulgarian speakers recorded the Bulgarian text – "Северният вятър и слънцето" as well.

The recordings were analyzed, segmented and labelled in Praat using the ToBI (Tone and Break Indices) labelling conventions: for English, we have followed the principles outlined in, e.g., Beckman et al. (2005), whereas for Bulgarian we have adopted the use of the ToBI conventions in Andreeva (2007) and Dimitrova and Jun (2015). Figure 1 shows the waveform, spectrogram, intensity and pitch tracks of the phrase „Северния вятър", along with several labelling tiers. The top labelling tier contains the ToBI pitch accent and phrase accent labels, the second tier from top shows the segmentation into syllables, with the pitch accented syllables transcribed in block capitals, the third tier from top shows the boundaries of the intonation phrase (IP), and the bottom two tiers contain the orthographic Bulgarian text and its English translation.

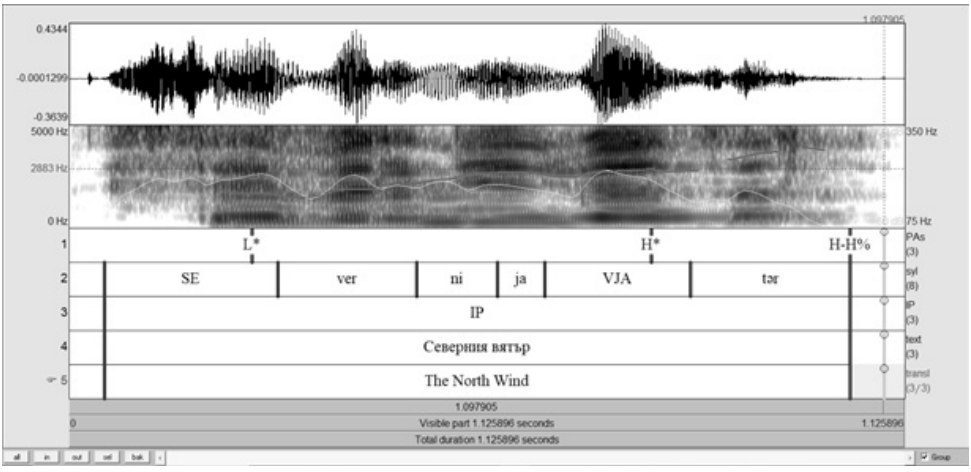

Fig. 1. Illustration of the segmentation and the ToBI labelling of the speech data.

The manually segmented and labelled files were analyzed using dedicated Praat scripts, with the help of which we obtained a number of fundamental frequency (F0) and durational measures. The Long-Term Distributional (LTD) measures in which we were interested were the following:

Fundamental frequency characteristics

- for pitch level – mean and median F0, measured in Hertz (Hz),
- for pitch span – F0 excursion, measured in Hz and converted to semitones (ST)

We analyzed pitch level and pitch span separately because, according to Ladd (1996), the two are characteristics which are partially related, but nevertheless should be viewed as distinct. Pitch level was defined as the overall height of a speaker's voice, whereas pitch span was taken to refer to the range of frequencies typically covered by a speaker. Fundamental frequency excursion (pitch span) was calculated as the difference between the maximum and minimum F0 values in a given intonation phrase; it was measured in Hz, and then converted to ST using the formula given by Reetz (1999).

Temporal characteristics

- mean syllable duration, measured in milliseconds (ms),
- intonation phrase (IP) and pause duration, measured in ms.

In addition, we performed counts of the number of Ips, pauses, stressed and unstressed syllables in the reading of each speaker.

We first compare the results obtained for the Bulgarian (B) and the Bulgarian English (BE) readings of the undergraduate students, in search for differences between the prosody of their L1 and L2. In addition, we also compare the English readings of the Bulgarian participants in the present experiment with the readings of native British English (RP) speakers in an attempt to shed light upon the way(s) in which the suprasegmental characteristics of Bulgarian English differ from those of native English modern RP pronunciation. We expect that both comparisons will yield statistically significant results: in the first case, we are comparing the L1 with the L2 performance of the same group of speakers, whereas in the second case, the comparison is between groups of L1 and L2 speakers which are similar in terms of group size, speaker age, gender, and education level. In a previous study (Dimitrova, 2019) we reported the results from a similar analysis of the recordings of "The North Wind and the Sun" for native British and American English which have been made publicly available by the International Phonetic Association (IPA). In that study, the caveat was made that those results could only be considered tentative and should be interpreted with caution because the native speaker data came from a single British (RP) and American (GA) speaker. In addition, there was little information about the age, education, etc. of those two English speakers (Dimitrova, 2019, p. 130). In the present paper, we comment on the representativeness of those data and compare the British English IPA recording with our newly obtained data from young female British English undergraduates.

The results which we obtained by measuring fundamental frequency (F0) are shown in Table 1, where RP 1 stands for the results for the group of 6 undergraduate native RP speakers, and RP 2 indicates the results for the single native RP female

reader of the official IPA version of the "North Wind and the Sun" fable reported in Dimitrova (2019).

Table 1. Fundamental frequency (F0) results (values rounded to the nearest whole in Hz; pitch span results shown in semitones ST). The features for which significant differences were found are marked with *.

	Bulgarian (B)	Bulgarian English (BE)	RP 1	RP 2
*Mean F0 (Hz)	226	220	199	188
*Median F0 (Hz)	223	215	197	179
*StDev F0 (Hz)	28.8	24.6	17.8	n.a.
*Min F0 (Hz)	177	185	169	135
*Max F0 (Hz)	297	275	240	269
*Pitch span (ST)	9.1	6.7	6.1	11.7

The F0 results were then also analysed statistically: Linear Mixed Models (LMMs) with the respective measure as dependent variable, "Speaker" as random factor, and "Language" (Bulgarian – B, Bulgarian English – BE and British Received Pronunciation – RP 1) as fixed factors were calculated, and Post-hoc tests were carried out.

For Mean F0, the results obtained for the Bulgarian (B) and the Bulgarian English (BE) readings of the female wundergraduates were very similar: about 226 Hz for the Bulgarian reading of "Северният вятър и слънцето", and about 220 Hz for Bulgarian English (the same six female Bulgarian speakers read the fable "The North Wind and the Sun" in English and "Северният вятър и слънцето" in Bulgarian). This similarity is altogether an unsurprising result given that the Bulgarian and the Bulgarian English readings were produced by the same speakers. The respective result for standard British English (Received Pronunciation, or RP) is 199 Hz – for the group of undergraduates, and 188 Hz – for the single speaker. Dimitrova (2019) points out that the significance of the mean and median F0 differences between the native and the non-native readings can only be demonstrated by an analysis which involves comparisons among groups speakers whose ages and backgrounds are comparable. The current results show that difference between Bulgarian and Bulgarian English, on the one hand, and RP 1, on the other hand, is

statistically significant [F (2, 25.43) = 45.5; p<0.0001]. (All statistical tests compare the Bulgarian with the Bulgarian English data, and the Bulgarian English with the RP 1 data; the RP 2 data, as already pointed out, come from a single speaker and any results from statistical tests would be questionable).

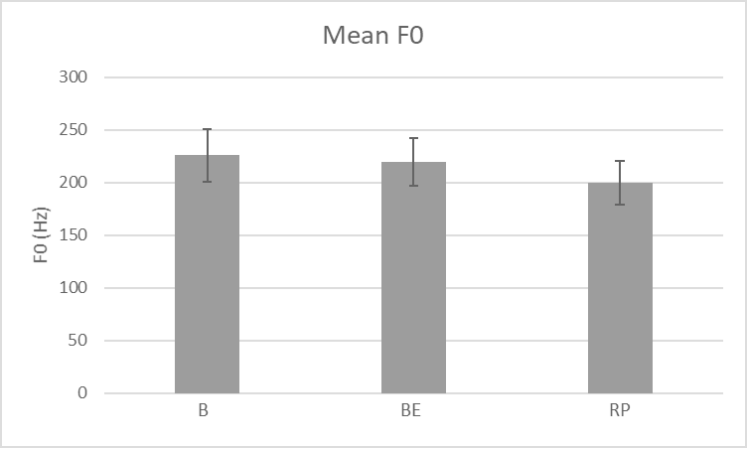

Figure 2. Mean F0 values and standard deviations (in Hz) for Bulgarian (B), Bulgarian English (BE) and British English Received Pronunciation 1 (RP).

The results for Median F0 are similar to those obtained for Mean F0, namely, 223 Hz for Bulgarian, 215 Hz for Bulgarian English, 197 Hz for the group of RP 1 speakers, and 179 Hz for the single RP 2 speaker. Again, the difference between Bulgarian and Bulgarian English, on the one hand, and RP 1, on the other hand, is statistically significant [F (2, 25.42) = 40.1; p<0.0001].

It is worth noting, at this point, that the RP 2 speaker (the speaker who read the representative IPA version of the fable) used lower Mean and Median F0 which, together with the considerably lower Minimum F0 measured for this speaker, gives us grounds to suppose that this was an older speaker than the undergraduates who comprised our current RP reference speaker groups. The F0 minimum values for Bulgarian, Bulgarian English and RP 1 were all significantly different from each other as well [F (2, 25.78) = 4.6837; p = 0.0184].

The Standard Deviation of F0 (StDev) also differed significantly [F (2, 23.82) = 27.0339; p<0.0001]. It was biggest in Bulgarian spoken as a mother tongue – 28.8 Hz, and smallest in native RP speech – 17.8 Hz. In terms of this feature, the value obtained for Bulgarian English (24.6 Hz) was again closer to Bulgarian, rather than to English. This result suggests that there is more F0 variation in both Bulgarian and Bulgarian English speech, than in native RP pronunciation, which does not

support the frequent observation that English intonation "goes up and down all the time", unlike Bulgarian intonation (see more on this below).

The Maximum F0 values which we obtained in our analyses were the highest for the Bulgarian speakers in their L1 Bulgarian readings (297 Hz). In the reading of the English text "The North Wind and the Sun", the Bulgarians used a maximum F0 of 275 Hz, which is considerably higher than the 240 Hz maximum used by the group of native RP 1 speakers. This may suggest the existence of an interesting suprasegmental difference between native Bulgarian and non-native Bulgarian English speech on the one hand, and native English RP speech, on the other hand. Tests show the difference to be statistically significant [$F(2, 25.3) = 43.5734$; $p<0.0001$].

The pitch span results which we obtained are as follows: the average span (F0 excursion) for the six Bulgarians speaking their mother tongue is 9.1 semitones, and only 6.7 semitones in Bulgarian English. For the British RP 1 group, the span is even narrower – 6.1 semitones. Statistical tests show that with respect to this suprasegmental feature, Bulgarian spoken as L1 is significantly different from Bulgarian English on the one hand, and from British English on the other hand [$F(2, 22.34) = 22.7072$; $p<0.0001$]. The highest measure for this characteristic was 11.7 ST, and it was obtained for the individual RP 2 speaker: this seems an idiosyncratic feature of the speaker. The pitch span results (in semitones) for Bulgarian, Bulgarian English and RP, along with the respective standard deviation values are shown in Figure 3.

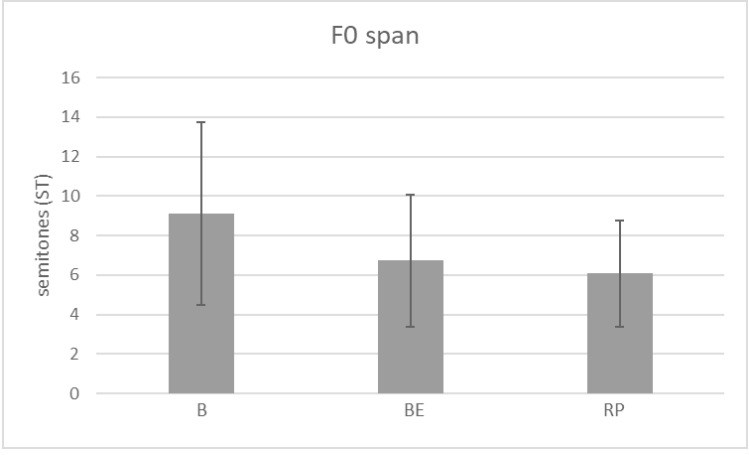

Figure 3. Pitch span measures (and standard deviations) between Bulgarian (B), Bulgarian English (BE) and British English (RP) (in semitones ST).

Our results for the L1 (Bulgarian vs. British English RP) readings of the two groups of speakers are in agreement with earlier results reported by Andreeva et al. (2014), who found significant differences between two language groups of speakers: in this study, the speakers of Germanic languages (German and English) used narrower pitch span (and lower pitch maxima) than the speakers of Slavic languages (Bulgarian and Polish), supporting the hypothesis that "linguistic communities tend to be characterized by particular pitch profiles" (Andreeva et al., 2014, p. 776).

Mennen (2007, pp. 63-64) offers an insightful discussion of current evidence on the influence of speakers' language background on their pitch range, or span. Drawing on Ohara's (1992) study of gender-dependent use of pitch levels in English vs. Japanese, she observes that "It is thought that cultures or languages have their particular 'vocal image', which reflects socio-culturally desired personal attributes and social roles, and that speakers choose a pitch (within their anatomical/physiological range) that approximates the vocal image they want to project" (Mennen, 2007, p. 64). Ladd (1996) likewise discusses such pitch changes, considering them variations in terms of pitch level (overall pitch height) and pitch span (frequency range). Our current findings add further evidence to support the importance of these parameters as already emerging in studies such as Mennen (2007, 2015), Zimmerer et al. (2014, 2015) and Andreeva (2016), among others.

However, the use of narrower F0 span in English than in Bulgarian does not corroborate the popular impressionistic observation by native English speakers that Bulgarian English speech sounds "flat" and "monotonous", and even "dull" and "uninterested". Conversely, Bulgarians often tend to perceive native English intonation as "affected" or "exaggerated". A similar observation about German listeners who think that English intonation is "over the top" has been noted by Eckert and Laver (1994, reported in Mennen, 2007, p. 64).

It may well be the case that the above impressions are due not to long-term F0 characteristics but to the use of certain pitch accents or "tones" which are absent from the pitch accent inventory of the foreign learner's L1. Such an account will be in line with the "systemic dimension" of Mennen's model and the inventory and distribution of the respective pitch accents, and probably also with the "frequency dimension" which specifies the frequency with which the structural elements are used. For English, we can tentatively hypothesise that one such pitch accent could be the "fall-rise" tone.

It may also be worth applying a somewhat different approach to pitch range modelling, as suggested by Patterson and Ladd (1999), who measured F0 values not just at the highest and the lowest point in an intonation phrase, but used other well-established landmarks such as initial peaks, as well as other accent peaks, valleys, and final lows in a sentence. It is clear that, with respect to F0 span, further research is needed to account for the seeming discrepancy between acoustic measurements and auditory impressions.

The temporal characteristics in which we were interested were mean intonation phrase duration, which was measured in milliseconds (ms), overall number of intonation phrases and pauses, speech tempo, measured in number of syllables per second (not reported in the present paper), mean syllable duration, and mean pause duration, the latter two also measured in milliseconds. These are shown in Table 2.

Table 2. Temporal measures for Bulgarian, Bulgarian English and RP. Duration measures are given in milliseconds (ms). The features for which significant differences were found are marked with *.

	Bulgarian (B)	Bulgarian English (BE)	Received Pronunciation (RP)
*Mean intonation phrase (IP) duration (ms)	1181	989	1621
*Number of Ips	140	178	98
*Number of pauses	68	88	55
*Mean syllable duration (ms)	137.7	203.6	183.8

The results obtained for the temporal features in which we were interested were then also analysed statistically: Linear Mixed Models (LMMs) with the respective measure as dependent variable, "Speaker" as random factor, and "Language" (Bulgarian – B, Bulgarian English – BE and British Received Pronunciation – RP) as fixed factors were calculated, and Post-hoc tests were carried out.

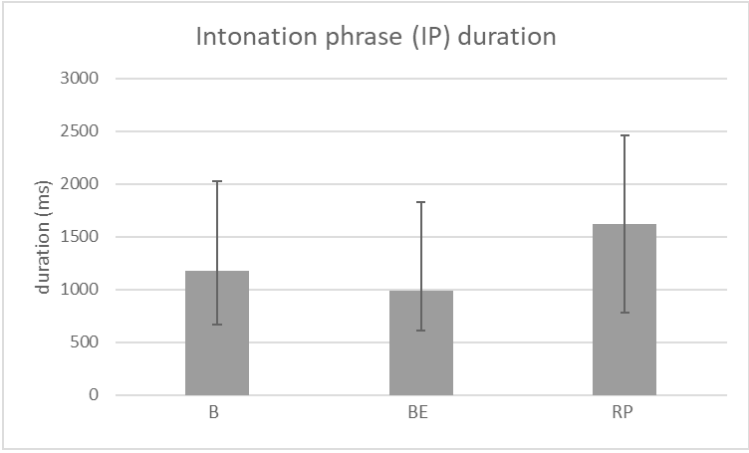

Figure 4. Intonation phrase (IP) duration differences (and standard deviations) between Bulgarian (B), Bulgarian English (BE) and British English (RP) (in milliseconds).

Intonation phrase duration differences between Bulgarian, Bulgarian English and RP were all statistically significant [$F(2, 26.33) = 37.3680$; $p<0.0001$]. However, the results for the Bulgarian text are not directly comparable with those for the English text. The intonation phrases produced by the Bulgarian speakers of English were much shorter – they were only about 60% of the duration of the Ips produced by the native RP speakers. Also, the overall number of Ips in the reading of the English text by the Bulgarians was much larger (178 vs. only 98 used by the RP speakers), and the number of pauses was also larger (88 in Bulgarian English vs. 55 in RP). These findings seem to confirm another well-known empirical observation, namely, that foreign learners of a language tend to produce shorter chunks of speech, and also tend to pause more often when speaking in a foreign language. And although this is commonly explained with the need for more planning time, our data point to the use of shorter speech chunks by foreign learners in reading tasks as well.

Mean syllable durations likewise differed significantly in Bulgarian, Bulgarian English and RP [$F(2, 10) = 106.5554$; $p<0.0001$], whereas mean pause duration was not significant and has therefore not been included in Table 2. Finally, the difference between Bulgarian English and RP speakers in terms of the number of syllables in the English text on which they put stress was also significantly different: the native speakers stressed on average 45 syllables, while the Bulgarian speakers of English put stress on 58 syllables, which is about 23% more in comparison with the native RP speakers [$F(1, 5) = 78.6691$; $p=0.0003$]. This result confirms yet another well-known empirical observation about the suprasegmental characteristics of Bulgarian English, namely, that non-native speakers stress more words. According to another view, non-native learners fail to de-stress words which constitute old information in connected speech: indeed, in the Bulgarian English reading of the text, there were on average 86 unstressed syllables, compared with 99 unstressed syllables found in the reading of the RP speakers. (The total number of syllables in the English text amounted to a total of 144, with the word "traveler" considered 3-syllabic irrespective of whether a given speaker pronounced it with 3 or with only 2 syllables. The Bulgarian text of the fable, on the other hand, comprised 200 syllables.)

The suprasegmental characteristics analyzed in the present paper are primarily related to the phonetic, or realizational, dimension of Mennen's L2 Intonation Learning theory (LILt). The next step in our ongoing research on the prosodic characteristics of Bulgarian English will be to focus on the "systemic dimension" of the LILt theory, and to use ToBI for the labelling of pitch accents, phrase accents and boundary tones. We will also draw comparisons between the distribution of those structural elements in Bulgarian vs. Bulgarian English, on the one hand, as well as in L1 vs. L2 English, on the other hand. But it should be borne in mind that, like many other learners of English world-wide, Bulgarians are exposed to a

range of native English accents whose suprasegmental characteristics are bound to influence their Bulgarian English pronunciation. A major difficulty in this respect turns out to be the relatively sparse comparative research available on the suprasegmental phonetics and phonology of English accents. Further analyses of the phonetic implementation of the categorical phonological elements comprising the "systemic dimension" of LILt such as, for example, the timing of pitch accents are also necessary. Finally, comparative investigations of the functioning of pitch accents, phrasal accents and boundary tones, as well as of their frequency of use, that is, of the "semantic" and "frequency" dimensions of LILt also need to be carried out.

In conclusion, the L2 Intonation Learning theory seems to provide a sound starting point for L2 prosody research but, as noted by Mennen herself, it should be treated as "an evolving or 'working' model, which is subject to change when more data are published" (Mennen, 2015, p. 17).

Acknowledgments

This research was supported by the research project "Prosodic aspects of Bulgarian in comparison with other languages with lexical stress", funded by Bulgarian National Science Fund, project № Кп-06-Н40/11 – 12.12.2019.

Bibliography

Andreeva, B. (2007), *Zur Phonetik und Phonologie der Intonation der Sofioter-Varietät des Bulgarischen*, PhD dissertation, PHONUS 12, Institute of Phonetics, University of the Saarland, Saarbrücken.

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F.,
Jügler, J., Oleskowicz-Popiel, M. and Trouvain, J. (2014), Comparison of pitch range and pitch variation in Slavic and Germanic languages, in Campbell, N., Gibbon, D. and Hirst. D. (Eds.), *Prodeedings of the 7th International Conference on Speech Prosody*, TCD, Dublin, Ireland, pp. 776–780.

Andreeva, B. (2016), Contrastive Prosody: Bulgarian vs. German, Kumulative Habilitationsschrift, Philosophischen Fakultät der Universität des Saarlandes, Saarbrücken.

Andreeva, B., Möbius, B., Demenko, G., Zimmerer, F. and Jügler, J. (2015), Linguistic measures of pitch range in Slavic and Germanic languages, in *Proceedings of Interspeech 2015. Speech beyond Speech: Towards a Better Understanding of the Most Important Biosignal.* ISCA, Dresden, pp. 968–972.

Beckman, M., Hirschberg, J. and Shattuck-Hufnagel, S. (2005), The original ToBI system and the evolution of the ToBI framework, in Jun, S.-A. (Ed.), *Prosodic Typology*, Cambridge University Press, Cambridge, pp. 9–54.

Best, C. (1995), A direct realist view of cross-language speech perception, in Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Timonium, MD: York Press, Baltimore, pp. 171–204.

Best, C. and Tyler, M. D. (2007), Non-native and second-language speech perception, in Bohn, O.-S., Munro, M. J. (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege.* J. Benjamins, Amsterdam, pp. 13–34.

Crystal, D. and Quirk, R. (1964), *Systems of Prosodic and Paralinguistic Features in English*, Mouton, The Hague.

Crystal, D. (1969), *Prosodic Systems and Intonation in English*, Cambridge University Press, Cambridge.

Dimitrova, S. and Jun, S.-A. (2015), Pitch accent variability in focus production and perception in Bulgarian declaratives, in Wolters, M., Livingstone, J., Beattie, B., Smith, R., MacMahon, M., Stuart-Smith, J., and Scobbie, J. (Eds.), *Proceedings of the 18th ICPhS*. University of Glasgow, Glasgow, paper number 832, 5 pp., available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0832.pdf (accessed 10 October 2021)

Dimitrova, S. (2019), Some prosodic characteristics of Bulgarian English, in Slavova, E., Bagasheva, A., Slavova, K., Tincheva, N., Muharska, R., Ishpekova, R. and Catalan, Z. (Eds.), *Traditions and Transitions,* Vol. 2, St Kliment Ohridski University Press, Sofia, pp 125–138.

Jun, S.-A. (Ed.), (2005), *Prosodic Typology: the Phonology of Intonation and Phrasing*, Oxford University Press, Oxford.

Jun, S.-A. (Ed.) (2014), *Prosodic Typology II: the Phonology of Intonation and Phrasing*. Oxford University Press, Oxford.

Eckert, H. and Laver, J. (1994), *Menschen und ihre Stimmen: Aspekte der vokalen Kommunikation*, Psychologie Verlags Union, Weinheim.

Estebas – Vilaplana, E. (2015), The learning of English intonation by Spanish speakers in a distance education environment, *PTLC 2015*, UCL, London, pp. 39–43.

Estebas – Vilaplana, E. (2018), The teaching and learning of L2 English intonation in a distance education environment: TL_ToBI vs. the traditional models, Linguistica 57 (1), 73–91.

Flege, J. E. (1995), Second-language speech learning: Theory, findings, and problems, in Strange, W. (Ed.), *Speech perception and linguistic experience: Issue in cross-language research*. Timonium, MD: York Press, Baltimore, pp. 229–273.

Flege, J. E. (2007), Language contact in bilingualism: Phonetic system interactions, in Cole, J. and Hualde, J. I. (Eds.), *Laboratory Phonology 9*, Mouton de Gruyter, Berlin, pp. 353–380.

Kuhl, P. K. (1991), Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories. Monkeys do not., *Perception and Psychophysics* 50 (2), pp. 93–107.

Kuhl, P. K. (1992), Speech prototypes: Studies on the nature, function, ontogeny and phylogeny of the "centers" of speech categories, in Tohkura, Y., Vatikiotis-Bateson, E. and Sagisaka, Y. (Eds.) *Speech, Perception, Production and Linguistic Structure*, IOS Press, Tokyo, Japan, pp. 240–264.

Kuhl, P. K. (2000), A new view of language acquisition, *PNAS – Proceedings of the National Academy of Sciences of the United States of America, October 24, 2000*, vol. 97 (22),

pp. 11850-11857; available at: https://doi.org/10.1073/pnas.97.22.11850 (accessed 28 September 2021).
Ladd, D. R. (1996), *Intonational Phonology.* Cambridge University Press, Cambridge.
Lehiste, I. (1970), *Suprasegmentals*, MIT Press, Cambridge, MA.
Lenneberg, E. (1969), *Biological Foundations of Language*. Wiley, New York.
Mennen, I. (2007), Phonological and phonetic influences in non-native intonation, In Trouvain, J. and Gut, U. (Eds.), *Non-native Prosody. Phonetic Description and Teaching Practice (Trends in Linguistics. Studies and Monographs (TiLSM) 186.* Mouton de Gruyter, Berlin, pp. 53–76.
Mennen, I. (2015), Beyond segments: towards an L2 intonation learning theory (LILt), in Delais-Roussarie, E., Avanzi, M. and Herment, S. (Eds.), *Prosody and languages in contact: L2 acquisition, attrition, languages in multilingual situation*, Springer Verlag, Berlin, pp. 171–188.
Ohara, Y. (1992), Gender dependent pitch levels: A comparative study in Japanese and English, in Hall, K., Bucholtz, M. and Moonwomon, B. (Eds.), *Locating Power: Proceedings of the Second Berkeley Women and Language Conference* 2, Berkeley, pp. 478–488.
Patterson, D. and Ladd, D. R. (1999), Pitch range modelling: linguistic dimensions of variation, in ICPhS 99: *Proceedings of the 14th International Congress of Phonetic Sciences,* University of California, Berkeley, San Francisco, pp. 1169–1172.
Pierrehumbert, J. (1980), *The Phonology and Phonetics of English Intonation*, PhD dissertation, MIT Press, Cambridge, MA.
Pierrehumbert, J. and Beckman, M. (1988), *Japanese Tone Structure*, MIT Press, Cambridge, MA.
Piske, T. (2007), Implications of James E. Flege's research for the foreign language classroom, in Munro, M. J. and Bohn, O.-S. (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege*, John Benjamins, Amsterdam, pp. 301–314.
Reetz, H. (1999), *Artikulatorische und akustische Phonetik.* Wissenschaftlicher Verlag, Trier.
Tyler, M. D. (2019), PAM-L2 and phonological category acquisition in the foreign language classroom, in Nyvad, A. M., Hejná, M., Højen, A., Jespersen, A. B. and Sørensen, M. H. (Eds.), *A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*, Aarhus University, Aarhus, pp. 607–630.
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, in Campbell, N., Gibbon, D. and Hirst, D. (Eds*.), Proceedings of the 7th Speech Prosody Conference, Trinity College Dublin, Ireland*, International Speech Communication Association, Dublin, pp. 1037–1041.
Zimmerer, F., Andreeva, B., Jügler, J. and Möbius, B. (2015), Comparison of pitch profiles of German and French speakers speaking French and German, in Wolters, M., Livingstone, J., Beattie, B., Smith, R., MacMahon, M., Stuart-Smith, J., and Scobbie, J. (Eds.), Proceedings of the 18th ICPhS. University of Glasgow, Glasgow, paper number

615, 5 pp., available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0615.pdf (accessed10 October 2021).

Assoc. Prof. Snezhina Dimitrova, PhD
snezhina@uni-sofia.bg
Sofia University „St. Kliment Ohridski“
15 Tzar Osvoboditel Blvd., Sofia 1504
Bulgaria

Доц. д-р Снежина Димитрова
snezhina@uni-sofia.bg
Софийски университет „Св. Климент Охридски“
бул. „Цар Освободител“ 15, София 1504
България