THE SIZE OF PROSODIC PHRASES IN NATIVE AND FOREIGN-ACCENTED READ-OUT MONOLOGUES

JAN VOLÍN

ABSTRACT

The objective of this study is to provide quantitative data concerning size of prosodic phrases in foreign-accented Czech. The speech production of Anglophone users of the Czech language is contrasted with that of Czech professional and non-professional speakers. Each of the three groups of speakers of Czech is represented by 12 speakers. The fourth group of speakers (also 12 subjects) are English professional news readers. They provide data pertaining to the mother tongue of the target group. As expected, the prosodic phrases produced by non-native speakers are shorter and our data provide basis for their modelling that can be used in perceptual testing. One of the interesting outcomes of the study is the revelation that although Czech professional speakers make longer phrases than English professionals if counted in syllables (10.78 against 7.76 syllable per phrase), if counted in words, the difference disappears (4.56 against 4.54 words per phrase). This suggests that semantic constraints on prosodic phrase length are stronger than purely structural ones.

Key words: prosodic phrase, prosodic boundary, foreign accent, clear speech, Czech, English

1. Introduction

Discussions of prosodic phrasing customarily start with the cases of contrastive representational meaning. Linguists are primarily concerned with, and laymen understandably interested in pairs of identical sequences of words which can be uttered so that they mean different things. If the sequence *Definitely not Archie* materializes as one phrase, then it may sound as a strong objection against *Archie*. If, however, there is a clear prosodic boundary after *not*, then Archie is offered as an alternative to something that was decidedly rejected: *Definitely not! Archie*. Moreover, we could speculate about shallow prosodic boundary in case of addressing Archie with a vocative tag: *Definitely not, Archie*. (This third case would presume that vocatives are typically separated from the message itself, which is a rather risky premise outside the declamational style.) Similarly, in a sentence like *I thought that you invited Kate and Amy Martin* it is unclear, whether *Amy Martin* is someone who was supposed to be invited or whether *Amy* was supposed to invite *Martin*.

In the recent decades, many studies have been devoted to investigating the prosodic cues that allow for disambiguation of analogous structures (see, e.g., Lehiste, 1973; Nespor & Vogel, 1983; Price, Ostendorf, Shattuck-Huffnagel & Fong, 1991; Pynte & Prieur, 1996; Carlson, Clifton & Frazier, 2001, etc.). It might be argued that in natural (meaning non-laboratory) settings, the disambiguation is simply provided by the context. In the example above, the interacting individuals should know, whether Amy's surname is Martin or whether she is acquainted with someone called Martin. Nevertheless, Schafer and her colleagues observed in their semi-spontaneous material that speakers signal syntactic differences with prosody even when the context fully disambiguates the structure (Schafer, Speer, Warren & White, 2000). This might suggest that competent speakers of a language use prosodic phrasing habitually to prevent misunderstanding or to serve a purpose other than disambiguation.

The question is whether the cases of potential ambiguity pose a real threat to speech communication. One might wonder how often during an ordinary working day such an ambiguous sentence is produced. We can quite probably testify that within our past few weeks' experience we have uttered thousands of propositions, yet we do not remember one that would expand the list of the above. Unclear meaning is more probably caused by incomplete information or differences in the context evoked in the minds of the provider and the receiver of the message. Does this render intonational phrasing irrelevant? Not in the least. Any time we talk to someone, this someone has to recover the meaning we are trying to convey and there is always some processing cost involved on the part of the listener. Proper phrasing can decrease the cost, the lack of thereof otherwise.

A number of experiments in the 1960s and 1970s demonstrated that speech is processed faster and its content remembered better if it is presented with clear phrasal prosody (e.g., Leonard, 1974; Martin, 1968; O'Connell, Turner & Onuska, 1968; Zurif & Mendelsohn, 1972). Many of the studies were probably inspired by the article by Epstein (1961) who showed that groups of non-words were more successfully recalled by respondents if they were presented with sentence morphology, i.e., if the non-words simulated conventional grammar. Yet, as it became clear soon afterwards, the effect of morphology in spoken stimuli only held if the non-words were presented with phrasal prosody. If presented with list prosody (i.e., as isolated items), the effect disappeared. Similarly, a list of isolated numerals is more difficult to remember than the same numerals prosodically grouped (e.g., Reeves, Schmauder & Morris, 2000).

In contrast, disrupted prosodic structure has been demonstrated to lead to longer reaction times in word, syllable or phoneme monitoring experiments (Meltzer, Martin, Mills, Imhoff & Zohar, 1976; Martin, 1979; Buxton, 1983; Tyler & Warren, 1987) or to compromise listeners' ability to retrieve the intended interpretation of an ambiguous utterance (Ferreira, Anes & Horine, 1996). Similarly, faulty turn construction causes awkward exchange of conversational floor, while proper boundary cues lead to successful transition of turns in conversations (Auer, 1996).

Another important question in the area of our present research involvement is that of syntax – prosody relationship. Due to the relatively long tradition of linguistic description, references to syntax are fairly easy to make and usually plausible enough to accept. We should always remember, however, that educational focus can cause bias. It is not necessarily true that what is taught at schools is somehow more real than what schools currently

ignore or neglect. Just because syntactic rules and units are part of elementary school syllabi, while prosodic structure is not, does not mean that prosody of real speech is in some sort of inferior position to syntax. The traditional belief that prosodic boundaries reflect syntactic structure (e.g., Selkirk, 1984; Price et al., 1991, but also, though not explicitly Kentner & Féry, 2013) is quite difficult to uphold outside the domain of laboratory speech.

Auer argued more than two decades ago that syntax and prosody do not serve one another. Rather, they complement each other to serve the communicative meaning and to manage the recipient's behaviour (Auer, 1996). Although this enlightened proposition is not as yet specific enough for precise phrase boundary predictions, there have been many attempts since to build boundary placement models for various speech materials (e.g., Cooper & Paccia Cooper, 1980; Gee & Grosjean, 1983; Taylor & Black, 1998; Parlikar & Black, 2011). Breen and colleagues discuss two fundamental options in this area of research: meaning-based approach and balance-based approach (Breen, Watson & Gibson, 2011). In their own experiments they also managed to obtain some practicable solutions, but they admit that precise modelling still requires more research. Our current study should contribute to that.

Foreign-accented speech adds one important aspect to the exploration of prosodic boundaries and that is the cognitive load. By this we mean lower-level (i.e., not intellectual) processing demands on the neurophysiology of the speaker's brain. A learner of a foreign language is constrained in the efforts to create proper prosodic phrasing, arguably by substantial detrimental processing factors (e.g., tedious search for words, uncertainty about morphosyntactic rules, or neurophysiological planning of articulatory gestures in phonotactically unfamiliar sequences). In foreign-accented speech, the prosodic boundaries can be involuntary or unplanned – the speaker just has to break the speech continuum when he struggles with the actual cognitive constraints. The results of this labour have to be mapped too for at least two reasons. First, the prosodic phrasing in foreign-accented speech must be eventually tested perceptually in rigorously planned experiments if we want to identify individual factors that impact on the listener. Second, various attempts to understand speech mechanisms have led to the appreciation of the fact that the devil is in detail. This study is motivated by ambition to provide clear, contextually grounded detail that will find its use in future research.

2. Method

The sample of professional native speakers was represented by news readers from respectable national radio stations. It was the BBC for English and Czech Radio (Český rozhlas) for the Czech language. Twelve established news readers (6 men + 6 women) were recorded for each language directly from a broadcast of news bulletins. (The professional experience of individual speakers was ascertained on the web pages of the respective radio stations.)

News reading exemplifies the so-called 'clear speech' – the speaking style used outside ordinary conversational settings, usually under special acoustic or social conditions. The use of clear speech in news reading is understandably essential due to the lack of visual cues for the listeners, the limited amount of shared context between the speaker and the

listeners, and due to relative semantic and syntactic complexity of the texts. Our tentative presupposition is that clear speech manifests prosodic structures more explicitly than common conversational speech thanks to greater production efforts exerted during its production (see also Dellwo et al., this volume).

The news bulletins were quite similar in form for both languages. They comprised 7 to 8 paragraphs (news items) with initial, final and occasionally medial greetings or contact phrases. The mean number of words in the English bulletins was 505, in the Czech bulletins it was 517.

The sample of professional news readers was complemented with twelve Czech speaking non-professionals: university students of 19–23 years of age (8 women + 4 men) who were asked to read out the text of one of the news bulletins in a recording studio. The students were given sufficient time to familiarize themselves with the text. They were well acquainted with both the recording studio and the experimenter who was present. Hence, we expect little impact of nervousness or performance anxiety. These non-professional readers were also advised to make longer pauses between the consecutive paragraphs to avoid performance stress.

Finally, the foreign accent bearers were Anglophone speakers (6 women + 6 men) living and working or studying in the Czech Republic for at least a year with proficiency in the Czech Language of at least B2 of CEFRL (Common European Framework of Reference for Languages). The length of residence was established in an interview after the recording, but since it did not correlate even remotely with the language proficiency, we do not report it.

In the graphs and tables below, the Czech and English professional speakers will be referred to as CzP and EnP respectively, the Czech non-professionals will be CzN, and the speakers of English-accented Czech will be represented by ECN.

Individual recordings were processed in Praat software package (Boersma & Weenink, 2019). The text was aligned with the sounds, and position of individual phones and words was estimated with Prague Labeller (Pollák, Volín & Skarnitzl, 2007) followed by manual corrections of boundaries. Syllabic peaks were established in a special tier with a Praat script.

Prosodic boundaries (or breaks) were located through auditory inspection. Two levels of division were sought, both compatible with ToBI break indices conventions (Price et al., 1991; Beckman & Ayers Elam, 1997) and with other similar recommendations (e.g., Xu, 2011). In this study the break index 4 (BI4) will be referred to as *major phrase boundary*. (Major prosodic phrase is called intonation phrase in some texts.) Such prosodic boundary is indisputable as it is signalled by multiple cues, especially by a very clear F0 pattern, decrease in tempo (i.e., lengthening of the phrase-final syllable or two), occasionally accompanied with a declination reset, change in phonation and amplitude, or specific juncture phenomena and pauses.

Minor phrase boundary (minor prosodic phrase is called intermediate phrase in some texts) is equivalent to what the ToBI transcription system labels as break index 3 (BI3). Such boundaries lack either the phrase-final lengthening or clear F0 pattern, or they may display weakened version of both. The BI3s also leave quite unambiguous feeling of discontinuity, but they require immediate restoration of the flow of speech. Thus, for instance, it would be unnatural to place a silent pause after them.

There were four groups of speakers (altogether 26 female and 22 male subjects) and the following four research questions were asked.

Research Question 1 – What is the mean length of a prosodic phrase in syllables
a) in Czech professional presentation of spoken texts?
b) in English professional presentation of spoken texts?
c) in Czech non-professional renderings of spoken texts?
d) in English-accented renderings of Czech spoken texts?

Research Question 2 – What is the mean length of a prosodic phrase in words?
with a), b), c) and d) subspecifications as above

Research Question 3 – What is the proportional representation of major and minor prosodic breaks in the spoken texts?
with a), b), c) and d) subspecifications as above

Research Question 4 – Is there any correlation between the articulation rate and prosodic phrase length?

To extract information about numbers of syllables, words, prosodic phrases and articulation rates, the scripting facility of the Praat software was used. Where appropriate, testing of statistical significance of differences was related to conventional $\alpha = 0.05$.

3. Results

Figure 1 presents the graphic answer to the research question 1a. Czech professional speakers produced phrases of 10.78 syllables on average. The longest phrases were made by the female speaker CzP02 – 13.2 syllables per phrase. Incidentally, the shortest phrases were also produced by a female speaker. CzP03 only used 8.7 syllables per phrase on average. The outcome then suggests that the male speakers form a more homogenous group, but since the size of the subgroup is only 6 individuals, this fact should not be overemphasized.

Figure 2 provides a set of results that is analogical to the previous one, but describes the phrase production in the groups of English professional news readers. Thanks to the
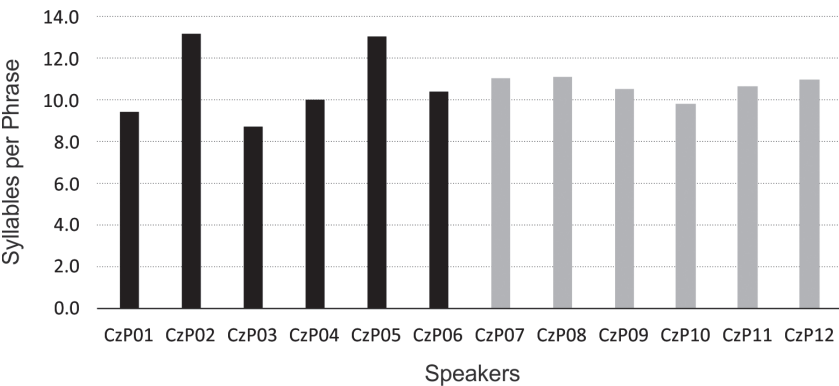

Figure 1. Mean lengths of major prosodic phrases in syllables by individual Czech professional speakers. Black columns = women, grey columns = men.
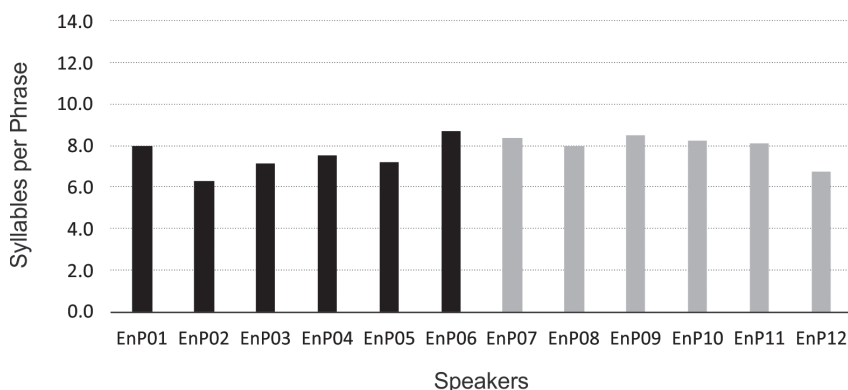
Figure 2. Mean lengths of major prosodic phrases in syllables by individual English professional speakers. Black columns = women, grey columns = men.
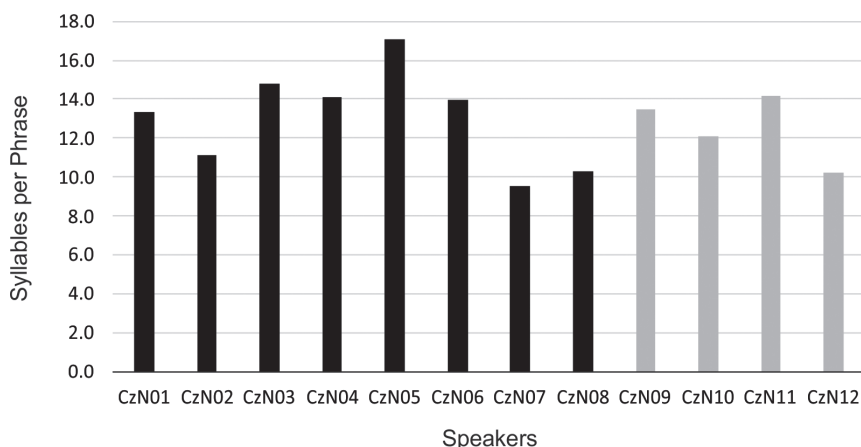

Figure 3. Mean lengths of major prosodic phrases in syllables by individual Czech non-professional speakers. Black columns = women, grey columns = men.

identical scaling of both graphs it is immediately noticeable that the English professionals produced shorter phrases – the mean length across the sample was only 7.76 per phrase, i.e., by three syllables fewer than in the Czech news reading. Interestingly, the highest and the lowest values were again produced by women: EnP06 produced phrases of 8.8 syllables on average, while EnP02 used merely 6.3 syllables. Similarly to the situation in the Czech professional sample, the values provided by men are again more balanced (with the same caveat).

Quite surprisingly, the axis scaling for the phrase production of the Czech non-professional speakers had to be changed (Figure 3). While the Czech professional news readers made prosodic phrases of 10.78 syllables (see above, Fig.1), the non-professional speakers reached the mean length of 12.89 syllables. Nine of the twelve non-professional speakers produced values above the mean of the Czech professionals. Figure 3 also exposes the lon-
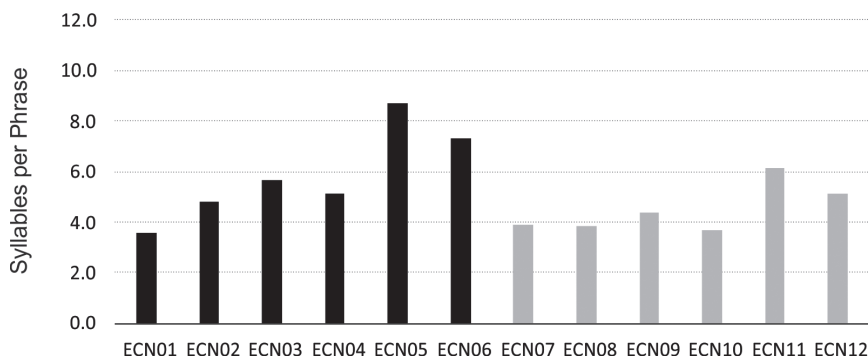
Figure 4. Mean lengths of major prosodic phrases in syllables by individual English speakers of Czech. Black columns = women, grey columns = men.

gest phrases in the sample: the speaker CzN05 created phrases with mean length of 17.1 syllables. That is almost 4 syllables more than the maximum in the Czech professional group. The shortest phrases were also produced by a woman: CzN07 delivered mean length of 9.5 syllables. (It has to be noted, though, that this sample is unbalanced gender-wise.)

The graph in Figure 4 had to be rescaled as well, but this time quite predictably. Based on everyday experience, foreign-accented speech can be anticipated to be more fragmented. This was, indeed, the case: the mean length of a prosodic phrase was only 5.22 syllables. Four of the twelve speakers even produced values under 4 syllables per phrase, three, on the other hand, exceeded 6 syllables per phrase. Yet again, it seems that the male speakers form a more homogenous group (and yet again we warn against over-generalizations from small samples).

Table 1. Mean length and variation of prosodic phrases in Czech professional news reading (CzP), English professional news reading (EnP), Czech non-professional news reading (CzN), and Czech spoken by Anglophone foreigners (ECN). Values of the arithmetic means and standard deviations are in syllables per phrase, coefficients of variation are percentages.

	CzP	*EnP*	*CzN*	*ECN*
mean	10.78	7.76	12.89	5.22
std. dev.	1.31	0.75	2.23	1.57
C_{var} (%)	12.15	9.66	17.30	30.07

Table 1 summarizes the results for the Research Question 1. It shows that in terms of syllables per phrase the longest units were produced by Czech non-professionals. These were followed first by Czech, and then by English professional speakers. As expected, the foreign accented Czech consisted of the shortest phrases. Although this study is designed to provide descriptive data and not to test hypotheses, one-way ANOVA was computed to ascertain the significance of the differences. The outcome was highly significant: $F(3, 44) = 55.83$; $p < 0.001$. (The general α was set to 0.05 – see Method.) Post-hoc Tukey HSD test returned high significance of all the differences between the four conditions.
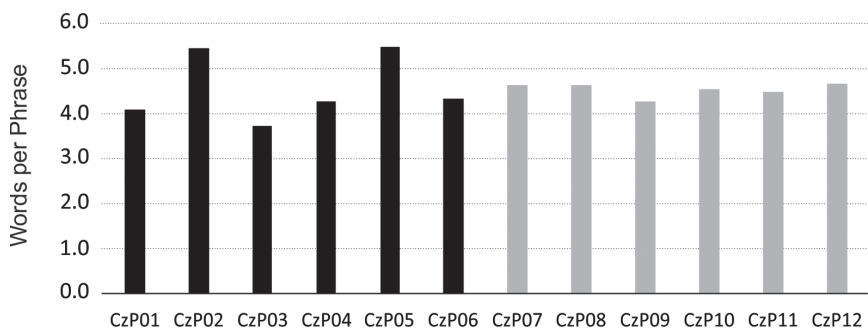
Figure 5. Mean lengths of major prosodic phrases in words by individual Czech professional speakers. Black columns = women, grey columns = men.
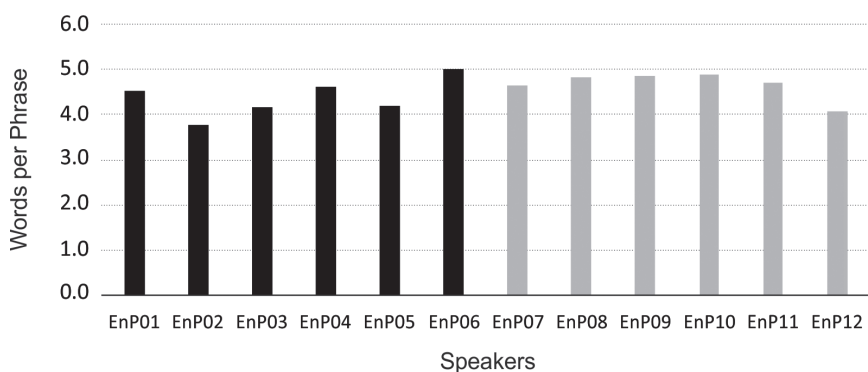

Figure 6. Mean lengths of major prosodic phrases in words by individual English professional speakers. Black columns = women, grey columns = men.

As to the variance in the data, Czech non-professionals produced the largest standard deviation, but after normalization by mean (i.e., computation of the coefficient of variation) the foreign-accented speech turned out to be most variable.

Figures 5 and 6 present mean lengths of prosodic phrases measured in words as produced by Czech and English professional news readers respectively. They are pertinent to Research Question 2 above (see final part of Method). Interestingly, despite the substantial difference between values expressed in syllables per phrase (Fig. 1 and 2), there is practically no difference in lengths expressed in words per phrase. The Czech grand mean is 4.56 and the English one is 4.54 words per phrase. The original difference of 3 syllables per phrase translates into negligible 0.02 words per phrase. This implies that semantic constraints are very similar for both languages, provided the word is the natural semantic building block. Structural constraints obviously differ. The explanation that offers itself first rests in the fact that Czech words are longer due to the rich inflectional system. In other words, there are much fewer monosyllables in Czech texts. There might also be the syllable phonotactics involved: Czech syllables avoid codas to much greater extent than the English ones. (In terms of syllable onsets, the complexity is comparable.)
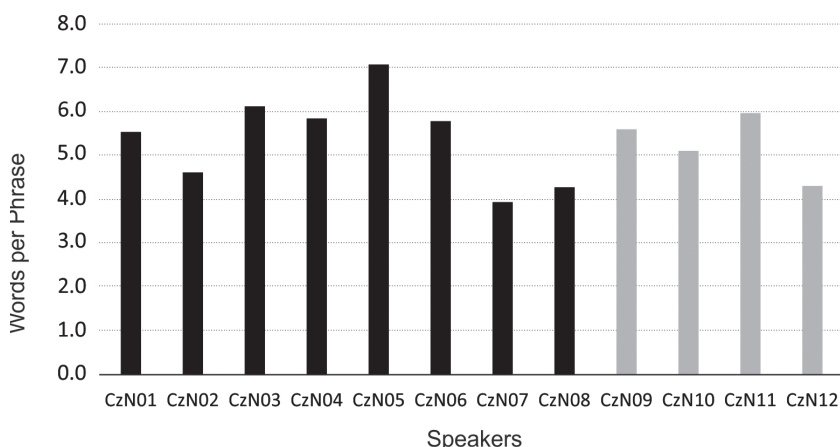
Figure 7. Mean lengths of major prosodic phrases in words by individual Czech non-professional speakers. Black columns = women, grey columns = men.

Figure 7 displays the mean values for the non-professional Czech speakers. The grand mean for this group exceeds both professional groups: it is 5.35 words per phrase. The non-professionals make their phrases by almost one word longer than Czech and English skilled news readers. There is, however, substantial variation within the non-professional group: the female speaker CzN05 makes phrases almost twice as long as the speaker CzN07.

Finally, but crucially for the primary motivation of this study, we have measured the lengths of prosodic phrases produced by Anglophone speakers of Czech. The results are displayed in Figure 8. The grand mean across the whole group is 2.19 words per phrase. That is less than half of the mean length produced by both Czech and English professional speakers (see Fig. 5 and 6). If we disregard speakers ECN05 and ECN06, the grand mean drops to 1.95 words per phrase. This signals quite a substantial number of phrases consisting of one word only, which contributes to the disfluent character of the foreigners' speech production. Indeed, out of 3132 prosodic phrases produced by ECN speakers there were 1722 (55%) containing just one word, of which 176 (about 10%) were monosyllables.

Table 2 summarizes the results for the Research Question 2. One-way ANOVA returned a highly significant effect: $F(3, 44) = 53.15$; $p < 0.001$, and post-hoc Tukey HSD test found no difference between English and Czech professionals, but both these groups were significantly different from Czech non-professionals and foreigners speaking Czech. Variation in the data is analogical to that of lengths in syllables per phrase (see C_{var} in Table 1 above).

Table 2. Mean length and variation of prosodic phrases in Czech and English professional samples (CzP and EnP respectively), Czech non-professional group (CzN), and Czech spoken by Anglophone foreigners (ECN). Values of the arithmetic means and standard deviations are in words per phrase, coefficients of variation are percentages.

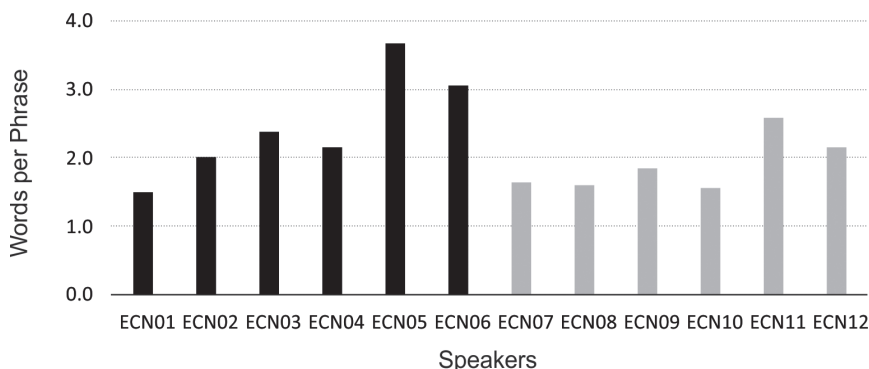|	*CzP*	*EnP*	*CzN*	*ECN*
mean	4.56	4.54	5.35	2.19
std. dev.	0.51	0.39	0.92	0.66
C_{var} *(%)*	11.13	8.53	17.17	30.19

Figure 8. Mean lengths of major prosodic phrases in words by individual English speakers of Czech. Black columns = women, grey columns = men.

The Research Question 3 asked about the relative proportions in the occurrence of minor and major prosodic breaks. Table 3 provides the answer to that. As explained above, the metric chosen for the ratio is the percentage of major prosodic boundaries from the entire set of boundaries. (Since only break indices BI3 and BI4 were included, 75% of major breaks, for instance, would mean 25% of minor breaks). Table 3 indicates that the relative incidence of major breaks was very similar across the four groups of speakers, specifically about 78%, which means that about 22% of the boundaries found were minor phrase boundaries. One-way ANOVA found the actual differences clearly insignificant ($p \approx 0.608$).

Table 3. Mean occurrences of major prosodic boundaries expressed as a percentage of the whole set of boundaries (i.e., major and minor – see Method).

	CzP	*EnP*	*CzN*	*ECN*
major phrase (%)	76.6	80.1	78.4	78.0
std. dev. (%)	6.5	7.3	6.1	5.4
C_{var} *(%)*	8.5	9.1	7.7	6.9

The fourth and final Research Question asked about a relationship between the mean length of the phrase and articulation rate. The Pearson correlation was found as very high: $r = 0.89$ and highly statistically significant ($p < 0.001$). Figure 9 depicts the situation with 48 data points, i.e., each speaker is represented by one data point.

The high correlation coefficient is clearly influenced by non-homogeneity of the whole assembly, especially by the behaviour of the group of foreigners speaking Czech (in the lower left part of the graph). After their exclusion, the correlation drops to $r = 0.51$, but stays highly significant nonetheless ($p < 0.001$). Faster talkers then produce longer prosodic phrases. This comes as no surprise, but it should be remembered that the primary objective of this study was not to discover new trends but, instead, to provide reliable quantitative data about Czech, English and English-accented Czech.
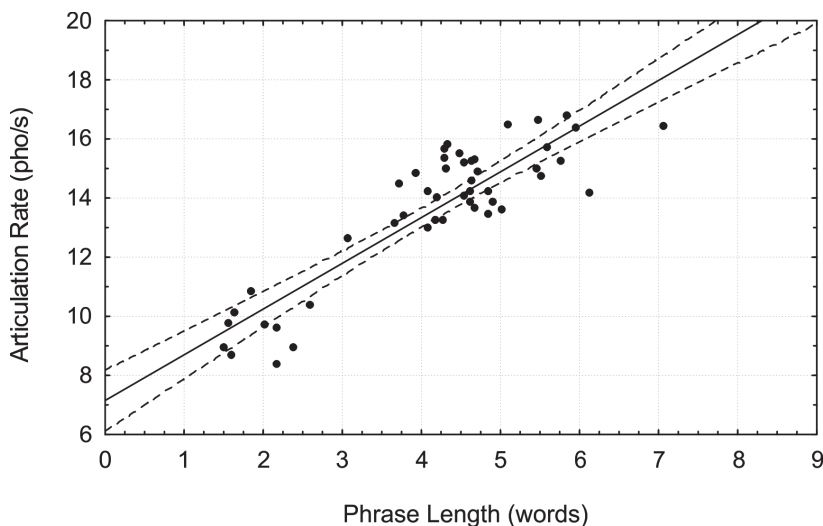
154

Figure 9. Scatterplot of 48 data points (1 for each speaker) capturing the relationship between articulation rate (in phones per second) and mean length of a prosodic phrase (in words).

4. Discussion

As expected, foreign-accented speech was found considerably more broken than native speech. L1 professionals, who are often considered 'model speakers', made about 22 major prosodic boundaries in every 100 words in our sample, whereas L2 speakers made more than 46 of those. Jun's observation that all languages use prosodic grouping even if different languages use them in very different ways (Jun, 2005), can be expanded then: even various groups of the same language users may build phrases in different ways.

Our material showed that a major prosodic phrase in foreign-accented speech very often consisted of one word only and cases when this was a monosyllabic grammatical word were not exceptional. The resulting impression of such fragmentations is typically that of struggle. Our results provide some practical guidance for future perception experiments, which should address the impact of abundant phrase boundaries on the listener. Ultimately, the impact of one's speech is what matters in everyday interactions in multilingual environment (*cf.* Lev-Ari & Keysar, 2010).

Somewhat unexpectedly, our sample of non-professionals produced significantly longer phrases than skilled speakers. Even though this trend is in the opposite direction than that of foreign-accented speech, it is quite probably not advantageous either. Larger continua of speech can pose extra demand on cerebral processing and listeners may find them as tiresome as the texts that do not 'hold together'. However, this can only be hypothesized if we accept the premise that professionals master the language better on all levels. The hypothesis still deserves empirical testing in the future perception experiments that should focus both on comprehension and on memory retention under different phrasing conditions.

One of the interesting details revealed in this study is that although foreigners make many more boundaries in spoken texts, the proportion of major and minor prosodic breaks is virtually the same as in other groups of speakers (see Table 3). This finding should be tested in other speech styles and genres. A speculative interpretation of this fact could state that the minor prosodic boundary is just an auxiliary agency while the major break is the principal way of prosodic grouping. Minor phrases are then used when an occurring semantic unit is too large and needs some sort of internal structure together with the necessity of preserving unity. Another possible explanation could be based on the ambiguity of the minor prosodic break. Analogically to the metrical structure of English, where despite the existence of secondary stresses and full unstressed syllables speakers evidently prefer primary stresses and reduced unstressed syllables in continuous speech, there might be a preference for a clear boundary or no boundary at all in prosodic 'chunking'. Be that as it may, the speech style of read-out news led to more than three quarters of all prosodic phrase boundaries being major.

The study also highlighted a thought-provoking relationship between two ways of measuring the size of prosodic phrases. Even though the results for two typologically disparate languages, Czech and English, differed in numbers of syllables per phrase, they were virtually identical in numbers of words per phrase. If we consider the syllable a basic structural unit, and the word a primary semantic unit, then our finding contributes to the debates on the relationship or interplay between the form and the meaning. It seems that our cognitive mechanisms are less constrained in terms of formal structures but more fastidiously tuned to certain 'amount of the meaning' (*cf.* Caplan & Waters, 1999 or, for instance, Hirotani, Frazier & Rayner, 2006). Naturally, our simple study does not allow for any far-reaching conclusions in this area, but rather invites experimenters with other language backgrounds to take read-out monologues (ideally news bulletins for direct comparison) and to try to replicate our measurements.

Future research should also investigate the acoustic and syntactic nature of the phrase boundaries. Even though informal observations suggest that the phonetic means of prosodic boundary markings are analogous in Czech, English and English-accented Czech, a detailed acoustic analysis might uncover interesting differences. The syntactic disparities, on the other hand, are quite obvious: foreign-accented Czech exhibits, for instance, some unusual breaks between the adjective and the modified noun, between the preposition and the following noun, or even between the first name and the surname of a person. The frequency of occurrence and other circumstances of such and similar cases should be known before further perceptual testing.

REFERENCES

Auer, P. (1996). On the prosody and syntax in turn-continuations. In: E. Couper-Kuhlen & M. Selting (Eds.) *Prosody in Conversation*, pp. 57–100. Cambridge: Cambridge University Press.

Beckman, M. E., & Ayers Elam, G. (1997). *Guidelines for ToBI Labelling, version 3.* The Ohio State University Research Foundation, Ohio State University.

Boersma, P. & Weenink, D. (2019). *Praat: doing phonetics by computer* [Computer program], Version 6.0.47, retrieved 8 February 2019, http://www.praat.org.

Breen, M., Watson, D. G. & Gibson, E. (2011). Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes* 26(10), pp. 1532–1562.

Buxton, H. (1983). Temporal predictability in the perception of English speech. In: A. Cutler & D. R. Ladd (Eds.) *Prosody: Models and measurements* (pp. 111–121). Heidelberg: Springer-Verlag.

Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22(1), 77–94.

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45(1), 58–81.

Cooper, W. E. & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

Dellwo, V., Pellegrino, E., He, L. & Kathiresan, T. (2019). The dynamics of indexical information in speech: Can recognizability be controlled by the speaker? *Acta Universitatis Carolinae – Philologica XX, Phonetica Pragensia XV*, pp. 57–75.

Ferreira. F, Anes, M. D. & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research* 25(2), 273–290.

Gee, J. P. & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15(4), 411–458.

Hirotani, M., Frazier, L., Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language* 54(3), pp. 425–443.

Jun, S. A. (2005). Prosodic typology. In: S. A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, pp. 430–458, Oxford University Press.

Kentner, G. & Féry, C. (2013). A new approach to prosodic grouping. *The Linguistic Review*, 30(2), 1–35.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 107–122.

Leonard, L. B. (1974). The role of intonation in the recall of various linguistic stimuli. *Language and Speech*, 16(4), 327–335.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093–1096.

Martin, J. G. (1968). Temporal word spacing and the perception of ordinary, anomalous, and scrambled strings. *Journal of Verbal Learning and Verbal Behaviour*, 7(1), 154–157.

Martin, J. G. (1979). Rhythmic and segmental perception are not independent. *Journal of the Acoustical Society of America*, 65(5), 1286–1297.

Meltzer, R. H., Martin, J. G., Mills, C. B., Imhoff, D. L. & Zohar, D. (1976). Reaction time to temporally displaced phoneme targets in continuous speech. *Journal of Experimental Psychology: Human Perception and Performance*, 2(2), 277–290.

Nespor, M., & Vogel, I. (1983). Prosodic structure above the word. In: A. Cutler & D. R.

Ladd (Eds.), *Prosody: Models and measurements*, pp. 123–140. Heidelberg: Springer-Verlag.

O'Connell, D. C, Turner, E. A. & Onuska, L. A. (1968). Intonation, grammatical structure, and contextual association in free recall. *Journal of Verbal Learning and Verbal Behaviour*, 7(1), 110–116.

Parlikar, A., & Black, A.W. (2011). A grammar based approach to style specific phrase prediction. In: *Proceedings of INTERSPEECH 2011*, Florence, Italy: ISCA, pp. 2149–2152.

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation* [Ph.D. dissertation]. MIT, Cambridge, MA. [Published in 1987 by Indiana University Linguistics Club, Bloomington].

Pollák, P., Volín, J. & Skarnitzl, R. (2007). HMM-based phonetic segmentation in Praat environment. In: *Proceedings of XIIth Speech and Computer – SPECOM 2007*, pp. 537–541.

Price, P., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustic Society of America*, 90(6), 2956–2970.

Pynte, J., & Prieur, B. (1996). Prosodic breaks and attachment decisions in sentence processing. *Language and Cognitive Processes*, 11(1–2), 165–192.

Reeves, C., Schmauder, A. & Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *Journal of Experimental Psychology – Learning, Memory, and Cognition*, 26(6), pp. 1638–1654.

Selkirk, E. (1984). *Phonology and Syntax*. Cambridge, Mass.: MIT Press.

Schafer, A. J., Speer, S. R., Warren, P. & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29(2), 169–182.

Taylor, P. & Black, A. W. (1998). Assigning phrase breaks from part-of-speech sequences, *Computer Speech and Language* 12(2), 99–117.

Tyler, L. K. & Warren, P. (1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language* 26(6), 638–657.

Xu, Y. (2011). Speech prosody: A methodological review. *Journal of Speech Sciences* 1(1), 85–115.

Zurif, E. B. & Mendelsohn, M. (1972). Hemispheric specialization for the perception of speech sounds: The influence of intonation and structure. *Perception & Psychophysics* 11(5), 329–332.

RESUMÉ

Primárním cílem této studie je poskytnout kvantitativní data týkající se délky promluvového úseku (prozodické fráze) v češtině s cizineckým přízvukem. Řečová produkce anglofonních mluvčích, kteří si osvojují češtinu jako cizí jazyk (L2) je porovnávána s češtinou profesionálních i neprofesionálních mluvčích, ale také s angličtinou jakožto rodným jazykem cílové skupiny. Každý ze čtyř vzorků mluvčích je reprezentován dvanácti mluvčími (tj. celkově *n* = 48). Materiálem jsou čtené monology, konkrétně rozhlasová zpravodajství, která reprezentují mluvu k neznámému publiku a jsou charakteristická požadavkem zřetelnosti.

Podle očekávání byly promluvové úseky v řeči s cizineckým přízvukem kratší a naše data poskytují konkrétní základ pro modelování tohoto rysu v percepčních testech. Tak např. 55 % promluvových úseků vyprodukovaných cizinci bylo jednoslovných a 10 % z těchto jednoslovných bylo dokonce jednoslabičných.

Jedním ze zajímavých výstupů studie je také zjištění, že čeští hlasatelé (profesionální mluvčí) produkují delší promluvové úseky než hlasatelé angličtí co do počtu slabik (v průměru 10.8 slabik na úsek v češtině proti 7.8 slabikám na úsek v angličtině), ale tento rozdíl se vytratí, pokud se délka promluvového úseku vyjádří v počtu slov na úsek (4.56 a 4.54 slova na úsek). Tento výsledek naznačuje, že v daném mluvním stylu je délka promluvového úseku (prozodické fráze) nejspíše vymezena sémanticky. Slovo je totiž jednotkou sémantickou, zatímco slabika jednotkou strukturní. Strukturní rozdíl tří slabik mezi češtinou a angličtinou odpovídá sémantickému rozdílu 0,02 slova.

Studie také přináší doklad o relativně stabilním poměru mělčích a hlubších prozodických předělů u všech čtyř sledovaných skupin mluvčích.

Jan Volín
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
E-mail: jan.volin@ff.cuni.cz