

PITCH RANGE OF INTONATION CONTOURS IN ENGLISH CZECH

JAN VOLÍN, DAMIEN GALEONE AND WESLEY JOHNSON

ABSTRACT

Pitch range is believed to code important information that is indispensable for correct decoding of spoken messages. Previous research found differences in pitch variation across languages like English, French, Bulgarian, Polish, Czech and German. In addition, differences in pitch range of foreign-accented and native speech were found in various types of speech material. In the present study a sample of sixteen English and American men and women produced recordings of spoken texts consisting of eight paragraphs taken from Czech news broadcasts. Manually corrected F0 tracks provided a possibility to extract four measures of F0 distributional dispersion in order to map global intonational habits of Anglophone learners of Czech as a foreign language. The extracted values were compared with reference values from earlier studies. The results in all four measures indicate that foreign accented Czech is spoken with a pitch range that is narrower than that of English and often even narrower than that of native Czech. Considering results of similar, albeit smaller, studies done earlier, we would attribute our findings to implicit uncertainty in the use of the foreign language, rather than to overcompensation.

Key words: speech melody, pitch range, Czech-accented English, English-accented Czech, F0 contours.

1. Introduction

Traditionally, learners of foreign languages aspired at proficiency comparable with that of the native users of the language. Common sense would have it that the educated native speakers' mastery was the appropriate model for a determined language learner. Current teaching insists less on perfection in pronunciation and sets more realistic goals for the majority of L2 learners. The overwhelming trend is to subscribe to the concept of comfortable intelligibility. Although this is a praiseworthy approach, we are far from understanding where exactly intelligibility starts and ends, and, above all, what makes it comfortable. Abercrombie's indication of speech "which can be understood with little or no conscious effort on the part of the listener" (Abercrombie, 1956: 37) sounds very

reasonable but might not be easy to specify in measurable features. In other words, comfortable intelligibility is a reasonable concept in common classrooms of mass education, but will not provide any useful guidance in empirical research unless comfort on various levels of (un)consciousness can be measured.

There is a relatively long tradition in linking intelligibility to segmental phonemic issues. Teachers are often trained to teach pronunciation of minimal pairs like *peak* and *pick*, *light* and *late* or *mouse* and *mouth*. That is by no means wrong, but the belief that such minimal pairs encapsulate the phonetics of the foreign language and that their training is sufficient to acquire nativelike accent is utterly misguided. At the current stage of our knowledge we should presume that any element of pronunciation can contribute to listeners' discomfort (Anderson-Hsieh, Johnson & Koehler, 1992; Derwing, Munro & Wiebe, 1998; Jilka, 2000; Derwing & Rossiter, 2003; Field, 2005; Kang, Rubin & Pickering, 2010) and, therefore, various features of foreignness should be held suspect of disrupting the smooth flow of perceptual processing. All aspects of sound patterning in speech may produce mismatches between the expected forms and the incoming acoustic signal (cf. Grossberg, 2003). These are likely to activate additional cognitive resources and, subsequently, put strain on vital brain capacities like attention or working memory (Van Engen & Peelle, 2014). Therefore, we consider it legitimate to expand our understanding of foreign accentedness into the area of speech prosody. Our concern in the present study is speech melody or intonation in the narrow sense of the word.

From a lay perspective the functions of intonation might seem purely ornamental. This illusion might be caused by the fact that intonation is not easy to access for conscious evaluation. However, intonologists or conversation analysts can provide abundant evidence concerning the importance of melodic events in speech for effective communication. Without clear prominences and phrasal breaks the speech becomes muddled, difficult to follow, and essential pragmatic and affective messages like friendliness or willingness to cooperate might be absent (e.g., Gilbert, 2014).

Volín and Poesová (2016) carried out an experiment in which they measured reaction times as indicators of the ease of cerebral processing on the part of listeners. They used semantically unpredictable utterances recorded by native speakers of English and by Czech learners of English. Test items were also created by hybridizing the F0 tracks: melodies produced by native speakers were implanted on the Czech-produced utterances and vice versa. Thus, there were four conditions for each of the utterances: (1) native English speech with native English intonation; (2) native English speech with Czech English intonation; (3) Czech English speech with native English intonation; (4) Czech English speech with its specific intonation. Conditions 1 and 4 were the original recordings (resynthesized without changes in order to equalize the technical quality of the sound), while conditions 2 and 3 were hybridized. The results showed that this plain swapping of F0 tracks had an impact on reaction times and that unaltered Czech English was the most difficult to process (Volín & Poesová, 2016). Unfortunately, the authors did not analyse the differences between the melodies in detail, but made it clear that one of the noticeable facts was the difference in the span of melodic movements.

Multiple attempts to describe melodic events in speech have produced large numbers of descriptive methods. The latest advances in speech technologies (speech synthesis, automatic speech recognition) seem to suggest that descriptions of speech melodies

(i.e., intonation in the narrow sense of the word) can be either cognitively accessible or technologically applicable but not both. Speech recognizers and synthesizers use massive computational algorithms that can produce the desired results but, unfortunately, cannot be turned into explanations of how speech melodies function. On the other hand, statements like “rising tune” or “monotonous melody”, which would make sense to most people, are too vague to be used in rigorous intonation modelling. Most linguists, psychologists and language teachers would like to see some sort of a compromise: a descriptive system (probably an open one, without a fixed preconceived inventory) that allows for a clear link between perceptual categories and specific melodic events (definable by specific physical properties).

One of the ways to describe melodic events is to consider the minimal building blocks of the phrase tunes (termed, e.g., pitch accents, tones), the pitch range in which they are produced, and the combinatory and distributional context in which they are used. Of this triad (*pattern – range – phonosyntax*) we will focus on the second aspect. For the purpose of detailed intonological description of local melodic events, Ladd (2008) proposes to consider pitch range a two-dimensional construct, with the dimensions of level and span. Level is known from older terminology as register, but this label is perhaps too easily confused with a long-term voice setting (bass, baritone, mezzosoprano, etc.) rather than a parameter that speakers change even within an utterance (e.g., the prosody of parenthetical clauses).

Nevertheless, our study focuses on more global attributes of speech so the dimension of level will not be investigated. We will address the question raised, for instance, by Hirst and Di Cristo, who wondered if various languages could be spoken with different overall pitch ranges (Hirst & Di Cristo, 1998: 42). At the time of writing their survey of intonation systems, they did not have any data to answer the question. This is because the seemingly simple matter of global pitch range can be meaningfully addressed only if the prerequisite speech samples are collected under comparable conditions and are of sufficient size. Comparable conditions are necessary since there might be differences in pitch range settings across various speaking styles, but also due to the fact that pitch range signals important affective messages that reflect the conditions under which the speech is produced (e.g., Patterson, 2000; Scherer, 2003). A sufficient size of the sample neutralizes the fact that individuals may differ in their habitual pitch range quite substantially. This should be also observed in our current data.

Studies comparing pitch range across languages do exist. Hirst himself, for instance, tested a rather complex set of F0 descriptors and successfully differentiated 10 English speakers from 10 French speakers (Hirst, 2003), yet it seemed that global pitch range did not play any role. He later added 10 Chinese speakers to demonstrate the merits of his method (Hirst, 2013). The report of Keating and Kuo (2012) concerning the difference between English and Chinese is also inconclusive with regard to pitch range, and, importantly, cautions researchers about the influence of the type of speech material. Mennen and her colleagues found that German spoken texts were produced with narrower pitch range than comparable texts in English (Mennen, Schaeffler & Docherty, 2007). Eight authors participated in a project to measure pitch ranges in four languages: Bulgarian, English, German and Polish (Andreeva et al., 2014). They used continuous texts read out by speakers on request. Interestingly, their results did not confirm the difference found

by Mennen and colleagues for English and German, but found significant differences between the two Slavic and two Germanic languages. In general, Polish and Bulgarian displayed greater pitch variation than English and German. A potential problem might stem from the fact that the group of authors did not record their own speech material. Instead, they used already existing corpora from which they selected speakers based on undisclosed criteria.

Volín, Poesová and Weingartová (2015) used an extensive sample of data to provide reliable reference values for English and Czech read monologues. They used texts of news bulletins read out by 32 professional newsreaders from national radio stations. Their results will be used in order to put the outcome of the present study in perspective.

It is only natural that the question of the overall pitch variation has entered the field of foreign or second language (in this study we use L2 for both) acquisition. If languages differ in the mean pitch range used, how will this fact influence the accented speech of L2 learners? Jun and Oh asked four learners and two native speakers to read a set of 40 specially constructed sentences. Their objective was to investigate the acquisition of Korean intonation by speakers of American English, placing special emphasis on phrasing. Among other things pitch range produced in foreign-accented Korean was narrower (Jun & Oh, 2000). The possibility of L2 learners using narrower pitch ranges is further corroborated by Lee (2014) who recorded eight Anglophone speakers learning French. Although he only measured phrase-final rises, he found that the learners spoke with narrower pitch range.

2. Method

Sixteen Anglophone speakers of Czech (eight women and eight men) were asked to read out a news bulletin originally broadcast on Czech Radio (a national broadcaster). They were all resident in Prague and spoke Czech at the levels of B1 to C1 of the CEFRL. The length of their residence in the Czech Republic varied from 1 to 20 years but this did not correlate with their L2 proficiency.

The subjects were given a print-out of the text of six paragraphs and were given sufficient time to get acquainted with it. The recordings were made in a sound treated room with a condenser microphone connected directly to a computer sound card. The recordings were saved in an uncompressed format at a 32-kHz sampling frequency using a 16-bit resolution. The spoken text was divided into breath-groups with a constraint on the excessive breathing producing breath-groups too short. Any breath-group shorter than 1.2 sec. was left with the adjacent one (preceding or following considering the prosodic closeness). Each speaker produced about 55 breath-groups with a mean duration of 5.2 sec.

F0 tracks were extracted in the speech analysis software Praat (Boersma & Weenink, 2014) with the autocorrelation algorithm. Individual values were taken in 10-ms steps and the contour was smoothed by a 10-Hz filter. Subsequently, all 872 contours were manually corrected since some of them contained octave jumps, spurious periodicities in voiceless regions or missing F0 values in soft breathy syllables. All contours were also interpolated through the voiceless regions to emulate the human percepts, which are also uninterrupted by the voiceless consonants (cf. Volín & Bartůňková, 2015).

Four correlates of pitch range, which are termed measures of data dispersion in descriptive statistics, were computed. In this study they will be referred to as: variation range (VAR), 80% percentile range (PER), interquartile range (IQR) and standard deviation (SD). Since comparisons of results across various studies are desirable, our understanding of the measures will be explained in more detail.

By *variation range* the span over the values or the distance between the lowest and the highest value (minimum – maximum) is understood. Although this measure is very popular, its disadvantage lies in the fact that it hinges on two extreme measure only, hence it is quite prone to error. We overcame this disadvantage in our study by computing the arithmetic mean of more than fifty breath-group ranges for each speaker. Thus, the range for a given speaker is not dependent on just two values, but on more than a hundred measurements. (Similarly, when a maximum or minimum is mentioned further in the text, it is not the absolute maximum of the given speaker, but the mean maximum averaged across all the breath-groups produced by that speaker).

Percentile range is a general term for distances between specified points in ordered rows of values. Even though, in theory, any two points can be selected, researchers use either memorable or otherwise justifiable numbers. In our study we will report the distance between the 10th and the 90th percentile. This measure is also mentioned as a possibility by Patterson and Ladd (1999) and used by Mennen et al. (2007) and Lee (2014). This 80 % percentile range will be referred to as PER for short. Again, the PER is computed for each breath-group produced by a speaker and the arithmetic mean for that speaker is reported.

Interquartile range (IQR) expresses the distance between the 25th and the 75th percentile in an ordered row of values. In other words, the lowest and the highest quarters of the data are disregarded and the variation range of the medial half of the ordered data is measured. This measure was also used in the above-mentioned study by Andreeva and colleagues (2014). The advantage of this measure is its stability: it is not influenced by extreme values. On the other hand, it can also lack important specific information if the domain of a phonetic feature happens to be outside the range of typical values.

Standard deviation is, in a sense, an improved concept of mean deviation. It also approximates the dispersion of the values around the arithmetic mean but, for the sake of generalizability, it weighs smaller distances from the mean differently from bigger distances, and takes into account the size of the sample from which it is calculated. Its use is widespread although it seems that it is sometimes forgotten that SD is designed primarily for symmetrical data. F0 values are usually asymmetrical – skewed to the right. Unlike ranges, SD values will be reported in Herz (Hz), which, relative to semitones (ST), is an exponential unit. Therefore, male and female results must be presented separately.

3. Results

The important aspect of the present study is the consideration of the reference values of both native Czech and native English. These were provided by Volín, Poesová and Weingartová (2015) and pertain to the same type of spoken texts. Figures 1 and 2 present the reference values together with the means of the range measurements obtained from the current material.

Figure 1. Mean values of three types of F0 range measures: variation range, 10–90 percentile range and interquartile range (see text) for English, Czech and Czech spoken by native speakers of English.

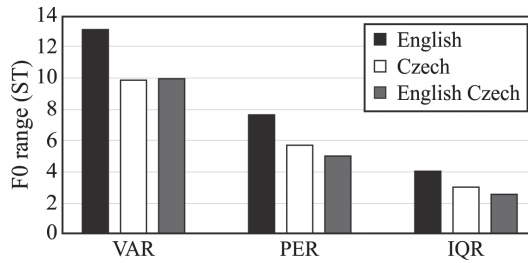
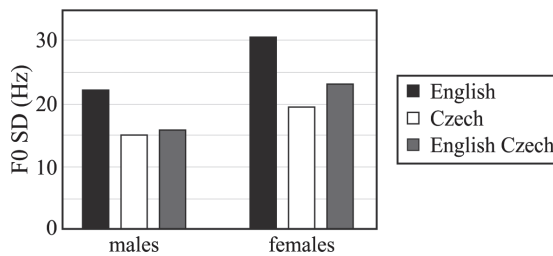


Figure 2. Mean values of F0 standard deviations for male and female speakers of English, Czech and English-accented Czech.



Since all of the differences between the Czech and English reference values were ascertained as highly statistically significant with $p < 0.001$ (Volín, Poesová & Weingartová, 2015), only t-tests for referential values were calculated to see whether the English-accented Czech (E-Cz) differed significantly. With regard to VAR, E-Cz was narrower than native English by 3.7 semitones, which was found highly significant: $t(15) = 7.39$; $p < 0.001$. Contrary to that the difference from native Czech was highly insignificant as it amounted to less than one tenth of a semitone ($p > 0.88$).

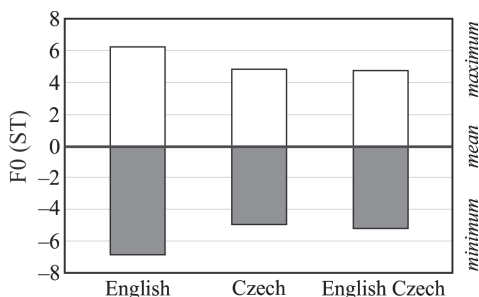
For PER the difference between native English and E-Cz was 2.6 ST and it was highly significant: $t(15) = 8.61$; $p < 0.001$. The difference between native Czech and E-Cz is only 0.68 ST but even this small number reached statistical significance since we were dealing with relatively stable concentrated values: $t(15) = 2.22$; $p < 0.05$.

The IQR produced the difference between native English and E-Cz of 1.5 ST and this was found to be highly significant: $t(15) = 8.93$; $p < 0.001$. Cz-E interquartile range was also narrower than that of native Czech by about 0.5 ST. Even this result reached statistical significance due to concentration of the values: $t(15) = 2.88$; $p < 0.05$.

Statistical significance of the standard deviation metric was calculated separately for men and women, i.e., with only 7 degrees of freedom (8 men and 8 women). Unlike Mennen et al. (2007) we used a two-tailed test, which is more rigorous. The lower SD for E-Cz compared with native English reached significance for both men and women: $t(7) = 5.03$; $p < 0.05$ and $t(7) = 3.27$; $p < 0.05$, respectively. As displayed in Figure 2, SD for E-Cz was higher in comparison with native Czech. This difference was found significant for women: Cz $t(7) = 2.48$; $p < 0.05$, but not for men: $t(7) = 0.34$; $p > 0.74$.

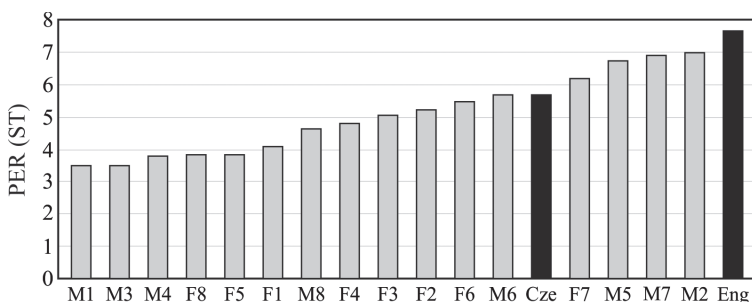
Apart from ranges that express the dispersion of F0 values we were also interested in the position of the lowest and highest values relative to the arithmetic mean. Figure 3 captures the situation. Clearly, there is no difference in the symmetry or asymmetry of the values: under each of the three conditions the speakers depart slightly further down from the mean than up. However, the magnitude of this difference extends to only fractions of a semitone. The limits of PER and IQR offered a very similar picture (on a smaller scale) and SD stretches to an equal distance up and down from the mean by definition.

Figure 3. Mean maxima and minima of F0 values relative to the arithmetic mean normalized to zero. The three columns represent native English, native Czech and English-accented Czech.



Another important question concerns the behaviour of individuals within the studied group. This is because mean group values are more reliable and applicable in predictions if individuals do not depart too far from the mean. Figure 4 presents PER values for each of the speakers in our sample together with the reference values for native English and Czech. It is the decomposition of the middle triplet of values from Figure 1.

Figure 4. Mean 80% percentile range (PER) for individual speakers from the sample (grey columns) together with the reference values for native Czech (Cze) and native English (Eng) represented by the black columns. M = male speaker, F = female speaker; numbering of the speakers bears no relevance to the results.



Obviously, what can be stated about the group does not necessarily hold for all the individuals. Four of the sixteen speakers produced values that fell between the Czech and English reference values. They would comply with the notion of interlanguage. However, majority of the speakers produced values below both referential points.

Figure 4 also reveals that there is no clear division between male and female speakers. The three narrowest ranges were produced by men, but so were the three broadest ranges. Similarly, we did not find any connection between the produced ranges and the length of residence of the speakers in the Czech Republic.

4. Discussion

Pitch range in spoken texts is not random or chaotic. It is used systematically and apart from universal functions (e.g., to signal affective arousal) it also seems to display language specific features. The main obstacle in describing the phenomenon clearly and comprehensively is the lack of methods that would be feasible to employ (given the underfinancing of phonetic research), and satisfactorily sensitive to important melodic events in speech.

Previous research produced interesting, even if sometimes contradictory data from attempts to compare various languages. Mennen, Schaeffler and Docherty (2007) observed a 2.2 ST difference between German and English spoken texts measured by 80% range (referred to as PER in the present study). Their sample was relatively small, but methodologically well managed. Later, in a larger study with different methodology the result was not confirmed (Andreeva et al., 2014). Volín, Poesová & Weingartová (2015) also found a 2 ST difference for 80% range, this time in a large and carefully controlled speech sample comparing Czech and English news reading.

The research in foreign-accented speech is naturally attempting to exploit the fact that languages might differ in their typical overall pitch ranges. However, the influence of the speech style, pragmatic context and affective charge of the communicative situations seems to be stronger than the global inherent tendency in the language. That poses specific demands on the research design. Not only the spoken texts, but also the recording conditions should be made comparable. Even the personality of the experimenter who is collecting the data is perhaps not to be underestimated. The present study tried to embrace these requirements and found that Anglophone speakers who learn Czech as a foreign language tend to use pitch ranges narrower than the ranges used in their mother tongue and often even narrower than those used in their target Czech.

Since the narrower pitch ranges have been reported in foreign-accented speech elsewhere, we might speculate that rather than a mutual interference of two intonational phonologies there is a unique trend in operation: the uncertainty of an L2 learner leads to pitch range compression. In a way, this could be a universal feature of foreignness – at least for situations in which the L2 learners “struggle”. As Major quite convincingly demonstrated, listeners unfamiliar with a language evaluate the degree of accentedness in it similarly to people who know the language (Major, 2007). Maybe pitch range is one of the cues. We should be cautious about the term *uncertainty* used above, though. Most probably it is a complex affective and cognitive structure that will be difficult to define and measure. (This might be further complicated by the fact that the subjects often unconsciously deny the phenomenon’s existence, or the opposite – they consciously claim it while in reality they do not possess it.)

As to further methodological problems, current practice in pitch range quantification seems to be quite crude. The suggestions of Hirst (2003, 2013) are not widely accept-

ed and, furthermore, they do not reflect syntactic/semantic architecture of utterances. A more linguistically motivated approach would require investigation of the local pitch ranges. One such possibility was suggested by Patterson and Ladd (1999), but, to the best of our knowledge, not pursued any further, perhaps because of its labour-extensive nature. Yet it is known that apart from setting the global span, speakers also expand or compress their pitch ranges within utterances depending on the actual communicative situation. This invites quantitative research and promises interesting results if adequate methods are found.

ACKNOWLEDGEMENTS

This study was supported by the Internal Grant Scheme at Metropolitan University Prague under subsection no. VVZ-52-01. The authors would also like to express their sincere thanks to all the volunteers who offered their voices for the purpose of intonation research.

REFERENCES

- Abercrombie, D. (1956). *Problems and Principles. Studies in the Teaching of English as a Second Language*. London: Longmans, Green.
- Anderson-Hsieh, J., Johnson, R. & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning*, 42, 529–555.
- Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., Oleskovicz-Popiel, M. & Trouvain, J. (2014). Comparison of pitch range and pitch variation in Slavic and Germanic Languages. In: *Proceedings of Speech Prosody 2014*, 776–780.
- Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer. Version 5.4.06, retrieved from <http://www.praat.org/>.
- Derwing, T., Munro, M. J. & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, 393–410.
- Derwing, T. M. & Rossiter, M. J. (2003). The Effects of Pronunciation Instruction on the Accuracy, Fluency, and Complexity of L2 Accented Speech. *Applied Language Learning*, 13, 1–17.
- Field, J. (2005). Intelligibility and the listener: the role of lexical stress. *TESOL Quarterly*, 39 (3), 399–423.
- Gilbert, J. (2014). Myth 4: Intonation is hard to teach. In: L. Grant (Ed.), *Pronunciation Myths: Applying Second Language Research to Classroom Teaching*, 107–136. Ann Arbor, MI: University of Michigan Press.
- Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31, 423–445.
- Hirst, D. J. (2003). Pitch parameters for prosodic typology: A preliminary comparison of English and French. In: *Proceedings of 15th ICPHS*, 1277–1280.
- Hirst, D. J. (2013). Melody metrics for prosodic typology: comparing English, French and Chinese. In: *Proceedings of Interspeech 2013*.
- Hirst, D. J. & Di Cristo, A. (1998). *Intonation Systems*. Cambridge: Cambridge University Press.
- Jilka, M. (2000). *Testing the contribution of prosody to the perception of foreign accent*. Dissertationsschrift zur Dr. phil, Fakultät für Philosophie der Universität Stuttgart.
- Jun, S. A. & Oh, M. (2000). Acquisition of second language intonation. In: *Proceedings of ICSLP (Interspeech)*, 73–76.
- Kang, O., Rubin, D. & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94(4), 554–566.

- Keating, P. & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *Journal of Acoustical Society of America*, 132, 1050–1060.
- Ladd, D. R. (2008). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lee, S. (2014). The realization of French rising intonation by speakers of American English. In: *Proceedings of Speech Prosody 2014*, 762–765.
- Major, R. C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29, 539–556.
- Mennen, I., Schaeffler, F. & Docherty, G. (2007). Pitching it differently: a comparison of the pitch ranges of German and English speakers. In: *Proceedings of the XVIth ICPHS*, 1769–1772.
- Patterson, D. (2000). *A Linguistic Approach to Pitch Range Modelling*. unpublished doctoral dissertation, University of Edinburgh.
- Patterson, D. & Ladd, D. R. (1999). Pitch range modeling: Linguistic dimensions of variation. In: *Proceedings of the XIVth ICPHS*, 1169–1172.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227–256.
- Van Engen, K. J. & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers In Human Neuroscience*, 8, 1–4.
- Volín, J. & Bartůňková, H. (2015). Assets and liabilities of simple descriptors of fundamental frequency tracks. In: Niebuhr, O. and Skarnitzl, R. (Eds.), *Tackling the Complexity in Speech*, 147–161. Praha: Faculty of Arts Press.
- Volín, J., Poesová, K. & Weingartová, L. (2015). Speech melody properties in English, Czech and Czech English: Reference and interference. *Research in Language*, 13(1), 107–123.

RESUMÉ

Studie se zabývá intonačním rozpětím v mluvě anglofonních cizinců, kteří se učí česky. Osm mužů a osm žen, jejichž mateřským jazykem je angličtina, ale kteří jsou schopni používat češtinu alespoň na úrovni B1 ESR, bylo požádáno, aby se seznámili s textem rozhlasového zpravodajství a přečetli jej na mikrofon v nahrávacím studiu Fonetického ústavu FF UK. V nahrávkách byly identifikovány nádechové úseky a čtyři typy korelátů intonačního rozpětí byly změřeny v křivkách průběhů základní hlasové frekvence F0. Tyto hodnoty byly porovnány s referenčními hodnotami pro češtinu a angličtinu známými z literatury. Ukázalo se, že Angličané a Američané ve své češtině používají užší intonační rozpětí než v angličtině a u některých ukazatelů dokonce i užší než bylo naměřeno pro češtinu. Ve shodě s výsledky podobných studií zahraničních je možno se domnívat, že návyk na určité intonační rozpětí se nepřenáší do osvojovaného cizího jazyka. Intonační rozpětí spíše odráží určitou živost sdělení nebo jistotu, s jakou mluvčí daný jazyk ovládá. I když se tedy české mluvené texty vyznačují užším intonačním rozpětím než odpovídající texty anglické, hodnoty pro češtinu s anglickým přízvukem se nenacházejí mezi hodnotami pro oba jazyky.