

THE RELATION BETWEEN SUBJECTIVE AND OBJECTIVE ASSESSMENT OF SPEAKING RATE IN CZECH RADIO NEWSREADERS

JITKA VEROŇKOVÁ AND PETRA POUKAROVÁ

ABSTRACT

This article examined objective and subjective speaking rate and their relation. Read speech of 22 Czech radio newsreaders (13 male and 9 female) constituted the material for the study. The objective rate was measured as speech rate and articulation rate, both expressed in syllables per second. Two domains, a single news report and an intonation phrase, were chosen as the units of observation. A perception test was used to establish the subjective assessment of speaking rate. The test was made up of the full news reports and the subjects were asked to rate their tempo on a scale slow – normal – fast. The median of speech rate and articulation rate for the individual news reports was 5.7 syll/s and 6.2 syll/s, respectively. In general, the listeners rated the stimuli tempo as normal and there was no significant difference between the subjective evaluation of female and male speakers. The possible factors influencing the relation between the subjective and objective rating are discussed; no simple and direct relationship between them was found.

Key words: speech rate, articulation rate, intonation phrase, perception, Czech, media speech

1. Introduction

1.1 Speaking rate in Czech media speech

Radio and TV news broadcasts fall among read texts that are very often examined because newsreaders and other media speakers are taken as promoters of standard speech. Regarding speaking rate, there is objective evidence that speech pronounced in the Czech radio and TV has accelerated in the last decades. Mean speech rate of radio newsreaders until 1994 was 5.1 syll/s (males) and 4.8 syll/s (females), in 1996 it was 5.3 syll/s (males and females) (all measured by Bartošek, 2000)¹ and 6.2 syll/s in 2002 (according to Palková, published in Palková et al., 2003). Mean speech rate of Czech TV

¹ The values are rounded to one decimal place.

weather forecast is 5.6 syll/s (Balkó, 1999). Researchers pointed out the fast speaking rate in the mass media and difficulties in speech intelligibility that might be caused just by it (Bartošek, 1995, Palková, 2004, Havlík et al., 2013).²

We decided to inquire into this phenomenon and to inspect the relation between the objective and subjective speaking rate in Czech news broadcasts, because in previous experiments researchers focused mainly on objective measurements without systematic inspection of the subjective perspective of listeners.

In the experiment, both the speech rate and articulation rate³ of Czech radio newsreaders were measured and a perception test was created to receive listeners' assessment. Then the measured (objective) speaking rate was compared with the listeners' ratings, i.e. the perceived (subjective) speaking rate, to establish the degree of correlation.

1.2 Variability in speaking rate

Speaking rate shows both the variability between and within speakers; there are many factors that affect it, both extralinguistic and intralinguistic. However, their influence on the speaking rate may not be direct; the factors may complement one another, but they may also be in contradiction (conf., e.g., Kohler, 1986).

Verhoeven et al. (2004) examining the Dutch corpus of semi-spontaneous speech found that age, sex and the dialect region affect speaking rate. According to their experiment younger people spoke faster than the older, men spoke faster than women and speakers from the West region, considered to be the linguistic centre of the Netherlands (sic!), showed higher speech rate compared to speakers from dialect regions further from the centre. Quené analysing the same corpus claimed that these factors (age, gender and centre/periphery) may play a role, however he considered the length of the intonation phrase to be the most important factor. He quoted findings of Nooteboom and Lindblom – Rapp from the 70s and explained that longer phrases containing more syllables showed a tendency towards a faster speaking rate and a shorter syllable duration.⁴ According to Quené's (2005) experiment, younger people tended to use longer intonation phrases and this was the cause of their faster speaking rate. However, Quené admitted that it was not possible to provide such a direct explanation for the other factors. There were other aspects, arising from the linguistic structure of texts, that caused changes of speaking rate, such as syllable structure (Pfitzinger, 2006), or phrase-final deceleration (for Czech Dankovičová, 2001; Volín & Skarnitzl, 2007), etc.

1.3 The relation between objective and subjective speaking rate

The variability of speech rate is also reflected in perception. Just as in speech production, there are many factors that affect our perception of speech rate. Kohler (1986), investigating German words and sentences, found overall duration, overall F0 level and

² In informal discussions listeners (especially the elderly) also complain that speaking rate in Czech media is very fast.

³ More details about these measurement in 1.4.

⁴ Ondráčková (1954a, b) (among others) refers to this phenomenon on the level of stress unit in Czech.

F0 movements to be important tempo cues. He examined not only the strength of the single factors, but pointed out that there was a relation between them and also between the factors and the production patterns. J. Koreman (2006) confirmed that articulation rate played an important role, yet not an exclusive one, for the perception of speech rate; other factors, e.g. pausing, disfluencies and other prosodic properties of speech, also determined the perceived speech rate. In his experiment the slow and fast speakers differed in their objective speech rates, however the grouping of the speakers according to the perceived speech rate did not match the categorisation based on the measured speech rate.

H. Pfitzinger (1998), who focused on local speech rate, claimed that linear combination of both the syllable and the phone rates (see 1.4 below) corresponded best to the perception of speaking rate.

1.4 Measurement of speaking rate – methodological notes

Speaking rate tells us, in general, the number of speech units produced by a speaker over a time unit. It is often expressed in syllables per second (syllable rate) or in phones per second (phone rate). It could also be expressed as the average syllable/phone duration. The interval unit could also be defined as vocalic and consonantal intervals (VC intervals) (conf. BonnTempo-Corpus (Dellwo et al., 2004)).

The chosen domain used to establish the speech rate did crucially influence the obtained results; it especially affected the degree of the speech rate variability (conf. Miller et al. (1984)). In the present study two domains were applied: a single item of a news bulletin and a unit we call an intonation phrase. The reason to choose an intonation phrase was based on the research of Dankovičová (1997, 2001). She examined the variability of articulation rate in Czech comparing three domains (breath group, syntactic phrase and intonation phrase) and she found that it was only the intonation phrase that had regular patterns of articulation rate. This unit was phonetically defined as a group of stress units combined together into a compact intonation unit. However, to determine this unit in Czech the sound properties of its boundaries were crucial (Palková, 1997, Daneš, 1957). Janoušková (2008) formalized and experimentally verified the hierarchy system of sound units in Czech and showed that a pause, a distinct melodic contour and the presence of final lengthening were important boundary markers of an intonation phrase. See the Appendix for an example of stimuli with intonation phrase boundaries marked.

Pauses influence both the production and the perception of speaking rate. They were also taken into account in the procedure of speech rate measurements. In calculating speech rate, pauses were counted into the duration of speech; however, in articulation rate measurements the duration of pauses was excluded. The principal question was to define the minimal pause duration to determine an articulatory pause. The usual value ranged between 0.2 and 0.3 s, following the works of Goldman-Eisler (1961). Hieke et al. (1983) recommended a value closer to 0.1 s; the duration of 0.13 s was used by Dankovičová (2001).

2. Experiment

2.1 Material

For the purpose of the experiment news broadcasts of the public Czech Radio (CRO) were used. The recordings were obtained from the web archive of CRO⁵ using the Cool-Edit96 software, sampling frequency 32 kHz, 16-bit amplitude resolution⁶; all the recordings were taken in the period between February–April 2008. The news bulletins were divided into single news reports; only the main speakers (i.e., the news readers) were taken into account, the speech of correspondents, analysts etc. was excluded as well as the news with backchannel sounds and disfluencies such as slips of the tongue. To eliminate yet another potential variable, the topic of the news was limited to “politics”; sports news and weather forecasts were also omitted. According to these criteria 22 speakers (13 male and 9 female) and their 22 news reports (1 recording per speaker) formed the corpus for the further procedure.⁷

2.2 Perception test

In order to keep the task as natural as possible, the news were presented to the listeners in continuous form, as retrieved from the source. The testing was carried out one year after retrieval of the material. Overall, the median length of the stimuli was 26.6 s, item length ranging between 24–32 seconds. Speech covered over 90.0% of each tested sample; except for one stimulus where the volume of pauses exceeded 10.0%. The stimuli contained 28 intonation phrases and 160 syllables on average and the number of intonation phrases/syllables was comparable in different stimuli. The intonation phrases that consisted of just one stress unit were the most numerous (46.3%).⁸

Each stimulus corresponded to one news report and it was played just once. Because of the total duration, the test was divided into two parts; each part lasted approx. 7 min., including 3 stimuli for the training session.

The listeners were Czech natives, students of the Czech language at the Faculty of Arts at Charles University in Prague, their mean age was 21 years. Seventeen listeners participated in both parts I and II (with an interval of 1 week in between the two parts); in addition, other 8 listeners participated in part II. All the listeners were females; a sufficient number of males was not available. The perception test was carried out in a sound treated lecture room. For each stimulus, the listeners judged the speaking rate of the speaker.

The listeners were presented with 3 main categories of speaking: slow (-1) – normal (0) – fast (1); they had the possibility to refine the evaluation using arrows to indicate a finer assessment within the basic category, in essence, it was a 9-point scale. For an illustration of the sheet form see Tab. 1 (part A). (Part B shows how the judgements were expressed numerically.)

⁵ <http://hledani.rozhlas.cz/iradio/>, (recorder February–April 2008)

⁶ Raw data: total duration 210 min., 24 newsreaders, total duration of news bulletins 4 min. on average.

⁷ In addition 3 more news reports were chosen for the training session of the perception test.

⁸ This finding is in accordance with Dankovičová (2001) and Hrachová (2016).

Table 1. Illustration of the answer sheet for the perception test (part A), with the corresponding numerical expression of the judgements (part B).

A	Speech Rate								
	Slow			Normal			Fast		
1.		-1			0			←	①
2.		-1			①				1
...									
B	-1.33	-1.00	-0.66	-0.33	0.00	0.33	0.66	1.00	1.33

2.3 Segmentation of the speech material

First, each news report was transcribed orthographically. To prepare the canonical transcriptions, the software *Convertor* (Laun, 2001) and the template *TRIPAC* (Janoušková, 2003) were used and consequently manually corrected. The intonation phrases were labelled using the software *Praat* (Boersma & Weenink, 2002–2010). A pause was classified as an articulation pause if its duration was at least 0.2 s (see 1.4 above). Both the labelling and the listening method yielded the same results regarding the position of pauses.

The segmentation was done according to the rules based on Machač & Skarnitzl (2009). It was necessary to solve special cases where a consonant at the beginning of a word was preceded by a glottal gesture followed by a schwa. Such a preglottalized sequence was regarded as a part of the given word, and was counted as an extra syllable. (Conf. the findings about preglottalization in news reading in Skarnitzl & Machač, 2009.)

2.4 Laboratory measurements

Speaking rate was expressed in syllables per second, both for speech rate (SR) and articulation rate (AR). Except for preglottalization mentioned in 2.3, the canonical number of syllables in words was used to calculate speaking rate; in Czech, syllable compression was not frequent in general and it appeared only once in our material (unlike the elision of consonants).

The entire news report corresponding to one stimulus of the perception test served as the unit of observation; speaking rate was measured in the whole recording: the duration of each item was divided by the number of syllables to receive SR. The item duration reduced by the total duration of internal pauses and divided by the number of syllables served for the calculation of AR.

Because of the possible variation of speaking rate within the stimuli and its potential influence on the subjective assessment, articulation rate was also calculated in yet another way.

Another value of a speaker's articulation rate (ARIP) was obtained by determining the mean articulation rate of all intonation phrases within each recording (where the into-

nation phrase AR was calculated as the duration of a single intonation phrase divided by the number of syllables it contained).

The values of SR, AR and ARIP for each stimulus (N = 22) were obtained.

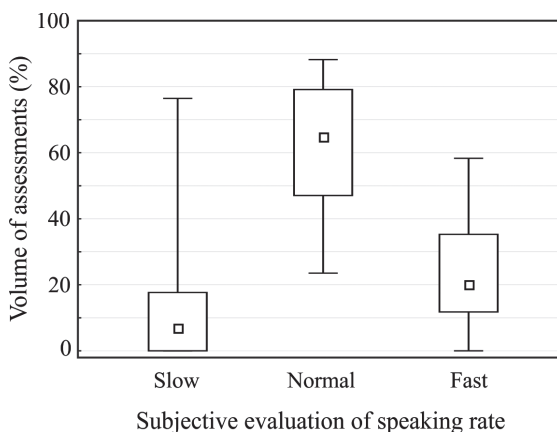
3. Results

3.1 Subjective speech rate

The speaking rate of every stimulus, i.e. each news report, was judged (see 2.3 above). Altogether there were 460 assessments available. The distribution of the global evaluation of speaking rate within the perception test was: slow (14.6%), normal (63.3%) and fast (22.2%). It is obvious that the listeners did not use all the categories equally in their assessments. The category normal speaking rate was chosen in two thirds of the assessments, while in almost one quarter of the cases the listeners assessed the speaking rate as fast.

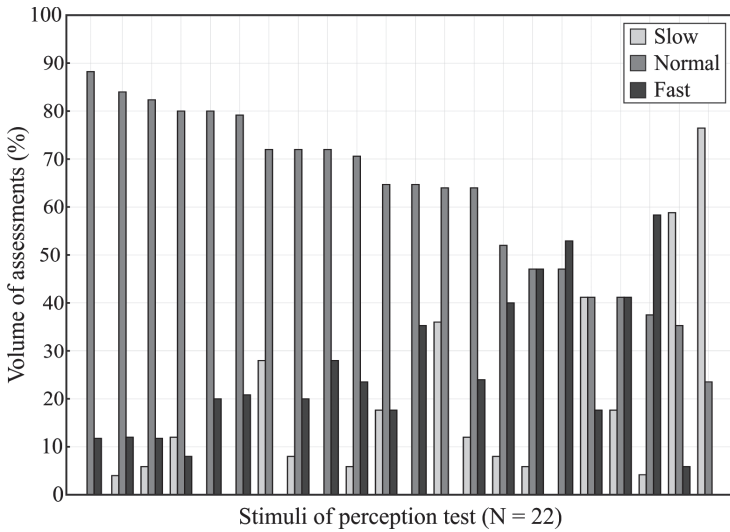
Subsequently the assessments of individual stimuli were examined. See Fig. 1 for the categories slow – normal – fast as a whole and Fig. 2 for individual stimuli.⁹ The category normal speaking rate was not only chosen by the listeners most often but it also showed the highest agreement in the rating of single stimuli in comparison with the other categories. The range of agreement of rating the individual stimuli within the category normal was 23.5–88.2%, while the agreement on speaking rate in category slow was 0.0–76.5% and 0.0–58.3% for the category fast. It can be seen that at least 20% of the listeners agreed that the speaking rate of any given stimuli was normal, whereas the categories slow or fast remained unused in the evaluation of some stimuli.

Figure 1. Subjective evaluation of speaking rate. The volume of slow – normal – fast ratings (in %). Indicated are the median, quartile range, and range.



⁹ All the tests, correlations and graphs were computed and visualized using software Statistica 12.0 (StatSoft).

Figure 2. Subjective evaluation of speaking rate of individual stimuli. The volume of slow – normal – fast ratings (in %).



For the following procedure, a recalculation of the subjective judgements was done (see Tab. 1 Part B) and the average evaluation of each stimulus was calculated. The tendencies shown above were confirmed: the speaking rate of stimuli was evaluated as normal on average.

According to the Mann-Whitney test there was no significant difference between the subjective evaluation of female and male speakers at $\alpha = 0.05$ (the U-value is 30, the critical value of U at $p < 0.05$ is 28).

3.2 Objective speech and articulation rate

Table 2. Objective measurement: speech rate (SR), articulation rate (AR), articulation rate based on articulation rate of intonation phrases (ARIP). In syll/s.

syll/s	SR			AR			ARIP		
	F	M	Total	F	M	Total	F	M	Total
Number of stimuli	9	13	22	9	13	22	9	13	22
Median	5.8	5.7	5.7	6.2	6.2	6.2	6.2	6.0	6.1
1st quartile	5.7	5.6	5.6	6.1	6.0	6.0	6.0	5.9	5.9
3rd quartile	6.0	5.9	6.0	6.4	6.5	6.5	6.3	6.3	6.3
Minimum	5.3	5.3	5.3	5.8	5.8	5.8	5.6	5.6	5.6
Maximum	6.2	6.2	6.2	6.6	6.9	6.9	6.5	6.8	6.8
Mean	5.8	5.8	5.8	6.2	6.3	6.3	6.2	6.1	6.1
Stand. dev.	0.3	0.3	0.3	0.3	0.4	0.3	0.3	0.4	0.3

For each stimulus, we measured speech rate (SR) and articulation rate (AR and ARIP)¹⁰.

Median values of all the 22 stimuli (in syll/s) were: SR 5.7, AR 6.2, ARIP 6.1. The values for SR were lower than for the articulation rates, and there were visible overlaps between AR and ARIP. See Table 2 and the box plot in Fig. 3. The differences, tested by means of a t-test for correlated measures, were significant at alpha = 0.05 not only for SR and articulation rates, but also for AR and ARIP; SR – AR: $t(21) = -17.4$; SR – ARIP: $t(21) = -10.9$; AR – ARIP: $t(21) = 4.5$. According to the Pearson correlation coefficients all the correlations (SR – AR, SR – ARIP, AR – ARIP) were very high ($r = 0.9$ in all three cases). The scatter plots (Fig. 4a–c) show these relations very clearly.

Figure 3. Speech rate (SR), articulation rate AR and ARIP of stimuli (N = 22). Indicated are the median, quartile range, and range.

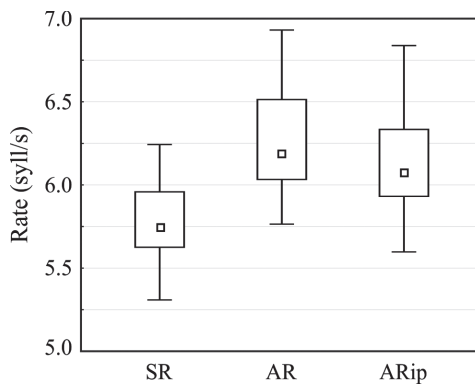
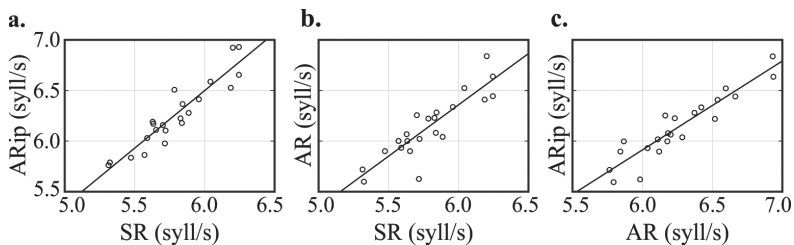


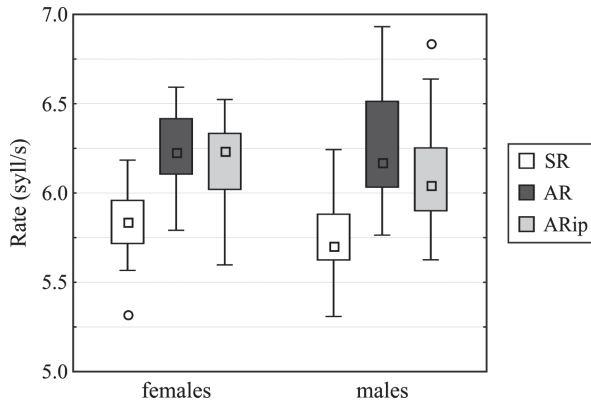
Figure 4. a. Articulation rate AR as a function of speech rate SR (N = 22), $r = 0.93$. **b.** Articulation rate ARIP as a function of articulation rate AR (N = 22), $r = 0.93$. **c.** Articulation rate ARIP as a function of speech rate SR (N = 22), $r = 0.88$.



The values of SR, AR and ARIP in female speakers were more compact than in males, but there were visible overlaps and according to the Mann-Whitney test, the differences between males and females in SR, AR and ARIP were not significant at alpha = 0.05. See Table 2 and box plot in Fig. 5.

¹⁰ AR: articulation rate counted within the whole stimulus. ARIP: articulation rate counted as mean articulation rate of the intonation phrases. See 2.5 above.

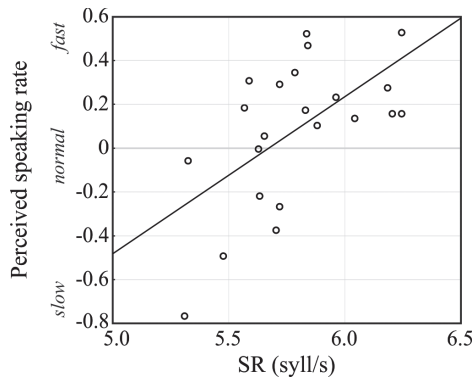
Figure 5. Speech rate (SR), articulation rate AR and ARIP of stimuli. Indicated are the median, quartile range, and range, as well as extreme values.



3.3 The relation between subjective and objective speech and articulation rates

Firstly, the relation between the subjective speaking rate and the measured values for SR, AR and ARIP was examined. Spearman rank correlation coefficients implied low/moderate correlations ($r_s = 0.5$ in all three cases). The scatter plot in Fig. 6 shows the relation between SR and the subjective speaking rate.

Figure 6. Subjective speaking rate as a function of speech rate SR (syll/s) (N = 22), $r = 0.54$



The influence of pauses, as another potential variable, was also examined. The Spearman rank correlation coefficient implied very low or low correlation between the amount of pauses and the examined parameters (SR: $r_s = 0.08$, AR $r_s = 0.36$, ARIP: $r_s = 0.30$, subjective speaking rate: $r_s = 0.05$).

It seems that there was no single correlate that would certainly describe the relation between the objective and the subjective evaluation of the tested stimuli in our corpus. This hypothesis could be illustrated by other specific examples. Speaker A was one of the

objectively fastest speakers in our corpus (SR: 6.2 syll/s, AR 6.9 syll/s). Subjectively his speech rate was evaluated as fast by 52.9% of listeners and as normal by 47.1% of listeners. Speaker B received similar subjective evaluation (fast in 58.3% and normal in 37.5%) but had lower objective values (SR: 5.8 syll/s, AR 6.2 syll/s). On the other hand, subjective speech rate of speaker C was evaluated as normal in 88.2%, although the objective values were higher (SR: 6.0 syll/s, AR 6.6 syll/s).

4. Discussion

The speaking rate of read radio bulletins was examined from both the subjective and objective point of view.

Regarding the objective rate, both the speech rate (SR) and the articulation rate (AR and ARIP) were calculated. All of these parameters were highly correlated with each other and at the same time the differences between them were significant. On the other hand, there was no significant difference found between male and female speakers; this finding was in accordance with the results of Veroňková and Janoušková (partially published in Veroňková, 2012), who examined the speaking rate of TV newsreaders.

The values measured in our experiment (median values in syll/s SR 5.7, AR 6.2, ARIP 6.1) corresponded to SR of Czech newsreaders obtained by other researchers in their studies from the last decade: 6.2 syll/s measured by Palková (in Palková et al., 2003) or 5.9 syll/s (males) and 6.1 syll/s (females) (Veroňková & Janoušková)¹¹. According to Palková (2004), the acceptable SR of news in Czech ranged between 5.5 and 5.8 syll/s. As far as our recordings are concerned, the SR of the tested news reports fell mostly to the higher end of this range, some even exceeded the limit (compare with a similar finding in Shevchenko & Uglova (2006) for TV news in the USA). Speaking rate of news presentation in professional speakers was consistently faster than SR in other genres or in non-professional speakers: the average SR of read speech pronounced by Czech university students was 4.7 syll/s (Balkó, 1999), 4.5 syll/s (Veroňková-Janíková, 2004); SR of guests performing in the radio broadcasts was 4.3 syll/s (Bartošek, 2000); SR of direct sport reports was 5.6 syll/s (ibid).

Speaking rate of the tested speech samples, i.e. news reports, was subjectively judged as normal on average. The same finding was examined by the above-mentioned study (Veroňková, 2012). Two factors could serve as an explanation for the low/moderate correlations between subjective judgement and objective measures of speech rate.¹² Firstly, the age of listeners may play a role. In both cases, the subjects were university students, i.e. younger adults. It is possible that the younger generation was more tolerant to higher speaking rates.¹³ This opinion can be supported by the findings of Verhoeven et al. (2004) and Quené (2005) (mentioned in 1.1 above) regarding the faster SR of younger

¹¹ The data are based on the material of TV news from 2003 published partially in Veroňková (2012).

¹² One of the anonymous reviewers of this paper also pointed out a third factor: "The fact that most people mostly chose "normal" as the response ... [is] affected by the response options they had. Choosing the medium option is a frequent response strategy."

¹³ According to the findings of Veroňková & Janoušková (2010) based on the material of TV news, younger people praise the (relatively) slower SR of newsreaders, however they do not complain about the objectively fast one.

people, together with the results of Schwab (2011) who confirmed the influence of the listeners' own objective SR to their subjective judgement of speaking rate. Secondly, the listeners probably took the type of communication and text into account – the SR within the genre, i.e. the news reporting, was subjectively being rated as normal. An experiment examining the subjective assessment of speech samples from different genres is being prepared.

In our experiment, which was focused on relatively long stimuli, no simple and direct relation between the subjective and objective rating was found. As it was discussed, there are many factors that could influence both the perception and the production of speech rate. It is clear that the longer the speech sample the harder it is to establish and mark off the one and most important factor as all of them are tightly bound together. There are several factors that deserve special attention in the following analyses of the corpora: F0 contour, pauses and the way of segmentation in general, the precision of articulation and phone rate. In the future studies, it would be useful to obtain the subjective speaking rate assessment of single intonation phrases from our corpora and to compare it both with the subjective rating of the given speaker and with the objective measurements.

REFERENCES

- Balkó, I. (1999). K fonetickému výzkumu tempa řeči a tempa artikulace v čteném textu a spontánním projevu. In: Jinakost, cizost v jazyce a literatuře, (pp. 38–44). Ústí nad Labem: UJEP.
- Bartošek, J. (1995). Jazyková kultura mluveného zpravodajství. In: J. Jančáková, M. Komárek & O. Uličný (Eds.), Spisovná čeština a jazyková kultura. Sborník z olomoucké konference 23.– 27. 8. 1993, Vol. 1 (pp. 166–171). Praha: FF UK.
- Bartošek, J. (2000): Mluvní tempo v rozhlasu a v televizi. In: Psychotrofon 1 (pp. 78–82). Olomouc: Univerzita Palackého.
- Boersma, P. & Weenink, D. (2002–2010). Praat: doing phonetics by computer. (Version 5.0.10, retrieved February 27th, 2008 and 5.1.23, retrieved January 1st, 2010), <http://www.praat.org>.
- Daneš, F. (1957). Intonace a věta ve spisovné češtině. Praha: Academia.
- Dankovičová, J. (1997). The domain of articulation rate variation in Czech. *Journal of Phonetics*, 25, 287–312.
- Dankovičová, J. (2001). The Linguistic Basis of Articulation Rate Variation in Czech. In: H. W. Wodarz (Ed.), *Forum Phonetikum 71*. Frankfurt am Main: Hector.
- Dellwo, V., Steiner, I., Aschenberner, B., Dankovičová, J. & Wagner, P. (2004). BonnTempo-Corpus and BonnTempo-Tools: A database for the study of speech rhythm and rate. In: Proceedings of the 8th ICSLP (pp. 777–780).
- Goldman-Eisler, F. (1961). The Significance of Changes in the Rate of Articulation. *Language and Speech*, 4, 171.
- Havlík, M., Jílková, L., Kaderka, P. & Mrázková, K. (2013). Analýza a hodnocení jazykové úrovně vybraných pořadů České televize za 1. pololetí 2013. Retrieved February 2016 <http://www.ceskatelevize.cz>
- Hiege, A. E., Kowal, S. & O'Connell, D. C. (1983). The trouble with „articulatory“ pauses. *Language and Speech*, 26, 493–511.
- Hrachová, J. (2016). Temporální charakteristika promluvočných úseků v řeči profesionálních mluvčích. Unpublished diploma thesis. Prague: Institute of Phonetics, Faculty of Arts, Charles University.
- Janoušková, J. (2003). TRIPAC. Originální aplikace šablony Microsoft Word pro potřeby fonetické transkripce češtiny. Interní nástroj Fonetického ústavu Filozofické fakulty univerzity Karlovy. Praha: Fonetický ústav FF UK v Praze.

- Janoušková, J. (2008). Shoda percepčního hodnocení hloubky prozodických předělů v závislosti na struktuře čteného textu. In: J. Volín & J. Janoušková (Eds.), *AUC Philologica 1, Phonetica Pragensia XI* (pp. 87–104). Praha: Karolinum.
- Kohler, K. J. (1986). Parameters of speech rate perception in German words and sentences: duration, F0 movement, and F0 level. *Language and Speech*, 29, 115–140.
- Koreman, J. (2006). The Role of Articulation Rate in distinguishing fast and slow speakers. Proceedings of the 3rd International Conference on Speech Prosody. Retrieved December 12th, 2009, <http://www.hf.ntnu.no/isk/koreman/Publications/2006/SpeechProsody2006.pdf>
- Laun, M. (2001). Convertor. Version 1.0.4. Software pro fonetickou transkripci češtiny.
- Machač, P. & Skarnitzl, R. (2009). *Fonetická segmentace hlásek*. Praha: Nakladatelství Epoque.
- Miller, J. L., Grosjean, F. & Lomanto, C. (1984). Articulation Rate and Its Variability in Spontaneous Speech: A Reanalysis and Some Implications. *Phonetica*, 41, 215–225.
- Ondráčková, J. (1954a). O mluvním rytmu v češtině. I. Pojetí mluvního taktu. *Slovo a slovesnost*, 15, 24–29.
- Ondráčková, J. (1954b). O mluvním rytmu v češtině. II. Členění mluvy. *Slovo a slovesnost*, 15, 145–157.
- Palková, Z. (1997). *Fonetika a fonologie češtiny*. 2nd edition. Praha: Karolinum.
- Palková, Z. (2004). Srozumitelnost řeči v rozhlasovém a televizním zpravodajství. In: Přednášky z XLVII. běhu Letní školy slovanských studií (pp. 97–109). FF UK: Praha.
- Palková, Z., Veroňková-Janíková J. & Hedbávná, B. (2003). Zvuková podoba rozhlasové češtiny. In: Proměna rozhlasového výrazu a tvaru. Sborník příspěvků z jarního semináře (pp. 20–39). Praha: Sdružení pro rozhlasovou tvorbu.
- Pfitzinger, H. R. (1998). Local Speech Rate as a Combination of Syllable and Phone Rate. In: Proceedings of ICSLP, Vol. 3 (pp.1087–1090). Sydney.
- Pfitzinger, H. R. & Tamashima, M. (2006). Comparing Perceptual Local Speech Rate of German and Japanese Speech. In: Proceedings of the 3rd International Conference on Speech Prosody (ICSP 2006), Vol. 1 (pp.105–108). Dresden.
- Quené, H. (2005). Modelling of Between-Speaker and Within-Speaker Variation in Spontaneous Speech Tempo. In: Proceedings of Interspeech (pp. 2457–2460). Lisbon.
- Schwab, S. (2011). Relationship between speech rate perceived and produced by the listener. *Phonetica*, 68, 243–255.
- Shevchenko, T. & Uglova, N. 2006. Timing in news and weather forecasts: implications for perception. Proceedings of the 3rd International Conference on Speech Prosody (ICSP 2006), Vol. 1 (pp. 5–8). Dresden.
- Skarnitzl, R. & Machač, P. (2009). Domain-initial coordination of phonation and articulation in Czech radio speech. *AUC Philologica 1, Phonetica Pragensia XII*, 21–35.
- Verhoeven, J., De Pauw, G. & Kloots, H. (2004). Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands. *Language and Speech*, 47, 297–308.
- Veroňková, J. (2012). Tempo řeči z různých stran. In: K. Hlínová & R. Vacula (Eds.), Sborník Asociace učitelů češtiny jako cizího jazyka. Praha: Akropolis, 203–223.
- Veroňková-Janíková, J. (2004). Dependence of individual speaking rate on speech task. In: Z. Palková & J. Veroňková (Eds.), *AUC – Philologica 1, Phonetica Pragensia X* (pp. 107–123). Charles University in Prague: The Karolinum Press.
- Veroňková, J. & Janoušková, J. (2010). Jak hodnotí posluchači zvukovou stránku projevů u moderátorů zpravodajství ČT1? In: S. Čmejková, J. Hoffmannová & E. Havlová (Eds.), *Užívání a prožívání jazyka* (pp. 115–119). Praha: Karolinum.
- Volín, J. & Skarnitzl, R. (2007). Temporal downtrends in Czech read speech. In: Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007) (pp. 442–445). Antwerpen: ISCA.
- <http://hledani.rozhlas.cz/iradio/> (retrieved February–April 2008)

APPENDIX

An example of a news report – stimulus of the perception test (orthographic form) showing intonation phrase boundaries (||). (P) refers to the noticeable pause.

Posádka raketoplánu Endeavour má za sebou úspěšný pátý výstup do volného kosmického prostoru. (P) Během šesti hodin připevnila na mezinárodní vesmírnou stanici několikametrový nástavec s kamerou. (P) Ta bude vesmírným lodím sloužit ke kontrole tepelného štítu,

aby se už neopakovala tragédie z roku dva tisíce tři, (P) kdy malá trhlinka způsobila rozpad raketoplánu Columbia (P) a všech sedm členů posádky zahynulo. (P) Astronauti z Endeavouru teď odpočívají před úterním odletem zpět k Zemi. (P)

Translation

The crew of the Endeavour space shuttle has completed the fifth extra-vehicular activity. During the six-hour EVA, they attached a several-meter long extender with a camera. It will serve for space shuttle heat shield tests so that a disaster, similar to the one of space shuttle Columbia from 2003 when a small crack caused its break-up and all seven crew members died, would not happen again.

The astronauts in the Endeavour space shuttle are now resting before they will take off back to the Earth on Tuesday.

RESUMÉ

Vztah mezi subjektivním a objektivním mluvním tempem u moderátorů českého rozhlasového zpravodajství

Experiment zkoumá objektivní a subjektivní mluvní tempo a jejich vzájemný vztah. Materiál tvoří nahrávky zpráv Českého rozhlasu (22 mluvčích – 9 žen a 13 mužů), od každého mluvčího 1 zpráva o délce 24–32 s. Subjektivní hodnocení bylo získáno na základě percepčního testu, položky tvořily jednotlivé zprávy. Posлуhači hodnotili vzorky na škále tempo pomalé – střední – rychlé. Pro každý vzorek bylo změřeno mluvní tempo celkové (MTC) a artikulační tempo (ve slabikách za sekundu). Artikulační tempo bylo počítáno dvojím způsobem; shodně s mluvním tempem celkovým v rámci položky (AT) a dále v rámci jednotlivých promluvových úseků (ATú). Naměřené hodnoty mediánu činí pro MTC 5,7 sl/s, AT 6,2 sl/s a ATú 6,1 sl/s. Mezi MTC, AT a ATú byla prokázána vysoká korelace; zároveň bylo ověřeno, že rozdíly mezi nimi jsou statisticky významné. Naopak mezi muži a ženami není ve zkoumaném vzorku rozdíl v tempu řeči významný. Posлуhači hodnotí položky z hlediska tempa obecně jako střední. Na produkci i percepci tempa působí vzájemně mnoho faktorů. To se potvrdilo i v našem experimentu; v rámci proměnných (MTC, AT, ATú, objem pauz) nebyl nalezen samostatný korelát platný pro všechny vzorky korpusu, který by jednoznačně charakterizoval vzájemný vztah mezi objektivním měřením a percepčním hodnocením. V další fázi výzkumu bude vhodné zaměřit se na roli melodického průběhu, výslovnosti a hláskového tempa, hlubšímu zkoumání podrobit pauzy a členění obecně. Užitečné by bylo percepční hodnocení vzorků na úrovni promluvových úseků a jeho srovnání s hodnocením tempa příslušného mluvčího.

Jitka Veronková
Institute of Phonetics
Faculty of Arts, Charles University
jitka.veronkova@ff.cuni.cz

Petra Poukarová
Institute of the Czech National Corpus
Faculty of Arts, Charles University