Univerzita Karlova

Filozofická fakulta

Fonetický ústav

Bakalářská práce

Tomash Kalynych

Melodické charakteristiky ukrajinské češtiny

The melodic characteristics of Ukrainian Czech

Praha 2018

Vedoucí práce: doc. PhDr. Jan Volín, Ph.D.

Poděkování

Děkuji doc. PhDr. Janu Volínovi, Ph.D. za velmi cenné rady, trpělivost a čas, který mé práci věnoval. Poděkování také patří všem respondentům za účast v nahrávání a rodině a přátelům za pomoc a podporu.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 30. května 2018

Tomash Kalynych

Abstrakt

Tato práce se zabývá charakteristikami řečové melodie češtiny rodilých mluvčích ukrajinštiny. Pořídili jsme záznamy čtené řeči 18 ukrajinských respondentů. Pro srovnání ukrajinské skupiny se vzorkem rodilých mluvčích češtiny jsme extrahovali řadu dlouhodobých frekvenčních parametrů. Kromě jazyka byly zkoumany i další možné faktory, jako například věk, pohlaví, míra vystavení cílovému jazyku a míra motivace k jeho osvojení. Ukrajinské respondentky dosáhly nižších hodnot výšky tónu než rodilé mluvčí češtiny. Naopak, ukrajinští muži dosáhly vyšších hodnot stejných parametrů ve srovnání s českými respondenty. Možné vysvětlení rozdílů mezi skupinami rodilých mluvčích češtiny a ukrajinštiny spočívá v rozdílných kulturních stereotypech v příslušných jazykových komunitách. Rodilí mluvčí ukrajinštiny vykázali úžší intonační rozpětí, což naznačuje nejistotu nebo úzkost spojenou s mluvením v cizím jazyce. Mezi zkoumanými frekvenčními parametry a sociolingvistickými parametry, jako například postoj k většinové populaci a délka pobytu v oblasti cílového jazyka byla nalezena mírna korelace. Analýza jednotlivých promluv ukázala velkou míru shody mezi dvěma jazykovými skupinami, což naznačuje, že jazykový obsah promluv významně ovlivňuje vybrané frekvenční parametry.

Klíčova slova: prozodie, řečová melodie, základní hlasivková frekvence, cizinecký přízvuk, interference

Abstract

This study focused on the characteristics of speech melody of Czech as spoken by native speakers of Ukrainian. Samples of read-out speech were obtained from 18 Ukrainian respondents. A number of long-term frequency measures was extracted in order to compare the Ukrainian respondents with a sample of native Czech speakers. Other possible factors such as age, gender, exposure and motivation to learn the target language were also investigated. Ukrainian female speakers reached lower values of pitch level measures relative to Czech female speakers. At the same time, Ukrainian male participants reached higher values for the same measures compared to Czech male speakers. A possible explanation of the disparity between the native Czech and Ukrainian speakers lies in different cultural stereotypes in the two language communities. Ukrainian speakers demonstrated a narrower pitch span, suggesting uncertainty or anxiety associated with speaking a foreign language. Of the sociolinguistic parameters, moderate correlations with the investigated frequency measures were found for the attitude towards the majority population and the length of residence in the target language area. The analysis of individual utterances has revealed a high level of agreement between the two language groups, indicating that the linguistic content of an utterance has a significant influence on the chosen frequency measures.

Keywords: prosody, speech melody, fundamental frequency, foreign accent, interference

Contents

List of abbreviations

CI	confidence interval
CSI	cumulative slope index
Cz-F	Czech female speaker(s)
Cz-M	Czech male speakers
Hz, kHz	hertz, kilohertz
F0	fundamental frequency
L1/L2	first/second language
LOR	length of residence
ms	milliseconds
SD	standard deviation
ST	semitones
ST/s	semitones per millisecond
Uk-F	Ukrainian female speaker(s)
Uk-M	Ukrainian male speaker(s)

1. Introduction

Foreign accent is an important aspect of communication, as it is largely known to influence or even decide the success of an interaction. Previous research has shown a significant role of prosody in determining overall accentedness of foreigner speech (Kamiyama, 2004; Trofimovich & Baker, 2006) and its direct effect on listener comprehension (Braun et al., 2011; Holm, 2007). Moreover, unusual prosodic patterns may influence the perception of speaker's personality (Volín et al., 2014). Nevertheless, suprasegmental phenomena in foreign accent remain understudied (Trouvain & Gut, 2007, p. v).

The Czech Republic has witnessed an increase of its foreign population every single year since 2010, with a clear rising trend spanning over 20 years. This study focuses on the largest immigrant group in the Czech Republic, Ukrainians. The fall of the Soviet system was followed by a sharp increase in mobility in Eastern Europe, and the number of Ukrainian citizens in the Czech Republic has since grown to comprise nearly a quarter of the country's foreign nationals (Czech Statistical Office, 2017; Leontiyeva, 2010, p. 60). Still, there is a scarcity of research on Czech as a second language of Ukrainian native speakers.

The aim of this study is to provide field data for the speech melody of Ukrainian Czech, as well as compare it to native Czech reference samples. Additionally, the study aims to take sociolinguistic correlates of foreign accent into account, comparing obtained data to measures of exposure, motivation and other factors. This contribution to the field of second-language prosody may also inform practical disciplines such as teaching of Czech as a foreign language, speech technology or forensic science.

The first chapter outlines the theoretical aspects of this study. The chapter begins with an introduction to speech melody, its nature and functions. A discussion of foreign accent and its social and communicational aspects then follows. Next, recent research in second-language prosody is exposed through several selected studies. The final part of the chapter is an overview of the design and aims of the study.

The second chapter describes the methodological part of the study (the recording procedure, the respondents' profile and data analysis). The results of the analysis are reported in the third chapter and discussed in the fourth chapter.

2. Theoretical background

2.1 Speech melody in general

Speech melody (or pitch movements) stands for the variations of height of speech sounds in the course of an utterance. The term "intonation" in the narrow sense is equivalent to the definition above and will be used this way throughout this study. Importantly, speech melody cannot be measured directly, being a concept based in perception. However, speech melody is closely related to the physiological and acoustic property of fundamental frequency (hereafter F0). F0, in its turn, stands for the rate of vibration of vocal cords and the frequency of sound waves they generate measured in hertz (Hz). As will be seen, there are some crucial differences between F0 and speech melody.

The syllable (or the syllable nucleus) is considered to be the lowest unit carrying pitch information. However, as has been demonstrated by House, most syllables in running speech are perceived as having a single level tone even if there is a change in F0 during their course (Hermes, 2006, pp. 42–43). Thus, speech melody is constructed above the syllable level: the relative changes in F0 between "high" and "low" syllables combine to form "falling" or "rising" pitch movements.

F0 is influenced by a range of additional factors. Some of them, such as inherent properties of different types of segments and their segmental environment ("microprosody"), are inconsequential for the perception of speech melody (see, for example, Ravuri & Ellis, 2008). Those that do influence intonational functions include duration and loudness.

F0 is a property of periodical sound waves, meaning that it is limited to voiced segments. On the contrary, speech melody is perceived continuously across the entire utterance, including its unvoiced portions. To account for this difference, interpolation is used for the unvoiced parts of F0 contours (see section 3.2 below, p. 21).

Using their vocal cords, speakers can control speech melody, for example, maintaining it or moving it higher or lower. Speech melody can be said to have six basic functions (Volín & Bořil, 2013). The lexical function stands for speech melody being used, together with the segmental phonemes, to distinguish lexical meaning. This function is only used in tonal languages or languages with pitch accent and does not apply to either Czech

or Ukrainian. The grammatical function refers to the usage of speech melody to mark clause boundaries, distinguish declarative statements from interrogative ones, disambiguate unclear syntactic structures, and for other grammatical purposes. Czech has been described as having a repertoire of three grammatically relevant nuclear accents; in the phrase "přišel pozdě" a low falling accent would indicate a neutral declarative sentence ("he came late"), while a low rising accent would indicate a neutral interrogative sentence ("did he come late?") (Palková, 1994, pp. 309–312). Similar pitch movements could be used to distinguish the same sentence types in Ukrainian.

Some other functions of speech melody are more closely related to pragmatic meaning. The affective function relates to expressing attitude or emotion, like excitement, sarcasm or anger. This function has been associated with properties of melody such as pitch range and pitch direction (Chun, 2002, p. 49). For Czech, expressive utterances have been linked to a wider pitch range, more dramatic pitch changes and frequent use of rising pitch movements (Palková, 1997). The discourse function concerns the role of speech melody beyond the sentence level. Its use includes controlling interactive structure of a conversation, particularly by marking conversational turns, or expressing speaker's conversational intentions (e.g. expectation of an answer). The accentual function stands for speech melody being used to mark information structure in an utterance by emphasising its various parts. Specifically, speech melody can be used to distinguish given and new information, or sentence focus and background (Grice & Bauman, 2007). In Czech, as well as in Ukrainian, speech melody may be the only and obligatory mean of producing word emphasis in a sentence. In the Czech sentence "dnes odpoledne půjdeme do kina" ("we are going to the movies today in the afternoon"), a pitch accent on any of the stress groups, except the last one (its neutral position), would reliably mark the respective item as the focus element (Palková, 1994, pp. 300–301). The last is the indexical function, which is perhaps most relevant to studying foreign speech patterns, but has only received a limited amount of scholarly attention to this date (Chun, 2002, p. 67). It refers to speech melody acting as a marker of membership in a sociolinguistic group based on gender, age, regional origin, occupation and other parameters. Intonation may also allow hearers to evaluate speakers on temporary traits such as their physical and psychological state.

2.2 Foreign accent

Foreign accent is an inevitable part of the process of learning a second language (hereafter referred to as L2). Intonational foreign accent, which is the primary focus of this work, seems to be especially persistent as it is found even in very proficient, near-native L2 learners (Mennen, 2004). Foreign accent has been identified as a potential cause of communication problems, incomprehensibility and discrimination, all while being a very salient property of speech (see Derwing & Munro, 2009 for an overview). Prosodic features were shown to have a key effect in determining overall degree of foreign accent and its comprehensibility (see below). Despite the importance of prosody, it is allegedly quite difficult to describe and teach (Trouvain & Gut, 2007). It could be assumed that most L2 learners, including Ukrainian learners of Czech, receive little to no instruction in prosody of their target language. For example, the most popular textbook of the Czech language for Ukrainian speakers only contains very brief theoretical references to Czech intonation (Danylenko, 2007).

Studies in foreign accent generally utilise the concepts of accentedness, comprehensibility and intelligibility. Accentedness stands for the strength of foreign accent in an utterance, comprehensibility is defined as the ease of understanding it, and intelligibility - as the extent to which it was actually understood. Accentedness and comprehensibility are usually treated as subjective measures. This involves either a self-assessment by the speakers themselves or an assessment by other listeners. Intelligibility, on the other hand, is more commonly measured objectively. Some of the common tasks used to obtain intelligibility scores are transcription of the stimuli, questions related to their content or paraphrasing tasks. It has been found that accentedness does not necessarily imply low comprehensibility or intelligibility (Munro & Derwing, 1995). However, a foreign accent might produce other adverse effects for non-native speakers.

A meta-analysis of 20 studies comparing the evaluations of standard and non-standard accents by Fuertes, Gottdiener, Martin, Gilbert and Giles (2012) found that non-standard speakers are consistently rated lower on personality traits such as intelligence or attractiveness. It has also been demonstrated that a stronger accent might lead to a stronger negative rating on solidarity and speech characteristics (Ryan et al., 1977). It is important to note that the negative effect on L2 speakers' personality judgements was also reported for intonational errors (Trim, 1988; as cited in Grabe, 1998).

A non-standard accent may also provide ground for discrimination in employment, citizenship or education rights (Ng, 2007). A stigmatised foreign accent may have a considerable negative impact on its speaker's employment and career. Applicants with a Mexican-Spanish accent are judged as less suitable for the job and less likely to be hired or promoted to a higher position (Hosoda, Nguyen, & Stone-Romero, 2012). In Europe, language tests are now a common requirement for obtaining residence and citizenship rights, and thus, indirectly, for access to benefits such as healthcare and education (Van Avermaet, 2009). Passing a language test, including tasks in speaking proficiency, is also required to obtain a residence permit or citizenship in Czech Republic (see Osměra, 2009 for a critical discussion). Finally, foreign-accented speakers may even be seen as less credible than native speakers (Lev-Ari & Keysar, 2010).

A large volume of research has focused on determining possible sources for non-native pronunciation. An often cited explanation is the so-called Critical Period Hypothesis, beginning with Lenneberg (1967). It holds that brain maturation constrains second-language learning skills, and thus is the reason for a foreign accent in adult learners. Although there is by now an abundance of counterevidence to the hypothesis, age effects on L2 pronunciation have also been attested by its critics (Flege, Yeni-Komshian, & Liu, 1999). Recordings of Korean L2 speakers of English were evaluated for accentedness and knowledge of morphosyntax. The participants differed according to parameters such as age of arrival in the United States, years of education in the United States and use of English and Korean. Of these, age of arrival in the United States and language use had a significant effect on the degree of foreign accent. Similar effect for both age of arrival and language use was found in other studies (Flege 1998, 1999; Flege, Frieda, & Nozawa, 1997).

Sociopsychological factors are another potential predictor of accent strength. Similarly to age of arrival, the number of years spent in the target-language country is often shown to correlate with a more native-like accent. However, the supporting evidence across different studies is conflicting, as the number of years only provides a rough index of L2 experience (Piske, MacKay, & Flege, 2001).

Moyer (2007) proposes to examine the connection between foreign accent and language attitudes. The 42 learners of English and 8 native speakers, residing in the United States,

were surveyed on their language learning background and aspects of their orientation concerning their target language. Their recordings of various reading and speaking tasks were then rated for strength of foreign accent. Of the examined factors, comfort with the idea of cultural assimilation (in the USA), intention for long-term residence, desire and perceived ability to improve one's accent, and having both personal and professional reasons for learning English were found to correlate significantly with foreign accent ratings, alongside self-ratings of fluency, the age of onset and length of residence in the country.

Building on the study by Moyer (2007) and communication accommodation theory, Gluszek, Newheiser and Dovidio (2011) further examined the relation between accent strength and social identity. A hypothesised causal model predicted an influence of identification with the host country on accent strength. 117 non-native speakers of English completed questionnaires on their demographic background, identification with American culture, communication challenges and lack of social belonging in the U.S., and provided speech samples, which were later rated for strength of foreign accent. The more the participants identified with American culture, the less strong was their accent rated by themselves and by other listeners. Self-ratings of accent strength were also a significant predictor of communication challenges (as was the number of years spent in the United States).

2.3 Current research in L2 prosody

Compared to segmental foreign accent, prosodic foreign accent has been studied to a lesser extent. Still, the importance of its overall contribution to the perception of non-native pronunciation as well as its intelligibility and comprehensibility has already been demonstrated in research.

The role of non-native prosody in evaluations of naturalness has been investigated by Kamiyama (2004). Recordings of short phrases read in French were obtained from native French speakers and Japanese learners. The recordings were then manipulated to create a set of stimuli for two perception tests. In the first test, the recordings were resynthesised using native and non-native segments and prosody (F0 and duration) in different combinations. In the second test, F0 and duration were manipulated in order to make the recordings of Japanese speakers appear more native-like (and vice versa for native speakers). The resulting stimuli were evaluated for naturalness by 17 native speakers of French. Prosody

was concluded to influence evaluations of naturalness to a high extent, as native-like prosody significantly improved the naturalness of utterances with non-native-like segments. However, the relative importance of prosody and segments remained unclear.

A similar conclusion was reached by Vieru-Dimulescu and de Mareüil (2005) regarding the role of prosodic features (speech melody and timing) in the perception of foreign accent. A set of sentences that are very similar in Spanish and Italian was recorded as read by a group of 9 participants: native speakers and Spanish-Italian bilinguals. The choice of languages of comparable status for the experiment reduced potential interference of various social factors. Resynthesis was used to combine the segments and pauses of each recording with F0 and duration parameters from a recording in the same or the other language. 60 French, Italian and Spanish listeners judged the original and resynthesised sentences as Spanish or Italian, along with other options such as Italian-accented Spanish. Prosody influenced the subjects' judgements more than segmental cues. Still, the differences in ratings were not highly significant, suggesting a comparable contribution of prosodic and segmental factors to identifying a foreign accent.

Holm (2007) used a comparable methodological approach in an exploration of the relative roles of intonation and duration in overall intelligibility. Speech recordings of 14 learners of Norwegian (native speakers of 7 different languages) were manipulated so as to match the intonation and duration of recordings obtained from a native speaker. F0 contour stylisation and manipulation of phoneme durations were used to create native-like and non-native-like stimuli pairs. In a following perception test, native Norwegian listeners wrote down the words they could understand, resulting in an intelligibility index. Manipulating intonation resulted in an increase in intelligibility scores for almost all of the L1 groups, although the increase was only statistically significant for English and German L1 speakers, while manipulating duration had a significant effect for L1 speakers of French, Tamil and Persian.

The effect of intonation on comprehension has been further studied by Braun, Dainora and Ernestus (2011). Dutch stimuli with an unattested but articulatorily possible intonation contour were obtained by resynthesising sentences read by a native speaker and imitation of the chosen contour by the same speaker. Thus, the observed effect of intonation was independent of other factors such as non-native segmental features. 24 native speakers

performed word-monitoring, lexical decision and semantic categorisation tasks while listening to the two sets of stimuli. The tasks were designed to investigate the effect of manipulation on lexical access and semantic integration. The unfamiliar intonation contour was found to slow down the participants' performance in all three tasks. Notably, lexical access was affected, even though in Dutch (just as in Czech) pitch is not used to distinguish otherwise identical words.

The aforementioned studies demonstrate that the influence of intonation patterns on various parameters related to foreign accent is comparable to that of segmental cues or might even surpass it. Non-native speech melodies were found to hinder comprehension both for Norwegian, a language with pitch accent, and, in a differently designed experiment, for Dutch, where pitch accent is not employed (similarly to Czech).

Trofimovich and Baker (2006) examined social factors influencing the acquisition of L2 prosody. Sentences produced by 30 adult Korean learners of English differing in amount of L2 experience (having about 3 months, 3 years and 10 years of the United States residence) were acoustically analysed for accuracy in production of suprasegmentals characterising speech melody (stress timing, tonal peak alignment) and speech fluency (speech rate, pause frequency and pause duration). Production of stress timing was found to be related to the amount of L2 experience, while observed speech rate, pause duration and frequency were found to be related to the speaker's age of arrival to the United States. This implies that the learning process is different for different suprasegmentals. Production of peak alignment appeared to have no relationship to either amount of L2 experience or age of arrival – possibly, the differences were too subtle to be perceived and overcome by learners even within 10 years of experience. Notably, suprasegmentals related to speech fluency were found to predict accentedness ratings better than those related to speech melody. However, the ratings were based on low-pass-filtered recordings, and may not be necessarily comparable to the perception of clear speech. The demonstrated complexity of the acquisition of L2 suprasegmental features warrants further exploration.

Finally, speech melody of Czech and Czech-accented English was recently studied by Volín, Poesová and Weingartová (2015). One of the aims of the study was to provide reference values for "pitch profiles" of read Czech and English. The reference material consisted of recordings of news bulletins in Czech and English read by 32 professional news readers.

The primary research question, however, was related to L2 acquisition. The second part of the recorded material consisted of English sentences read by 8 non-professional native speakers of Czech and English. This second set was compared to the reference material to test the interference theory of L2 acquisition, which would predict the F0 values for Czech English to lie in between those measured for Czech and English reference data. Both samples were analysed for various measures of pitch level and pitch span, as well as for intonational downtrend. For the reference sample, significant differences were found for all three parameters: English speakers produced higher pitch level, wider pitch span and a steeper downtrend gradient than the Czech speakers. The pitch span of the Czech English material was the narrowest of all the examined groups. This indicates that there may be additional factors influencing a foreign accent apart from the interference from the speaker's native language. Further descriptions of second-language speech melody are desired, which motivates the task at hand.

2.4 This study

The main objective of this study was to collect and analyse fieldwork data on Ukrainian Czech (the variety of Czech found in native speakers of Ukrainian). The study aims to be an exploration of L2 speech melody characteristics.

Speech samples were to be obtained from 18 respondents. First, the respondents' speech recordings would be processed to extract F0 tracks, representing the F0 values during the course of a speech unit. Manually corrected and interpolated F0 tracks would provide an approximation of perceived speech melody. These F0 tracks would then be used to extract measures of pitch level and pitch span, which could then be compared to analogous data obtained from native speakers of Czech. The methodological procedure described here closely follows that used in the aforementioned work by Volín, Poesová and Weingartová (2015). However, this study includes an additional measure to observe the amount of F0 variation, the cumulative slope index (CSI), designed by Hruška (2016).

The acquisition of second-language prosody has been shown to be a very complex phenomenon. The learning process may be different depending on the suprasegmental feature in question (Trofimovich & Baker, 2006; see above). The second-language acquisition is largely viewed in terms of interference, with non-native speech interpreted as a result

of an incomplete transition from L1 to L2. However, recent evidence demonstrates that transfer from L1 cannot always account for the prosodic features observed in L2, suggesting that other factors may be present (Volín, Poesová, & Weingartová, 2015; Santiago & Delais-Roussarie, 2015; Gut & Pillai, 2015). This complexity has informed the present study in its additional aim, which was to analyse the variation in the speech melody parameters against the background of the speakers' sociolinguistic characteristics. Based on previous research in L2 acquisition (see Section 2.2 above), exposure to the Czech language and motivation to learn the Czech language were selected as the topics of a short questionnaire administered to the respondents. The effects of age, gender and length of residence in the Czech Republic were also investigated.

This study addresses the following research question: What differences in F0 measures are there between the Ukrainian Czech and native Czech groups, if any? Additionally, what differences can be found between groups defined by gender, exposure to Czech and other sociolinguistic parameters? The null hypothesis assumes no differences in speech melody of Ukrainian Czech and native Czech speakers. The alternative hypotheses account for possible differences. We are also interested in the contribution to the results of the individual speech units with respect of their linguistic contents.

3. Method

3.1. Collecting material

The material for the present study consists of recordings of a read-out passage in Czech. After the recording, additional sociolinguistic data were collected using a questionnaire. Read-out speech allows to control the content of obtained material, resulting in a better comparability of any extracted speech parameters. On the other hand, insufficient familiarity with L2 orthography can obscure the picture of L2 errors in respondents' speech. Read-out speech phenomena are not necessarily generalisable to spontaneous speech, as readers and speakers have different pragmatic goals.

The 18 participants (7 men and 11 women) were recruited for an interview about the life of Ukrainians in the Czech Republic. Of the participants, 11 were author's acquaintances; the remaining 7 were suggested by other participants or recruited through advertising on social media websites. 15 participants were recruited in Prague, the other 3 in Brno. All of them were fluent speakers of Czech, who had spent most of their lives in Ukraine and have lived in the Czech Republic for 1 to 7 years. The youngest respondent was 19 years old and had just finished high school. The oldest respondent was over 40 years old. For practical reasons, the majority of participants were students or had just finished studying at university. However, this group varied as to the type of education they were enrolled in. This was found to affect the way of learning Czech. For example, those respondents, who entered universities as doctoral students (4 respondents), tended to rely more on self-education. All of the respondents participated in the study on a voluntary basis. The list of participants together with their social characteristics is provided in Appendix A.

A sample of a corresponding size, gender distribution and social characteristics was assembled for reference purposes. The sample consisted of 18 native speakers of Czech (7 men and 11 women). The speakers in this group were all university students born in the area around Prague. The size of the sample limited the possibility of constructing a representative model of the Czech language community. For this reason, a homogenous sample reflecting a specific social profile was preferred.

The participants were recorded individually in a quiet room with a ZOOM H1 digital recorder, using a 48-kHz sampling frequency and 16-bit resolution. When possible,

the recordings were obtained in a sound-treated room at the Faculty of Arts, Charles University. In other cases, it was ensured that the recording conditions were as favourable as possible. Of the total 20 recordings, 2 have not been included in this study due to their quality being insufficient for F0 analysis. Before the recording, the participants were informed about the recording procedure and ensured that their anonymity would be protected (in Ukrainian). The elicitation procedure included recording some other speech material, however, that will not be part of the present study.

The recording lasted from 18 to 46 minutes (32.5 on average) and was conducted in Czech. The part analysed in the present study lasted from about 1 to 5 minutes (3 minutes on average). The participants were asked to read a short (151 words) written narrative describing an everyday conversation of a boy and his mother, and repeat any dysfluent sentences and phrases (the text is provided in Appendix B). All of the elicitation materials, including the narrative used in the reading task, were initially developed and provided by the Institute of Phonetics (Faculty of Arts, Charles University).

After the recording, the participants were instructed to fill out a short questionnaire in Czech consisting of 7 items. The questionnaire was aimed at evaluating the participants' exposure to the Czech language and motivation to learn it. The first part concerned exposure to Czech and contained the four following questions: "How often do you hear Czech?", "How often do you speak Czech?" (7-point scale: "several times a day", "a couple of times a day", "once in several days", "once a week", "once in several weeks", "once a month", "less often"), "How often do you watch films or other videos in Czech?" (7-point scale: "daily", "a couple of times a week", "once a week", "once a month", "a couple of times a year", "less often", "never"), "Do you live with Czechs?" ("yes" or "no"). The second part concerned motivation and included three statements and a scale of answers indicating a degree of agreement with them: "Czech is very important for my (future) work", "I have a positive attitude towards Czech people", "I would like to stay in the Czech Republic for a long period" (7-point scale: "completely agree", "agree", "somewhat agree", "difficult to decide", "somewhat disagree", "disagree", "completely disagree").

3.2. Pre-processing material

The original stereo 48-kHz recordings were first resampled to 32-kHz frequency mono signal, 16-bit resolution. Amplitude peaks containing noise (e.g. coughing, paper rustling) were removed or deamplified to be level with the rest of the recording. The recordings were then amplified by up to 15 dB in order to normalise the signal amplitude. The following steps in pre-processing the material were performed with the use of Praat (Boersma & Weenink, 2014) and additional scripts provided by the Institute of Phonetics in Prague. The text was split into 20 ideal breath groups, stretches of speech between where a speaker is expected to pause for taking breath (Lieberman, 1967). These breath groups are referred to as utterances in Section 4.4.

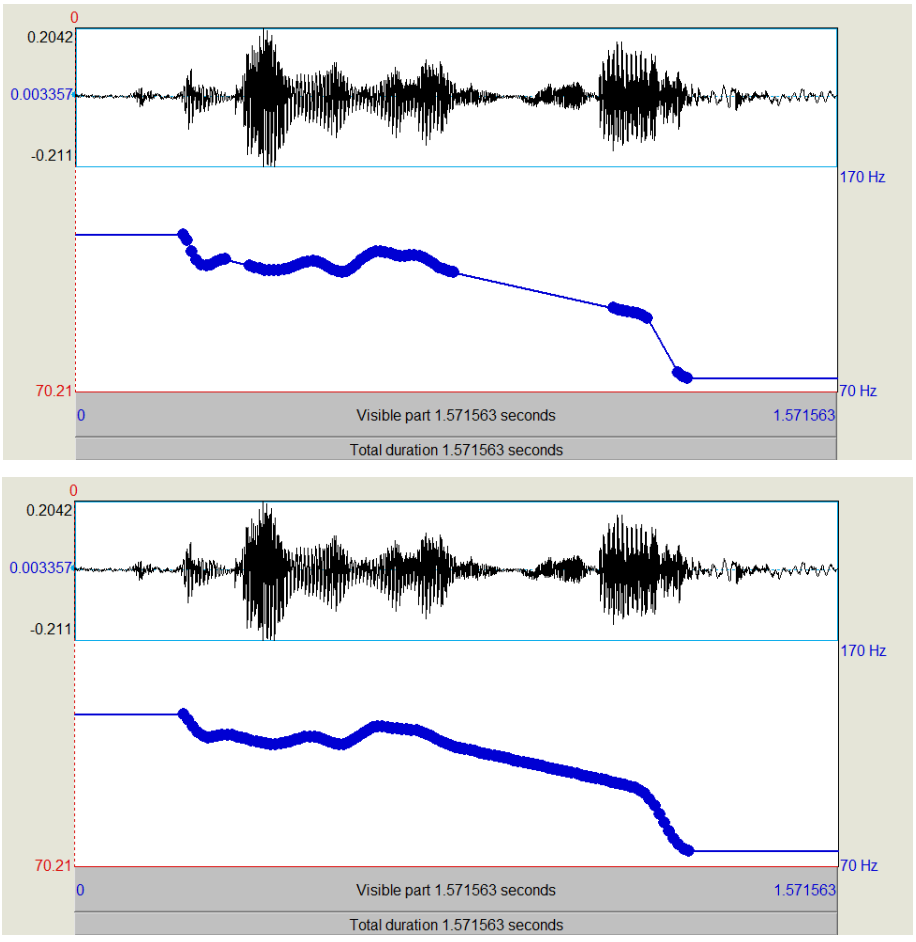

Figure 3.1. An example of an F0 contour before (above) and after interpolation (below).

Next, Praat's autocorrelation method was used to obtain F0 tracks with a 10-ms step between measuring points and convert them to Praat PitchTier objects smoothed by a 10-Hz filter. The PitchTier objects were manually corrected to reduce the impact of F0 measurement errors for voiceless segments or creaky phonation. Such errors also arose from the F0 extractor

predicting halved (subharmonic) or doubled (diplophonic) F0 values instead of the perceived pitch. Finally, the corrected PitchTier objects were interpolated (see Figure 3.1 above). In the previous step, the F0 contour is intermittent, only being defined for voiced segments. Interpolation involves including unvoiced segments by building straight lines between points with measured F0, producing an uninterrupted F0 contour. Interpolation results in a better approximation of the human perception of speech melody, which is also continuous.

3.3. Extracting data

To compare the material obtained from native Ukrainian participants and reference material from Czech native speakers, 13 descriptive frequency measures were calculated, characterising long-term properties of individual utterances. The measures were obtained using scripts for Praat and R (for CSI; see References) provided by the Institute of Phonetics. Three of the measures reflected the overall height of speaker's voice, or pitch level. Mean F0 represents the central tendency of a speaker's level. Similar to the mean, median F0 reduces the influence of outlying values on the result. F0 baseline, defined as the seventh percentile, i.e. the level above the lowest 7% of the F0 values in an utterance, serves as another measure of a speaker's level, more stable across various speaking styles and recording conditions (Lindh & Eriksson, 2007).

Five measures were chosen to represent pitch span, or the range of frequencies in a speaker's utterance. Variation range stands for the difference between the minimum and maximum F0 values. Two additional range measures reduce the influence of extreme F0 values: 80-percentile range discards the lowest and highest 10% of the F0 values, quartile range discards the lowest and highest 25%. Standard deviation describes the amount of variation in F0 values relative to the mean. Cumulative slope index (CSI), designed by Hruška (2016) and used by Hruška and Bořil (2017) for speech style identification, describes the amount of variation in F0 based on frequency differences between points in an F0 contour and the duration of the respective utterance.

Four measures were used to describe F0 range limits. As pitch span tends to have asymmetric properties, it is useful to inspect the difference in pitch variation below and above a speaker's average. The F0 range limits measures used here are: minimum F0, maximum F0,

10^{th} percentile and 90^{th} percentile of F0 values. The percentile measures omit extreme values to provide a more reliable picture of a speaker's pitch variation.

Finally, downtrend gradient represents the average rate of F0 declination in an utterance. The mean, the median, the baseline and the standard deviation were calculated in hertz (hereafter Hz); the three measures of variation range were calculated in semitones (ST); the downtrend gradient and the CSI were measured in semitones per second (ST/s). The measures of variation range were normalised by each speaker's average F0. For this reason, they represent deviations from individual average values.

The data extraction was followed by analysing individual values and means of the obtained measures. 95 percent confidence intervals (hereafter CI) are presented alongside average values for the speaker groups. Two-tailed t-test was used for between-group comparison ($\alpha = .05$). Statistical testing has revealed the ambiguous normality of a small portion of data (with p-values within the range of 0.01–0.05). Nonetheless, the versatility of the t-test allows applying it to the slightly skewed distributions at hand. Bar plots were used to illustrate long-term group means. A more informative description of the same data using dot plots can be found in Appendix C. For the sociolinguistic analysis, Kendall's tau-b correlation coefficients were calculated, optimal for ranked data with ties such as Likert scale values. The obtained coefficients were then tested for statistical significance.

4. Results

4.1. Measures of pitch level

The two groups of participants, Czech and Ukrainian native speakers, exhibit differences in the chosen pitch level measures. As specified in the paragraphs below, some of these differences were found to be statistically significant, others were insignificant. Also included were measures of F0 range limits (minimal and maximal F0, 10th and 90th percentile values of F0). Measures of range are discussed in the following section. The F0 measures discussed below are, in the following order: mean, median, baseline, minimum and maximum, 10th and 90th percentile.

The physiological difference between male and female speakers renders cross-gender comparison of raw mean F0 impractical. Therefore, mean F0 (as well as other measures in Hz) is analysed separately for male and female respondents. Nonetheless, our primary interest lies in the factor of language. Ukrainian female speakers displayed a lower mean F0 than Czech female speakers: $t(20) = 4.46$; $p < 0.001$. The direction of the difference was the opposite with the male participants: compared to the Czech male speakers, the Ukrainian male speakers demonstrated a higher mean F0: $t(12) = 2.38$; $p < 0.05$. Mean values in Hz for male and female participants in both groups are shown in Figure 4.1 below.

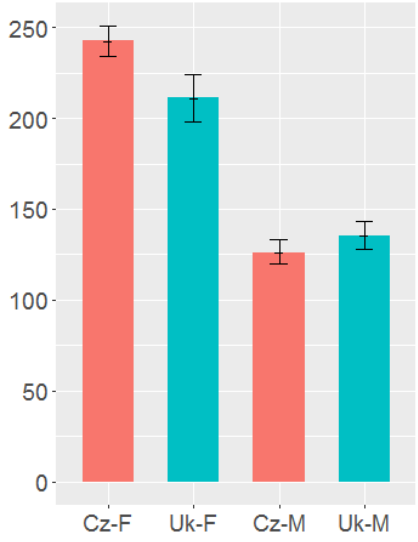

Figure 4.1. Mean F0 (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

The mean F0 values of female participants are expectedly higher than those produced by male participants in either group (given the raw, non-normalised values in Hz).

The 95% confidence intervals for the two groups of male speakers are overlapping; however, the difference in their mean F0 was statistically significant. Mean F0 in the native Czech group was 242 Hz (female; CI: 234 to 251) and 126 Hz (male; CI: 120 to 133). The native Ukrainian group had a mean F0 of 211 Hz (female; CI: 198 to 224) and 135 Hz (male; CI: 128 to 143).

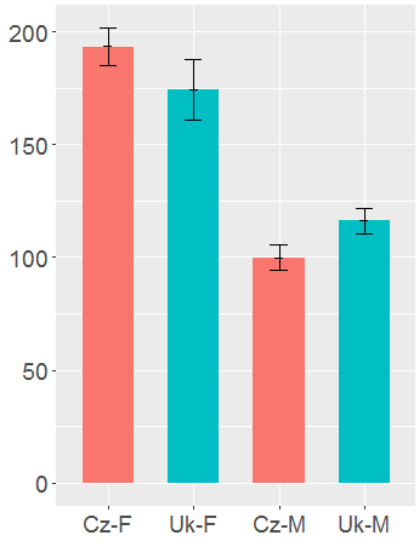

Figure 4.2. F0 baseline (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

Similar difference was observed for the other measures of pitch level, namely median F0 (female: $t(20) = 4.64$; $p < 0.001$; male: $t(12) = 2.17$; $p < 0.05$) and baseline value (female: $t(20) = 2.63$; $p < 0.05$; male: $t(12) = 5.10$; $p < 0.001$). The median F0 values for all speaker groups and individual speakers were close to their mean F0. The discrepancies for individual speakers were in the range of -4 to 9 Hz (-2 to 2 Hz for most). Compared to the two measures of central tendency, the baseline value, defined as the seventh percentile (see Section 3.3 above), was more similar for female speakers across the two language groups and less similar for male speakers, as can be seen in Figure 4.2 above. The individual values along with confidence intervals are listed in Table 4.1 below.

Several other measures describe the deviations in a speaker's pitch level above or below their average. As mentioned in Section 3.3, these measures have been normalised to every speaker's overall mean F0. The differences in minimal and maximal F0 values of the investigated groups can be seen in Figure 4.3 below.

Speaker group	Mean	95% CI	Median	95% CI	Baseline	95% CI
Cz-F	242	234 to 251	242	236 to 249	193	185 to 202
Cz-M	126	120 to 133	126	120 to 132	100	94 to 106
Uk-F	211	198 to 224	210	196 to 223	174	161 to 188
Uk-M	135	128 to 143	134	128 to 141	116	111 to 122

Table 4.1. Values and 95% confidence intervals for F0 level (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Baseline refers to the 7th percentile (see Lindh & Eriksson, 2007).

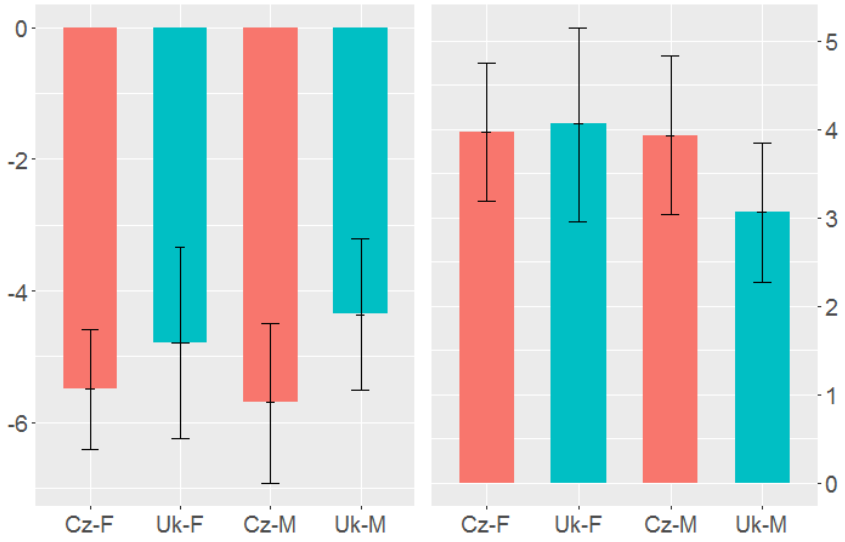

Figure 4.3. Minimal (left) and maximal (right) F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

Compared to Ukrainian speakers, the native Czech group diverged more from their average F0 in both directions. Czech speakers attained a minimum F0 of -5.6 ST (CI: -6.2 to -4.9) and a maximum F0 of 4.0 ST (CI: 3.4 to 4.5). At the same time, Ukrainian speakers only attained a minimum of -4.6 ST (CI: -5.5 to -3.7) and a maximum of 3.7 ST (CI: 3.0 to 4.4). The difference was more pronounced in the minimal values, where it approached statistical significance: $t(34) = 1.80$; $p < 0.09$ (maximum: $t(34) = 0.67$; $p > 0.05$). Ukrainian female speakers were found to produce both greater minimal ($t(20) = 0.91$; $p > 0.05$) and greater maximal F0 values ($t(20) = 0.15$; $p > 0.05$) than Czech female speakers. Nevertheless, the maximal F0 values were generally comparable, resulting in a smaller overall range in the Ukrainian female speaker group: $t(20) = 0.48$; $p > 0.05$. Ukrainian male speakers deviated less from their average values than Czech speakers. This is true for both maximal

(t(12) = 1.77; p ≈ 0.1) and minimal (t(12) = 1.96; p < 0.08) F0 values. As Figure 4.3 indicates, male and female Czech native speakers reached virtually the same maximal and minimal F0 values (maximum: t(16) = 0.07; p > 0.05; minimum: t(16) = 0.32; p > 0.05). The mean values in the native Czech group were a minimum F0 of -5.7 ST and a maximum F0 of 3.9 ST for the male speakers, a minimum of -5.5 ST and a maximum of 3.9 ST for the female speakers. At the same time, Ukrainian male speakers produced maximal and minimal F0 closer to the average than in the Ukrainian female group (maximum: t(16) = 1.48; p > 0.05; minimum: t(16) = 0.48; p > 0.05). In the Ukrainian group, female speakers produced a minimum of -4.8 ST and a maximum of 4.1 ST, while male speakers only reached a minimum of -4.4 ST and a maximum of 3.1 ST. However, only the differences between the two male groups were marginally significant.

Figure 4.4 below illustrates the inter-group difference in the 10th and 90th percentile values of F0. Again, the Ukrainian group stayed closer to their average than the native Czech group, with the difference in the 10th percentile value being statistically significant: t(34) = 2.35; p < 0.05 (90th percentile: t(34) = 1.20; p > 0.05). As with the maximal and minimal F0, both groups of Czech speakers attained highly similar values (90th percentile: t(16) = 0.17; p > 0.05; 10th percentile: t(16) = 0.11; p > 0.05). In comparison, Ukrainian female speakers were closer to their average in both directions. However, the difference with the Czech female group was not statistically significant (90th percentile: t(20) = 0.32; p > 0.05; 10th percentile: t(20) = 1.09; p > 0.05).

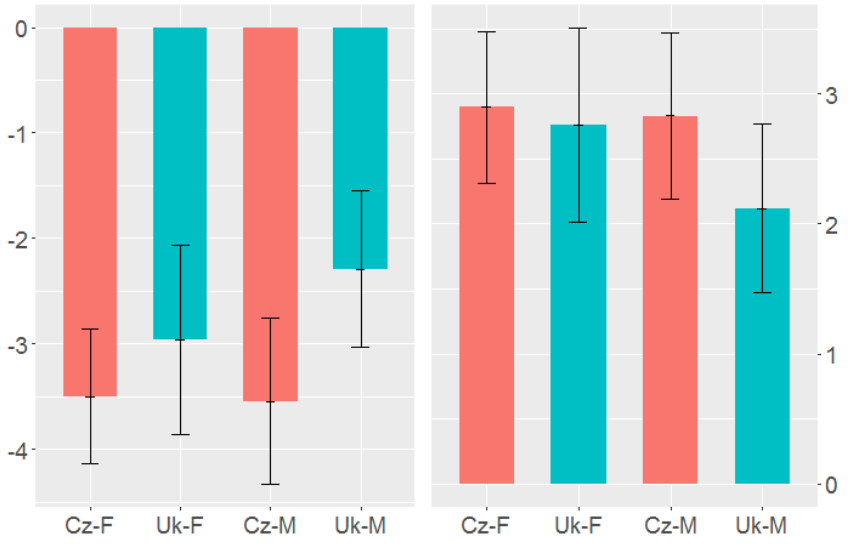

Figure 4.4. Average 10th (left) and 90th (right) percentile values of F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

Ukrainian male speakers are distinct from all other groups with the values closest to their average. As Figure 4.4 reveals, this trend becomes clearer when substituting absolute maximum and minimum F0 values for the 90th and 10th percentile. Namely, Ukrainian male speakers produced a higher 10th percentile value and a lower 90th percentile value compared to the Ukrainian female respondents (90th percentile: $t(16) = 1.35$; $p > 0.05$; 10th percentile: $t(16) = 1.19$; $p > 0.05$). Furthermore, their divergence from the Czech male speakers followed the same direction, producing a statistically significant difference in the 10th percentile value (90th percentile: $t(12) = 1.90$; $p < 0.09$; 10th percentile: $t(12) = 2.84$; $p < 0.05$). The individual values and confidence intervals for the normalised measures of F0 level can be found in Table 4.2 below.

A

Speaker group	Minimum F0	Maximum F0	10th percentile	90th percentile
Cz-F	-5.5	4.0	-3.5	2.9
Cz-M	-5.7	3.9	-3.6	2.8
Uk-F	-4.8	4.1	-3.0	2.8
Uk-M	-4.4	3.1	-2.3	2.1

B

Speaker group	Minimum F0	Maximum F0	10th percentile	90th percentile
Cz-F	-6.4 to -4.6	3.2 to 4.8	-4.1 to -2.9	2.3 to 3.5
Cz-M	-6.9 to -4.5	3.0 to 4.8	-4.3 to -2.8	2.2 to 3.5
Uk-F	-6.2 to -3.3	3.0 to 5.2	-3.9 to -2.1	2.0 to 3.5
Uk-M	-5.5 to -3.2	2.3 to 3.9	-3.0 to -1.6	1.5 to 2.8

Table 4.2. Values for F0 range limits (Table A) and 95% confidence intervals (Table B) (in ST) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian).

To summarise, native Czech speakers deviated more from their average pitch level when compared to native Ukrainian respondents. Czech speakers made more use of lower frequencies than Ukrainian female speakers, but their high pitch boundaries were similar. Czech speakers departed more below and above their average relative to Ukrainian male speakers. The difference between the two language groups as well as between the two groups of male respondents was more prominent below the average, where it was statistically significant (for the 10th percentile value). Another finding is that while both Czech groups were in agreement for all four measures, the Ukrainian groups displayed different behaviour.

Ukrainian female speakers attained values farther from their average than did Ukrainian male speakers. All of the mentioned trends were stronger for the measures of the 10th and 90th percentile, which omit extreme values. The group means were asymmetric – the observed deviations were larger below the average values than above them. However, several speakers produced an asymmetry in the opposite direction.

4.2. Measures of pitch span

Five F0 measures were chosen to represent differences in pitch span. The measures are discussed in the following order: variation range, 80-percentile range, quartile range, standard deviation (SD), cumulative slope index (CSI). The analysis of these measures corroborates the findings described above. The values of overall F0 range, based on the minimum and maximum F0, can be seen in Figure 4.5 below.

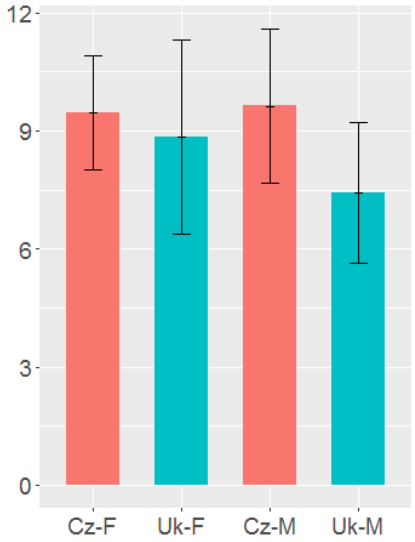

Figure 4.5. Average variation range of F0 (in semitones)
for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian).
Whiskers indicate 95% confidence intervals.

As seen in the figure, the Czech speakers produced a higher pitch range than the Ukrainian speakers, namely that of 9.5 ST (CI: 8.5 to 10.6) against 8.3 ST (CI: 6.7 to 9.8) for Ukrainians. However, this difference was not statistically significant: $t(34) = 1.40$; $p > 0.05$. Between the two gender groups, native Czech speakers displayed a very similar F0 range with the average of 9.5 ST for female and 9.6 ST for male speakers: $t(16) = 0.16$; $p > 0.05$. At the same time, the two native Ukrainian groups attained more distant values, with 8.9 ST for female and 7.4 ST for male speakers: $t(16) = 0.95$; $p > 0.05$. For both genders,

the Ukrainian group produced a smaller F0 range than the Czech group. Still, this difference was only marginally significant for the male speakers (female speakers: $t(20) = 0.48$; $p > 0.05$; male speakers: $t(12) = 2.04$; $p < 0.07$).

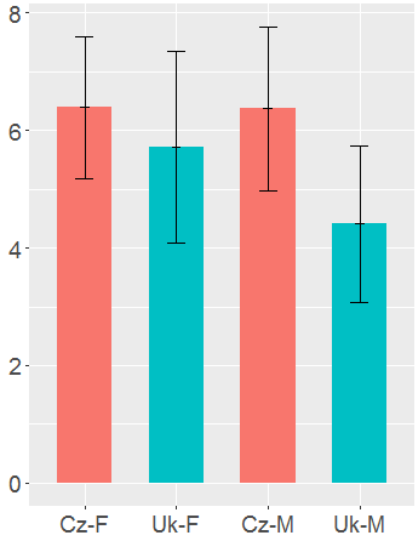

Figure 4.6. Mean values of the 80-percentile range of F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

The contrast becomes more pronounced when extreme F0 measures are discarded. The 80-percentile F0 range (between the 10th and the 90th percentile) was significantly different between the two languages for male speakers (both genders: $t(34) = 1.85$; $p < 0.08$; male speakers: $t(12) = 2.50$; $p < 0.05$; female speakers: $t(20) = 0.75$; $p > 0.05$). The quartile range (between the 25th and the 75th percentile) further emphasises the same tendency (both genders: $t(34) = 1.98$; $p < 0.06$; male speakers: $t(12) = 3.11$; $p < 0.01$; female speakers: $t(20) = 0.68$; $p > 0.05$). The variation in the 80-percentile range is shown in Figure 4.6 above.

Again, the two native Czech groups arrived at analogous values: $t(16) = 0.03$; $p > 0.05$. Ukrainian male speakers demonstrated a narrower range than Ukrainian female respondents, however, the difference was not statistically significant: $t(16) = 1.29$; $p > 0.05$. The quartile range showed very similar results within the two language groups (Czech speakers: $t(16) = 0.01$; $p > 0.05$; Ukrainian speakers: $t(16) = 1.68$; $p > 0.05$).

Figure 4.7 below provides the values for standard deviation (SD), measured in Hz. We are aware that SD is primarily meant to be used for symmetrical distributions. F0 values are a type of asymmetrical data, being skewed to the right, as demonstrated in the discussion

of range limits in the previous section. As with the other raw measurements in Hz, SD is treated separately for the two gender groups. The general difference between male and female speakers is apparent in Figure 4.7. Czech speakers produced a SD of 33.4 Hz (female; CI: 26.8–39.9) and 17.4 Hz (male; CI: 13.6–21.2). Ukrainian speakers reached 26.5 Hz (female; CI: 19.0–34.1) and 13.4 Hz (male; CI: 9.3–17.5). Clearly, the mean values in the Ukrainian group were lower than those in the Czech group for both genders. However, this difference was not statistically significant (female speakers: $t(20) = 1.52$; $p > 0.05$; male speakers: $t(12) = 1.74$; $p > 0.05$).

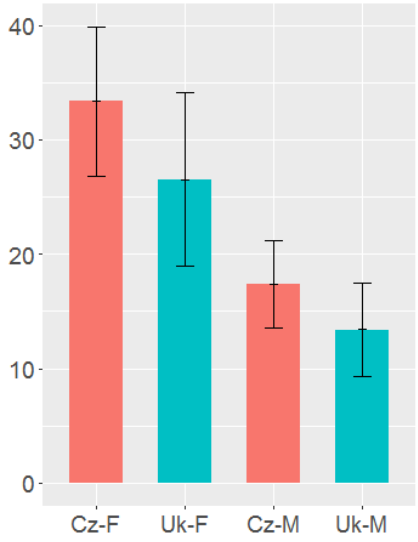

Figure 4.7. Standard deviation of F0 (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

Figure 4.8 below illustrates the mean values of the cumulative slope index (CSI), measured in semitones per second (ST/s). This parameter, designed by Hruška (2016), is based on the rate and intensity of F0 changes. In other words, CSI represents the variability of a F0 contour. The mean CSI produced by the native Czech speakers was 14.4 ST/s (CI: 12.9 to 15.9), while the Ukrainian respondents produced the value of 10.6 ST/s (CI: 8.5 to 12.6). Subdivided by gender, the Czech group reached 14.1 ST/s (female; CI: 12.1 to 16.1) and 14.8 ST/s (male; CI: 12.1 to 17.5). Correspondingly, Ukrainian speakers reached 11.1 ST/s (female; CI: 7.8 to 14.4) and 9.7 ST/s (male; CI: 7.2 to 12.1). The range of the values in our data is similar to that described by Hruška and Bořil (2017), with two thirds of all individual values ranging from 5 to 15 ST/s.

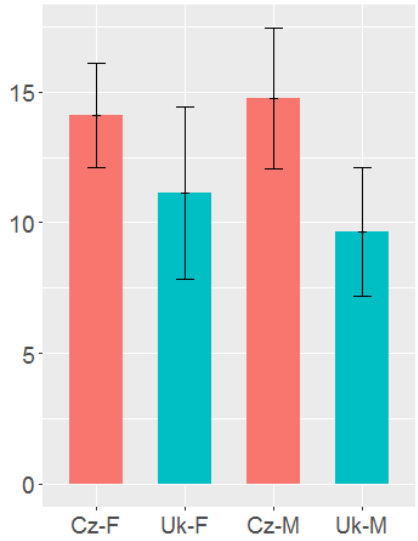

Figure 4.8. Cumulative slope index (in semitones per second) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

As can be seen, for both male and female speakers the mean CSI values were higher in the Czech group. The difference was statistically significant between the two language groups as well as between the two male groups (both genders: $t(34) = 3.20$; $p < 0.05$; male: $t(12) = 3.42$; $p < 0.01$; female: $t(20) = 1.73$; $p \approx 0.1$). Within the native Czech group, male speakers produced a higher mean compared to the female speakers. Among Ukrainian speakers, it was the female group that reached the higher value. However, both differences were miniscule and statistically insignificant (Czech speakers: $t(16) = 0.46$; $p > 0.05$; Ukrainian speakers: $t(16) = 0.73$; $p > 0.05$). The individual values and confidence intervals for all pitch span measures can be found in Table 4.3 below.

To summarise, native Ukrainian speakers consistently exhibited lower measures of pitch span when compared to the native Czech group. While Czech male and female speakers displayed very little difference in their values, there was more variability within the Ukrainian group, with Ukrainian male speakers producing the lowest measures out of all respondents on average. Consequently, the difference between the Czech and Ukrainian male groups was statistically significant for all pitch span parameters (except the marginally significant difference in variation range). These findings are analogous to the analysis of F0 range limits in Section 4.1 above.

A

Speaker group	Variation range (ST)	80-perc. range (ST)	Quartile range (ST)	SD (Hz)	CSI (ST/s)
Cz-F	9.5	6.4	3.4	33.4	14.1
Cz-M	9.6	6.4	3.4	17.4	14.8
Uk-F	8.9	5.7	3.0	26.5	11.1
Uk-M	7.4	4.4	2.1	13.4	9.7

B

Speaker group	Variation range (ST)	80-perc. range (ST)	Quartile range (ST)	SD (Hz)	CSI (ST/s)
Cz-F	8.0 to 10.9	5.2 to 7.6	2.7 to 4.1	26.8 to 39.9	12.1 to 16.1
Cz-M	7.7 to 11.6	5.0 to 7.8	2.7 to 4.1	13.6 to 21.2	12.1 to 17.5
Uk-F	6.4 to 11.3	4.1 to 7.3	2.2 to 3.9	19.0 to 34.1	7.8 to 14.4
Uk-M	5.6 to 9.2	3.1 to 5.7	1.4 to 2.8	9.3 to 17.5	7.2 to 12.1

Table 4.3. Values of F0 span (Table A) and 95% confidence intervals (Table B) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian).

4.3. Downtrend gradient

Lastly, downtrend gradient, measured in semitones per second, stands for the average decline of F0 values in the course of an utterance. A lower gradient value indicates a larger decline.

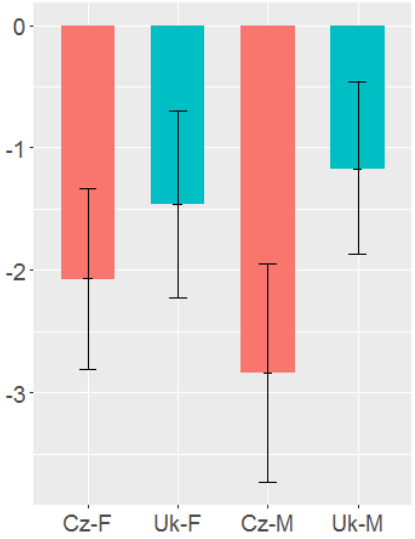

Figure 4.9. Average values of downtrend gradient (in semitones per second) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

Average values for the investigated groups can be seen in Figure 4.9 above. Czech speakers decrease their F0 by -2.4 ST/s (CI: -2.9 to -1.8) – more than Ukrainian speakers with the value of -1.3 ST/s (CI: -1.8 to -0.9). The largest gradient was produced by Czech male speakers, whose average decline was -2.8 ST/s (CI: -3.7 to -2.0), while Czech female speakers only reached -2.1 ST/s (CI: -2.8 to -1.3). The direction of the difference was the opposite in the Ukrainian group. While Ukrainian female respondents reached a value of -1.5 ST/s (CI: -2.2 to -0.7), it was the Ukrainian male speakers who produced the smallest average decline of -1.2 ST/s (CI: -1.9 to -0.5). The difference between the two language groups across gender as well as for the two male groups was statistically significant (both genders: $t(34) = 2.95$; $p < 0.01$; male: $t(12) = 3.61$; $p < 0.01$; female: $t(20) = 1.28$; $p > 0.05$). Contrary to the measures discussed in the previous two sections, the Czech group was less homogenous than the Ukrainian group. Still, the variation within each of the two groups was not statistically significant (Czechs: $t(16) = 1.52$; $p > 0.05$; Ukrainians: $t(16) = 0.60$; $p > 0.05$). All individual speakers except one produced a negative mean gradient.

4.4. Analysis by utterance

The previous sections considered the variation in chosen long-term F0 measures in the course of the whole recording. An analysis of this kind can certainly provide insights about global pitch differences between the investigated groups (see Chapter 5 below, p. 46). However, it is only at the local level that oppositions in the actual use of speech melody as part of the language code can be discerned. The local variance in the chosen pitch measures is illustrated in the figures below in this section. In each figure, dots represent group mean values for one of the 20 corresponding utterances (see Section 3.2, p. 20). The dots and the lines connecting them allow inspecting the changes in F0 measures over the course of the text.

The discussion of pitch parameters is followed by an analysis of F0 contours. Individual contours were examined visually and acoustically in order to identify the variance in choice and production of melody for each utterance. The findings are related to inter-group differences in F0 measures. In the examples of contours provided below, frequency range and duration can be found within the respective figure.

Group variation in mean F0 can be seen in the Figure 4.10 below. Due to physiological gender differences, it was preferred to display the data for the two gender groups separately.

As the utterance level reveals, there was a high amount of agreement in mean F0 values across both gender and language groups. All four groups appear to exhibit a lowering trend towards the end of the text. Downtrend also occurred in measures of pitch range limits and pitch span, described below. A similar lowering of pitch termed "supradeclination" has been observed at paragraph level for read-out material (see Wichmann, 2000). The most prominent places of divergence in mean F0 values were Utterances 5, 8, 11 and 14. The trend for median F0 values was in close agreement with mean F0. However, unlike the mean, the median value for Utterance 1 was higher relative to Utterance 2 in the Czech group, adding another place of discrepancy between the two language groups.

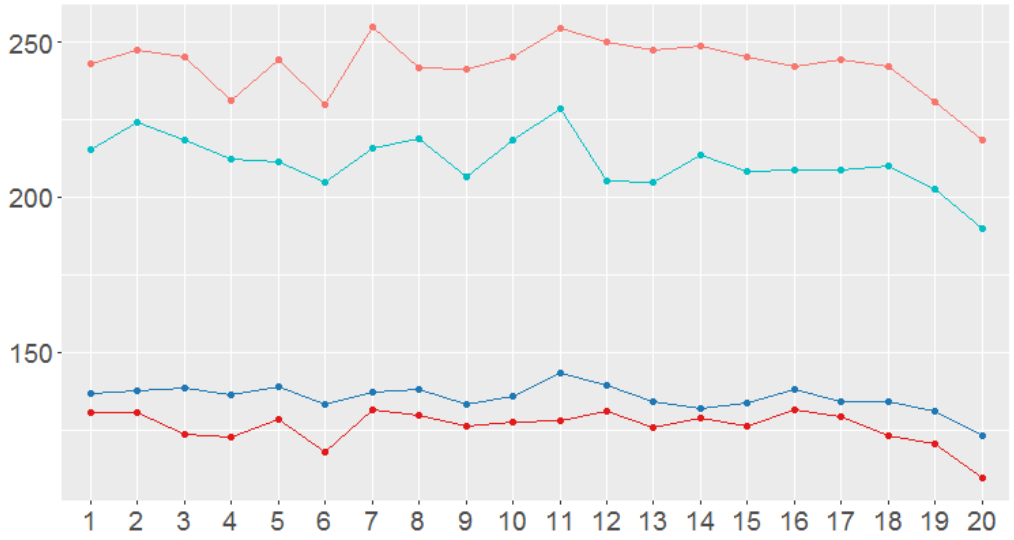

Figure 4.10. Mean F0 (in Hz) per utterance for two genders (darker colours – male, lighter colours – female) and languages (red – Czech, blue – Ukrainian).

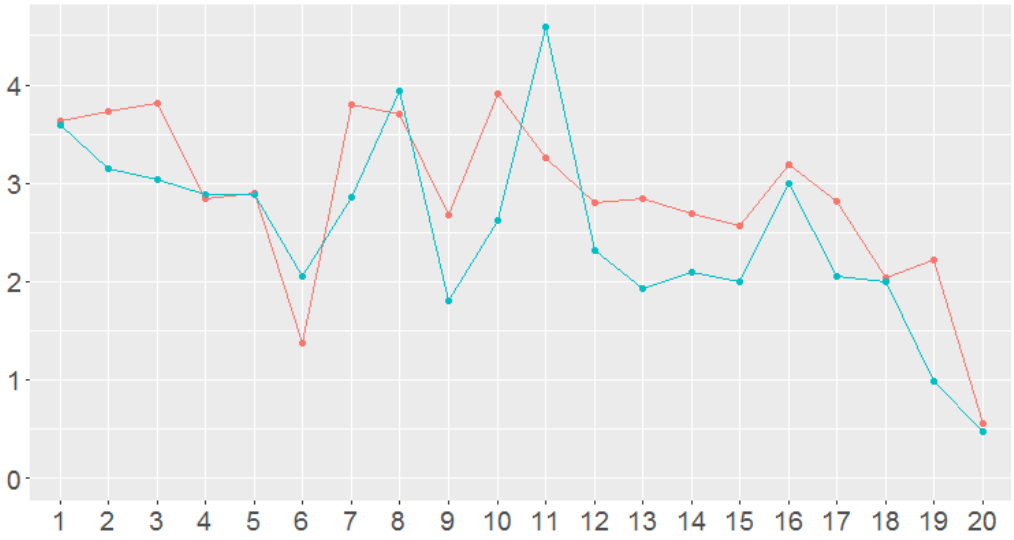

Figure 4.11. Average 90th percentile values of F0 (in ST) per utterance for two languages (red – Czech, blue – Ukrainian).

Figure 4.11 above represents group variance in the 90th percentile values of F0. Again, for the two groups, the relationship between the utterances and the values observed was largely parallel, including a general downward trend. In most cases, native Czech and Ukrainian respondents deviated from their mean values to a similar extent, the maximal observed difference being 1.3 ST (Utterance 10). Divergence between the two groups was most prominent for Utterances 3, 7, 11 and 19. The maximum F0 values followed a similar pattern.

The average 10th percentile values of F0 per group and utterance are illustrated in Figure 4.12 below. Again, significant agreement is visible despite the difference in average values between the two groups. Given the average distance between the values for Czech and Ukrainian respondents, the most divergent values are found for Utterances 4, 12 and 16. The minimum F0 values for the two groups were less concordant. However, a higher amount of variation is expected, since minimum F0 also includes the extreme F0 measures below the 10th percentile.

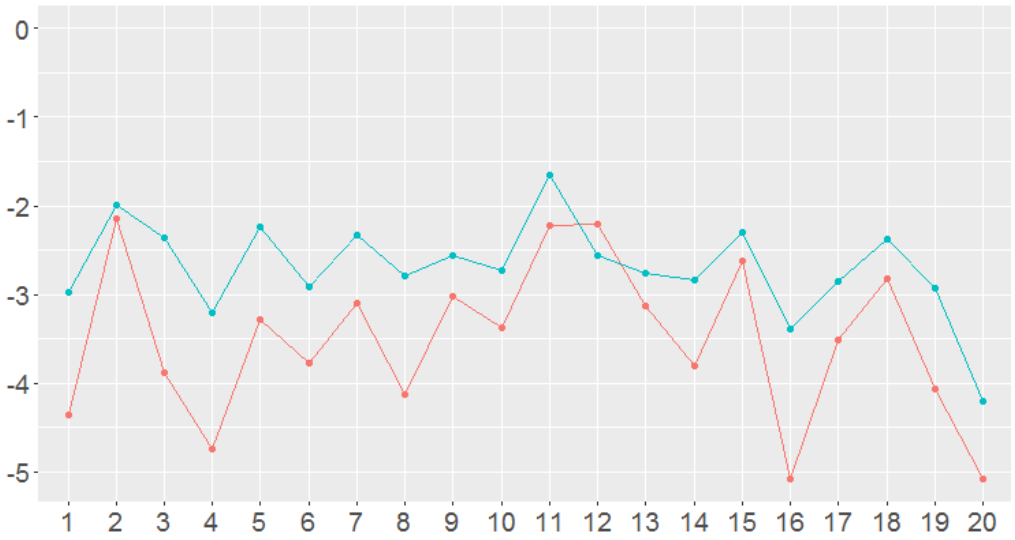

Figure 4.12. Average 10th percentile values of F0 (in ST) per utterance for two languages (red – Czech, blue – Ukrainian).

Group values of the CSI per utterance are represented in the figure 4.13 below. The local changes for the two language groups appear to coincide considerably. Differing trends can be seen for Utterances 3, 7, 8, 11 and 16. As with the other measures, the duration of F0 contours seemed to have an influence on the values observed. In a short utterance, a pitch movement might extend over a significant portion of the contour, resulting in a technically correct but misleading value. This was the case for Utterance 11, a short polar

question, where an outlying point of 48.9 ST/s was produced by a native Ukrainian speaker (Uk-F-10). Still, the group difference for Utterance 11 remained even after the speaker was removed from the sample (see Figure 4.14).

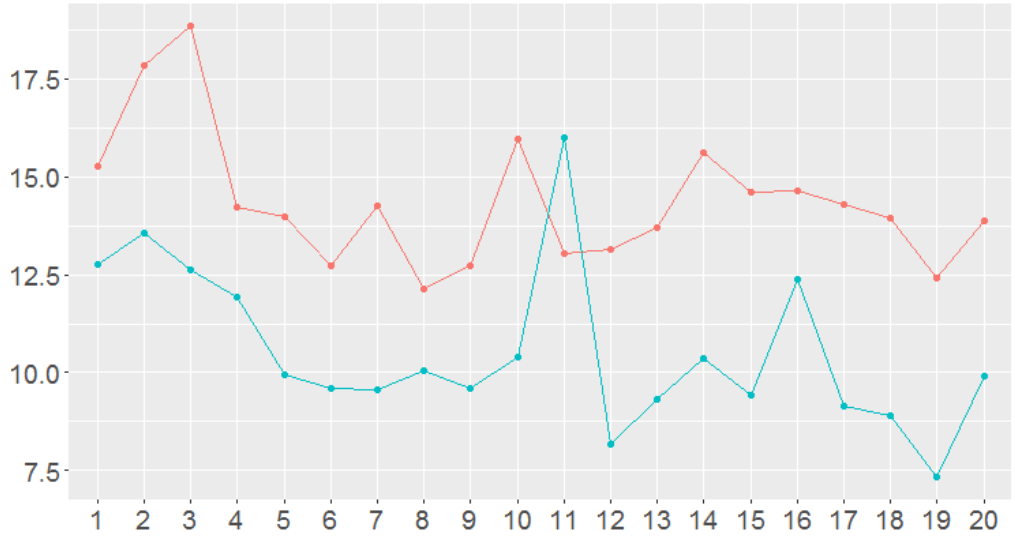

Figure 4.13. Cumulative slope index (in semitones per second) per utterance for two languages (red – Czech, blue – Ukrainian).

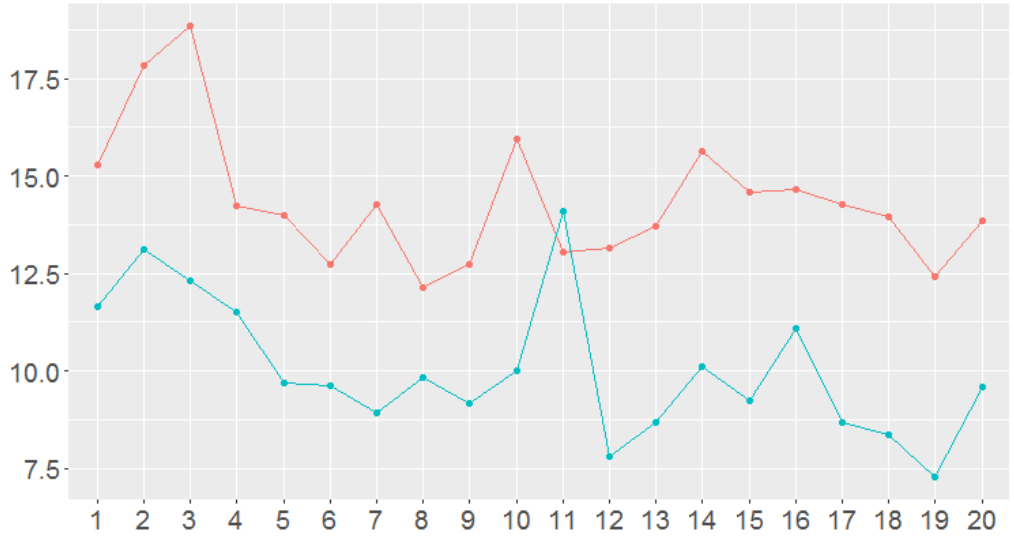

Figure 4.14. Cumulative slope index (in semitones per second) per utterance for two languages (red – Czech, blue – Ukrainian), speaker Uk-F-10 removed.

Finally, Figure 4.15 below illustrates the group values for downtrend gradient per utterance. As can be seen, the values for gradient demonstrate the highest amount of inter-group agreement out of all the F0 measures. This suggests a close relationship between the magnitude of gradient and the content of individual utterances. The largest divergence in gradient values was found in Utterances 3, 16 and 19.

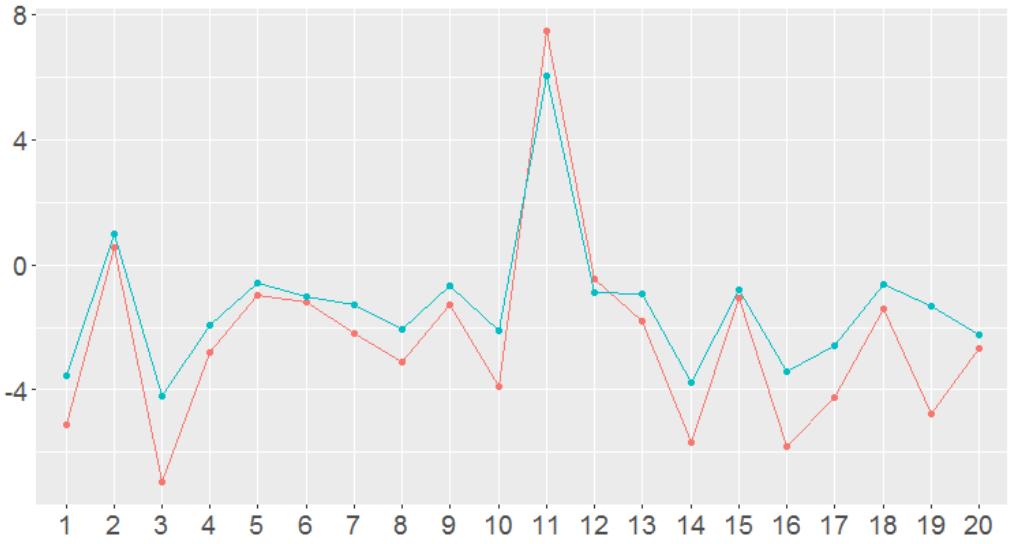

Figure 4.15. Downtrend gradient (in semitones per second) per utterance for two languages (red – Czech, blue – Ukrainian).

To summarise, the F0 measures for the two language groups were most divergent for Utterances 1, 3, 4, 5, 7, 8, 11, 12, 14, 16 and 19. These Utterances (as well as Utterance 2) are analysed below, apart from Utterances 1, 4, 5 and 12.

Utterance 2, "Milánku, už máš hotový úkol?" ("Milánek, have you done your assignment already?"), starts with a vocative. In colloquial Czech of Central Bohemia, it seems to be usually marked by a rise-fall melody. This melody was followed by most native Czech speakers in our sample and was reflected as a local peak in their F0 contours. At the same time, most Ukrainians produced the same portion of the utterance without a prominent rise. Examples of different F0 contours can be seen in Figure 4.16 below. The majority of speakers in both groups resolved the utterance with a fall-rise melody aligned with the last stress group, the word "úkol" ("assignment"). This is the standard polar question melody in Czech (Palková 1994, p. 308). One of the participants, a Ukrainian speaker, displayed the non-standard rise-fall melody known as the "Prague question" (see Palková 1994, p. 313). The values of F0 baseline, 10^{th} percentile and CSI obtained for Utterance 2 were relatively higher for Czech than for Ukrainian speakers. This was possibly related to an additional F0 peak.

Utterance 3, "Kdy ho budeš psát?" ("When are you going to write it?"), is a content question. Most speakers from both language groups adhered to the same melody, producing a rise aligned with the initial stress group "kdy ho", immediately followed by a fall. This is

the unmarked melody for Czech content questions (Palková 1994, p. 307). The divergence in mean F0 for the two male groups was seemingly caused by an initial upstep produced by one of the Ukrainian male speakers. The divergence in the 90th percentile values, CSI and other pitch span measures was limited to the two female groups, with Ukrainian female speakers attaining relatively low values, reflecting a narrower pitch range. The reasons for this are unclear.

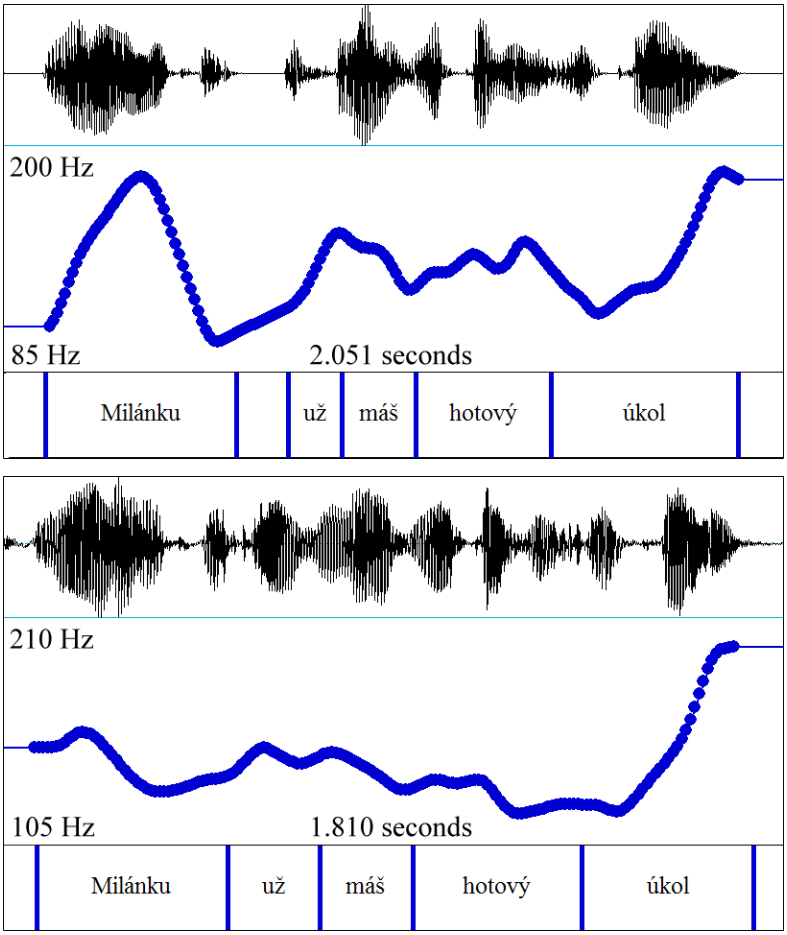

Figure 4.16. Utterance 2 "Milánku, už máš hotový úkol?". F0 contour examples of a native Czech (Cz-M-3, above) and a Ukrainian speaker (Uk-M-5, below).

Utterance 7, "Vlastně jsem tě chtěl poprosit jestli mi s tím nepomůžeš" ("Actually, I wanted to ask you if you'd help me with that"), consists of two clauses. For native Czech speakers, each clause contained a rise-fall movement, the fall being used to mark lexical prominence. Thus, there were two pitch falls aligned with the words "poprosit" ("ask") and "nepomůžeš" ("help"), as well as steady rising before them. However, about a half of the Ukrainian speakers only produced the first of the two rise-fall movements, their pitch moving down thereafter. Two of these speakers produced a minor rise aligned with the penultimate syllable,

"mů" in "nepomůžeš". These differences were a possible reason for the lower CSI in the Ukrainian group. Examples of F0 contours are provided in Figure 4.17 below.

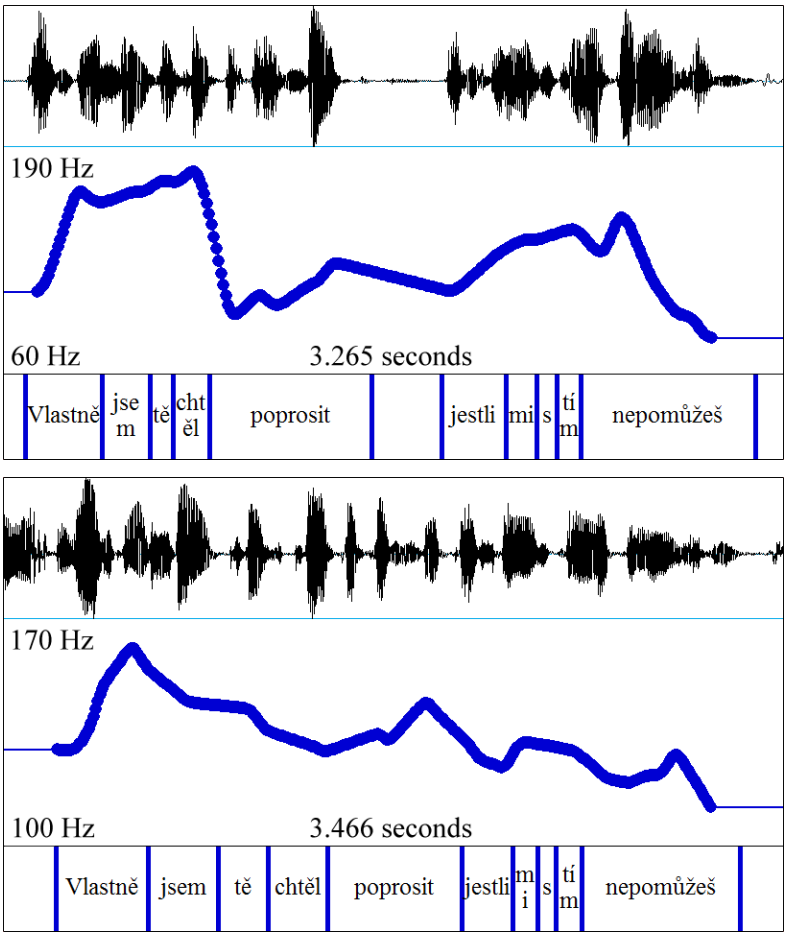

Figure 4.17. Utterance 7 "Vlastně jsem tě chtěl poprosit jestli mi s tím nepomůžeš". F0 contour examples of a native Czech (Cz-M-1, above) and a Ukrainian speaker (Uk-M-2, below).

For Utterances 8 and 14, visual inspection of F0 contours revealed no features consistently different for any group of speakers. Acoustic analysis suggested a relationship between shifts in F0 measures and degree of pitch reset. Both preceding utterances, 7 and 13, were produced with a falling end by all respondents. The Utterances 8 and 14 followed with an initial upstep, varying in relative magnitude for different speakers. The amount of pitch reset could be perceived as the degree of engagement. This observation was supported by analysing the individual speakers' F0 measures. An audibly prominent upstep correlated with a relative rise in mean F0 and the 90th percentile of F0 values. The presence of a prominent upstep was more common for Ukrainian female speakers. This contributed to the inter-group difference for the measures mentioned above.

Utterance 10, “Jak dlouho nám ten úkol zabere?” (“How long will the assignment take us?”), is the second content question in the text. As in the Utterance 3 above (p. 37), the majority of both Czech and Ukrainian speakers produced a rise aligned with the emphasised word, “dlouho” (“long”), followed by a fall (see Figure 4.18 below).

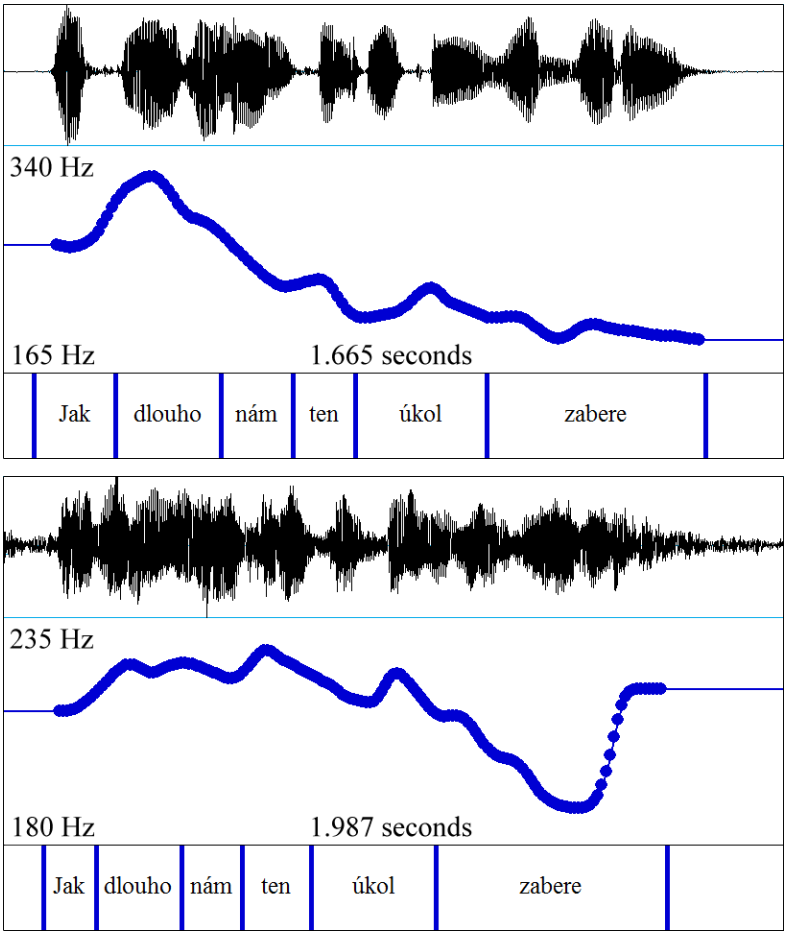

Figure 4.18. Utterance 10 “Jak dlouho nám ten úkol zabere?”. F0 contour examples of a native Czech (Cz-F-6, above) and a Ukrainian speaker (Uk-F-6, below).

Two Ukrainian female speakers diverged in placement of the F0 peak, producing the fall after the word “jak” (“how”) or after the first syllable in “dlouho” rather than the last one. A different melody was obtained from three other Ukrainian female speakers. Their F0 contours featured a fall-rise movement aligned with the final stress group in the utterance, the word “zabere” (“take”). However, the three speakers differed in their production of stress and pitch rise. Two speakers produced an initial stress, followed by a rise aligned with the second syllable in the stress group. The other speaker produced both stress and pitch rise on the last syllable, which is stressed in the related Ukrainian word (see Figure 4.18). A final rise or rise-fall melody was observed for three native Czech speakers, resulting in a high mean value of minimum F0 in the Czech group. Relative to Utterance 9, the median F0 value was

higher among Ukrainian respondents, but lower for native Czechs. This followed from the occurrence of an initial upstep in Utterance 9, produced mainly by Czech speakers.

In Utterance 11, "Bude to těžké?" ("Is it going to be difficult?"), the majority of speakers from both language groups conformed to the standard Czech polar question melody. The speakers produced a strong rise aligned with the final syllable of the last stress group, here "ké" in the word "těžké" ("difficult"). Only one speaker in each language group displayed the non-standard "Prague question" melody (see comments on Utterance 2 above, p. 37). Although both native Czech and Ukrainian respondents aimed at the same melody, the details of their contours differed. Most native Czechs produced the penultimate syllable ("těž" in "těžké") higher relative to the preceding stretch of the utterance, resulting in a gradual final rise. Most Ukrainians, however, produced the same syllable lower than the preceding part, the result resembling a fall-rise melody. This distinction is exemplified in Figure 4.19 below.

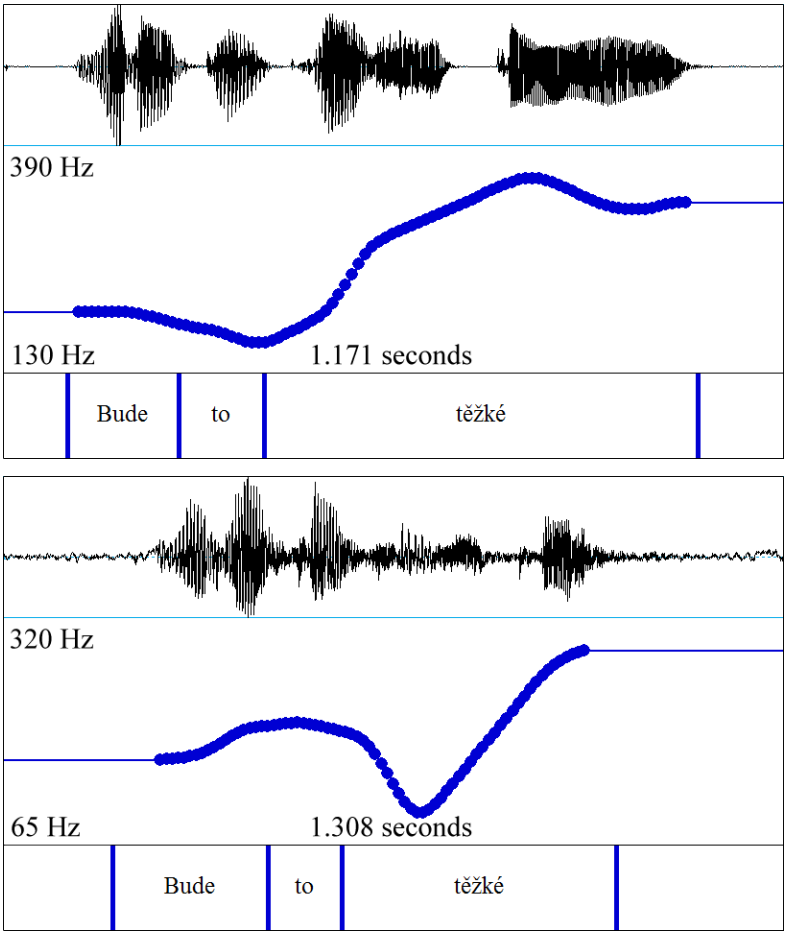

Figure 4.19. Utterance 11 "Bude to těžké?". F0 contour examples of a native Czech (Cz-F-11, above) and a Ukrainian speaker (Uk-F-10, below).

Another identified inter-group difference concerned segment duration. The majority of Ukrainian participants did not reach the native Czech duration for the utterance-final long vowel. On average, the Ukrainian group attained a mean duration of about 160 ms against the native Czech mean of about 220 ms. On the other hand, two Ukrainian speakers reached durations exceeding those found in the entire Czech group, suggesting target overshoot. Given the small length of Utterance 11, the duration of a single vowel can substantially influence any measures derived from the resulting F0 contour.

The Ukrainian group reached a relatively high CSI value in Utterance 16, "A potom půjdu hrát fotbal" ("And then I'll go play football"), while the Czech group reached a moderate value (see Figure 4.13 above). Of the native Ukrainian speakers, four produced the utterance with a final pitch rise rather than the fall expected for a declarative sentence, contributing to the high gradient value observed in the Ukrainian group. However, mean CSI as well as the 10th percentile for the Ukrainian group remained high even after the diverging speakers' values were removed.

The two language groups also demonstrated different trends for Utterance 19, "Jen míč budu muset přifouknout" ("I'll just have to pump the ball up"). Relative to Utterances 18 and 20, the Ukrainian group produced lower values of all of the pitch span measures, while the values in the native Czech group were higher. The largest inter-group difference in the values of downtrend gradient was also observed for Utterance 19, with native Czechs reaching the lower value. The majority of Ukrainian speakers produced filled and unfilled pauses within the utterance, suggesting a restriction in reading or comprehension. As most pauses occurred in the same place (before the word "přifouknout"), the cause for hesitation possibly lies in the lexical domain. Fatigue is another likely source of the observed difficulties. The uncertainty associated with hesitation is seemingly reflected in the narrower pitch span and smaller downtrend within the Ukrainian group.

4.5. Sociolinguistic analysis

Statistically significant correlations were found between the F0 measures produced by speakers of Ukrainian Czech and their sociolinguistic parameters. The participants' answers on the questionnaire evaluating exposure to the Czech language and motivation to learn it

are summarised below. Appendix A lists the age and length of residence in the Czech Republic for each participant.

Four questions were used to measure the subjects' exposure to the Czech language. Of the 18 participants, 15 stated they heard Czech several times a day, while 3 stated they only heard it a couple of times a day. Concerning active use of Czech, 10 participants stated they spoke it several times a day, 7 – a couple of times a day, the one remaining participant – once in several days. The participants' responses about their frequency of watching films or other videos in Czech were distributed the following way: daily – 1; a couple of times a week – 3; once a week – 6; once a month – 5; a couple of times a year – 2; less often – 1; never – 0. Only two respondents had Czech cohabitants at the time of the survey. The second part of the questionnaire contained three questions and aimed at assessing motivation to learn the Czech language. Its results are presented in Table 4.4 below.

	Czech is very important for my (future) work	I have a positive attitude towards Czech people	I would like to stay in the Czech Rep. for a long period
completely agree	7	6	3
agree	3	9	6
somewhat agree	4	3	3
difficult to decide	3	-	5
somewhat disagree	1	-	1
disagree	-	-	-
completely agree	-	-	-

Table 4.4. Questions on motivation to learn Czech by number of responses for each item.

As can be seen from the data above, the distribution of answers was unequal for all of the items on the questionnaire. Due to the small number of categories, it was preferred to analyse the results of Question 1 (frequency of hearing Czech), Question 2 (frequency of speaking Czech) and Question 4 (cohabitation with Czechs) with the help of two-tailed t-test. For Question 2, the respondent with the lowest self-assessed frequency was placed in the category for "a couple of times a day". Age and length of residence were analysed using Pearson correlation coefficient (for linear correlation). Kendall rank correlation coefficient was also used to account for non-linear effects. The precise age was unknown

for one respondent (Uk-F-11), who was treated as having age 40 for the purposes of the testing. Again, all F0 parameters measured in Hz (mean, median, baseline, standard deviation) were analysed separately for the two genders.

The self-assessed frequencies of hearing and speaking Czech had a high amount of agreement for individual respondents. However, statistically significant results were only obtained for the frequency of speaking Czech. The observed mean value of F0 baseline was higher for female speakers who spoke Czech several times a day (183 Hz) than for those lower on the frequency scale (153 Hz): $t(9) = 2.82$; $p < 0.05$. The group with a higher frequency of speaking also produced a higher mean value for downtrend gradient (-1 against -1.8 ST/s) and a lower value for CSI (10 against 11.3 ST/s) compared to the other group, the differences being marginally significant (gradient: $t(16) = 2.05$; $p < 0.06$; $t(16) = 0.67$; $p < 0.06$).

A moderate negative correlation was found between the frequency of watching films or videos in Czech and the downtrend gradient ($\tau = -0.41$; $p < 0.05$). The two participants who had Czech cohabitants did not produce significantly different values of F0 measures compared to the other 16 respondents. The values of gradient were also negatively correlated with the perceived importance of the Czech language for work ($\tau = -0.56$; $p < 0.01$). Willingness to stay in the Czech Republic for a long period was not found to correlate with the participants' results for any of the investigated F0 measures.

A number of moderate correlations was observed for the degree of positive attitude towards Czech people. Statistically significant correlation was found for maximum F0 ($\tau = 0.45$; $p < 0.05$), the 90^{th} percentile value of F0 ($\tau = 0.48$; $p < 0.05$) and CSI ($\tau = 0.40$; $p < 0.05$). Marginally significant correlations were found for the variation range ($\tau = 0.32$; $p \approx 0.1$), percentile range ($\tau = 0.37$; $p < 0.07$) and quartile range ($\tau = 0.32$; $p \approx 0.1$). A marginally significant correlation was also found for mean F0 produced by male speakers ($\tau = 0.58$; $p < 0.09$). Mean and median F0 had the same ranking for male speakers, producing the same rank coefficient.

A significant negative correlation was observed between age and median F0 for female speakers: $\tau = -0.49$; $p < 0.05$. Similarly, marginally significant correlations occurred between age and female speakers' mean ($\tau = -0.41$; $p < 0.09$) and baseline F0 ($\tau = -0.45$; $p < 0.06$) values. The lower the speaker's age, the higher values were observed for pitch level, as is

expected for physiological reasons. However, no statistically significant results were found when Pearson correlation coefficient was used.

Finally, length of residence in the Czech Republic was found to negatively correlate with the speaker's values for downtrend gradient, with analogous results for both linear ($\rho = -0.52$; $t(16) = 2.45$; $p < 0.05$) and rank correlation ($\tau = -0.43$; $p < 0.05$). Additionally, statistically significant and marginally significant linear correlations were found between the speaker's lengths of residence and median F0 values (male: $\rho = 0.73$; $t(5) = 2.41$; $p < 0.07$; female: $\rho = 0.62$; $t(9) = 2.34$; $p < 0.05$). Marginally significant correlations were also observed for mean F0 (male: $\rho = 0.70$; $t(5) = 2.20$; $p < 0.08$; female: $\rho = 0.53$; $t(9) = 1.87$; $p < 0.1$). For male speakers, a marginally significant rank correlation was also found ($\tau = 0.55$; $p < 0.10$; mean F0 = median F0).

The results described above should be analysed with caution given the size of the sample. Moreover, false positive results are to be expected when a large number of statistical tests is performed. Due to a small difference between the chosen responses, Questions 1 and 2 were not found to be a reliable measure of the frequency of hearing or speaking the target language. The effects for age were dissimilar across the two statistical tests. The possible reason for this is that the age variation in the sample is insufficient to be a substantial factor in the speakers' language abilities. Out of all the examined sociolinguistic parameters, the most credible results were found for attitude towards the Czech majority population and the length of residence in the Czech Republic.

5. Discussion

This study aimed to provide field data on Ukrainian Czech and compare it to a native Czech reference sample. The study focused on the characteristics of L2 speech melody in the examined population. A number of acoustic correlates of speech melody was selected in order to analyse the difference between the two samples. Apart from the contribution of native language to speech melody, we attempted to investigate other possible factors such as age, gender, exposure and motivation to learn the target language. We were also interested in the effect of the individual speech units and their linguistic contents on the results.

Samples of read-out speech were obtained from 18 native speakers of Ukrainian who were also fluent speakers of the Czech language. The recordings were processed to extract F0 tracks, representing the F0 values during the course of a speech unit. Corrected and interpolated F0 tracks were used to extract long-term frequency measures characterising individual utterances. The 13 utilised F0 parameters included measures of pitch level (mean and median F0, baseline), pitch level boundaries (maximum and minimum F0, the 10^{th} and 90^{th} percentiles of F0 values) and pitch span (variation range, 80-percentile and quartile range), as well as the downtrend gradient. To evaluate the difference between the two language groups, individual values and means of the obtained measures were analysed.

Due to physiological gender differences, the pitch level parameters, measured in hertz, were assessed separately for male and female respondents. Relative to Czech female speakers, Ukrainian female speakers reached significantly lower values for mean and median F0 as well as for F0 baseline. On the other hand, Ukrainian male respondents reached significantly higher values for all three measures compared to the Czech male group. A possible explanation of the disparity between the native Czech and Ukrainian speakers lies in potentially different cultural stereotypes in the two language communities. Previous research has supported the idea that speakers from different cultural groups would employ different pitch levels to conform to a socioculturally desired vocal image (see, for example, Van Bezooijen, 1995). For instance, while some cultures place a stronger emphasis on stereotypical masculinity, associated with a lower pitch, other cultures might display a higher preference for traits associated with higher pitch such as cheer or liveliness.

Czech respondents deviated more from their average pitch level when compared to Ukrainian speakers. This difference was statistically significant for the low pitch boundary, represented by an adjusted pitch level measure (10^{th} percentile of F0). For all groups and most speakers, the deviations observed were larger below the average values than above them. Another finding is that while both Czech groups were in agreement for all four measures, the behaviour of Ukrainian speakers differed by gender. Ukrainian female speakers attained values further from their average than Ukrainian male speakers.

The measures of pitch span agreed with the measures of F0 range limits. Native Ukrainian speakers consistently exhibited lower measures of pitch span when compared to the native Czech group. At the same time, the measures within the Ukrainian group also displayed more variability than those obtained for Czech respondents, with Ukrainian male speakers producing the lowest measures out of all respondent groups. However, the differences in pitch span measures were only significant for male speakers. Similarly, the Ukrainian group displayed significantly higher values of downtrend gradient relative to the native Czechs. Again, when the two gender groups were examined separately, the difference in gradient values was only found to be significant for male speakers. The reason for the different results found for male and female Ukrainian speakers are not entirely clear. A higher L2 proficiency could possibly explain this disparity, as the values in the female group were closer to those attained by the native Czech respondents. A deeper examination of the population is needed to exclude the possibility of sampling bias.

It has been theorised that a narrower pitch span signalises uncertainty or anxiety associated with speaking a foreign language (Volín, Poesová, & Weingartová, 2015). Narrower pitch ranges have been reported in other studies on foreign-accented speech, including for non-native speakers of Czech (Volín, Galeone, & Johnson, 2017). On the other hand, a wide pitch range may be perceived as a sign of insincerity or overacting. Speakers of different languages exhibit variation in fundamental frequency range (Mennen, Schaeffler, & Docherty, 2012), suggesting that L2 pitch span can also be subject to native language interference.

The utterance analysis has revealed a high level of agreement between the two language groups, suggesting that the linguistic content of individual utterances influences the chosen F0 parameters to a high degree. The discrepancies in F0 measures correlated with different

production of speech melody between the two groups. Different melody choices, absence and misplacement of F0 peaks were observed for some Ukrainian speakers.

An additional aim of the study was to investigate possible sociolinguistic correlates of foreign accent, comparing the obtained data on Ukrainian Czech speech melody to a number of sociolinguistic measures. Taking into account the previous research on foreign accent and L2 acquisition, age, length of residence in the target language area, occupation, exposure to the target language and motivation to learn the target language were selected for the sociolinguistic analysis. Exposure and motivation were estimated using a short questionnaire administered to the respondents in the end of the recording procedure (see Section 3.1, p. 19).

Moderate correlations were observed for a number of sociolinguistic factors, most prominently the attitude towards the majority population and the length of residence in the target language area. The participants who evaluated their attitude towards Czech people more positively also produced more native-like values of CSI and the upper range limit, indicating a wider pitch span. The marginally significant correlations for measures of pitch span and central tendency (for male speakers) pointed in the same direction. A greater length of residence in the Czech Republic correlated with steeper downtrend gradient values, closer to those attained by the native Czech respondents. However, Ukrainian female participants residing in the Czech Republic for a longer time also reached higher values for median F0, increasing their distance from the Czech female group. While the length of residence is a well-known factor in second-language acquisition, the factors related to social identity, such as attitude towards the majority population, warrant special attention. It is conceivable that a speaker's attitude towards the majority population is closely related to the degree of integration in the target community and other social characteristics. Similar findings were reported in the study of non-native speakers of English in the U.S. by Gluszek, Newheiser and Dovidio (2011; see Section 2.2, p. 13), who identified a relationship between accent strength and the identification with American culture.

The size and sociolinguistic characteristics of our sample limit the possibility to generalise the findings of this study to the whole population of Ukrainian speakers. However, this study aimed to investigate a subset of that population, specifically Ukrainian speakers living in the Czech Republic. The acoustic variables employed do not equal the subjective percept

of intonation (p. 9). Nevertheless, the examination of acoustic correlates of perceived pitch and pitch movements is the current established approach of studying speech melody.

The results of our study suggest several possible directions for future research. First, it would be beneficial to obtain and analyse samples of other speech types. Recording spontaneous speech would eliminate the influence of the participants' reading skills on the material. A different experiment design would also provide an opportunity to test the effect of a less controlled environment on the non-native respondents' language behaviour. Another direction for improvement includes collecting a broader sample of speakers to validate the findings of this study. The relationship between linguistic content of an utterance and its pitch properties also warrants further investigation. Finally, the exploration of sociolinguistic factors provided here could be followed by a more detailed examination of the attitudes of Ukrainians living in the Czech Republic and their relationship to the acquisition of the Czech language.

6. References

Boersma, P., & Weenink, D. (2017). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.28, retrieved 16 May 2017 from http://www.praat.org/

Braun, B., Dainora, A., & Ernestus, M. (2011). An unfamiliar intonation contour slows down online speech comprehension. *Language and Cognitive Processes, 26*(3), 350-375.

Chun, D. M. (2002). *Discourse Intonation in L2: From Theory and Research to Practice*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Czech Statistical Office (2017). *Cizinci v ČR podle státního občanství v letech 1994 – 2016 (stav k 31. 12.)*. [Data file]. Available from https://www.czso.cz/csu/cizinci/4-ciz_pocet_cizincu

Danylenko, L. (2007). *Čes'ka mova*. Kyjiv: Dovira.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476-490.

Flege, J. E. (1998). The Role of Subject and Phonetic Variables in Second-Language Learning. In: M. C. Gruber, D. Higgins, K. S. Olson & T. Wysocki (Eds.), *Papers from the 34th Annual Meeting of the Chicago Linguistic Society, Volume II, The Panels* (pp. 213-232). Chicago: Chicago Linguistic Society.

Flege, J. E. (1999). Age of learning and second-language speech. In: D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* (pp. 101-132). Hillsdale, NJ: Lawrence Erlbaum.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169-186.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*, 120-133.

Gluszek, A., Newheiser, A-K., & Dovidio, J. F. (2011). Social psychological orientations and accent strength. *Journal of Language and Social Psychology, 30*(1). 28-45.

Grabe, E. (1998). *Comparative Intonational Phonology: English and German.* Wageningen: Ponsen en Looien.

Grice, M., & Bauman, S. (2007). An introduction to intonation – functions and models. In: J. Trouvain & U. Gut (Eds.), *Non-Native Prosody* (pp. 25-51). Berlin: Mouton de Gruyter.

Gut, U., & Pillai, S. (2015). The Question Intonation of Malay Speakers of English. In: E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Language Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*, 51-70. Berlin: Springer.

Hermes, D. J. (2006). Stylization of pitch contours. In: S. Sudhoff et al. (Eds.), *Methods in empirical prosody research* (pp. 29-62). Berlin: Walter de Gruyter.

Holm, S. (2007). The relative contributions of intonation and duration to intelligibility in Norwegian as a second language. In: J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th ICPhS* (pp. 1653–1656). Saarbrücken: IPA.

Hosoda, M., Nguyen, L. T., & Stone-Romero, E. F. (2012). The effect of Hispanic accents on employment decisions. *Journal of Managerial Psychology, 27*(4), 347-364.

Hruška, R. (2016). Vlastnosti kontur základní frekvence s ohledem na segmentální složení promluv (Unpublished bachelor's thesis). Charles University, Prague.

Hruška, R., & Bořil, T. (2017). Temporal variability of fundamental frequency contours. *AUC Philologica 3/2017*, 35-44.

Kamiyama, T. (2004). Perception of foreign accentedness in L2 prosody and segments: L1 Japanese speakers learning L2 French. In: Bel, B., & Marlien, I. (Eds.), *Proceedings of Speech Prosody 2004* (pp. 721-724). Nara, Japan: ISCA.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Leontiyeva, Y. (2010). *Ukrajinci v ČR*. In: T. Hirt (Ed.), *Vietnamci, Mongolové a Ukrajinci v ČR: Pracovní migrace, životní podmínky, kulturní specifika* (pp. 59-80). Praha: Ministerstvo vnitra.

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*, 1093-1096.

Lieberman, P. (1967). *Intonation, perception and language*. Cambridge, Massachusetts: MIT Press.

Lindh, J., & Eriksson, A. (2007). Robustness of long time measures of fundamental frequency. In: *Proceedings of Interspeech 2007* (pp. 2025-2028). Antwerp: ISCA.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32, 543-563.

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In: J. Trouvain & U. Gut (Eds.), *Non-Native Prosody* (pp. 53-76). Berlin: Mouton de Gruyter.

Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of America*, 131(3), 2249-2260.

Moyer, A. (2007). Do Language Attitudes Determine Accent? A Study of Bilinguals in the USA. *Journal of Multilingual and Multicultural Development, 28*(6), 502-17.

Munro, M. J., & Derwing, T. M. (1995). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning 45*(1), 73-97.

Ng, S. H. (2007). Language-Based Discrimination: Blatant and Subtle Forms. *Journal of Language and Social Psychology, 26*(2), 106-122.

Osměra, R. (2009). Podoba zkoušky z českého jazyka pro účely získání trvalého pobytu. *MigraceOnline.cz*. Retrieved from http://migraceonline.cz/cz/e-knihovna/podoba-zkousky-z-ceskeho-jazyka-pro-ucely-ziskani-trvaleho-pobytu

Palková, Z. (1994). *Fonetika a fonologie češtiny: s obecným úvodem do problematiky oboru*. Praha: Karolinum.

Palková, Z. (1997). Modelling intonation in Czech: neutral vs. marked TTS F0-patterns. In: A. Botinis (Ed.), *Intonation: Theory, Models, and Applications* (pp. 267-270). Athens, Greece.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics, 29*, 191-215.

R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ravuri, S., & Ellis, D. P. W. (2008). Stylization of pitch with syllable-based linear segments. In: *2008 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 3985-3988). Las Vegas: IEEE Signal Processing Society.

Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech, 20*, 267-273.

Santiago, F., & Delais-Roussarie, E. (2015). The Acquisition of Question Intonation by Mexican Spanish Learners of French. In: E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Language Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*, 243-270. Berlin: Springer.

Trim, J. L. M. (1988). Some contrastive intonated features of British English and German. In: J. Klegraf & D. Nehls (Eds.), *Essays on the English language and applied linguistics on the occasion of Gerhard Nickel's 60th Birthday*, 235-49. Heidelberg: Julius Groos.

Trofimovich, P., & Baker, W. (2006). Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquisition, 28*, 1-30.

Trouvain, J., & Gut, U. (2007). *Non-Native Prosody*. Berlin: Mouton de Gruyter.

Van Avermaet, P. (2009). Fortress Europe? Language policy regimes for immigration and citizenship. In: G. Hogan-Brun, C. Mar-Molinero & P. Stevenson (Eds.), *Discourses on Language and Integration: Critical perspectives on language testing regimes in Europe* (pp. 15-44). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and speech*, 38(3), 253-265.

Vieru-Dimulescu, B., & de Mareüil, P. (2005). Contribution of prosody to the perception of a foreign accent: a study based on Spanish/Italian modified speech. In: *Proceedings of the ISCA Workshop on Plasticity in Speech Perception* (pp. 66-69). London.

Volín, J., Galeone, D., & Johnson, W. (2017). Pitch range of intonation contours in English Czech. *AUC Philologica 3/2017*, 55-64.

Volín, J., Poesová, K., & Skarnitzl, R. (2014). The impact of rhythmic distortions in speech on personality assessment. *Research in Language*, *12*, 209-216.

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.

Wichmann, A. (2000). *Intonation in text and discourse* (pp. 102-122). Harlow, UK: Longman.

Appendix A: List of participants

The table below contains the list of participants recorded for this study along with their sociolinguistic characteristics at the time of recording. The 'Participant' column gives the coded IDs for each participant, specified for gender ('F' for female and 'M' for male). The next two columns, 'Age' and 'LOR', provide the participants' age and length of residence in the Czech Republic in years, approximated to the nearest whole number. The 'Occupation' column categorises participants as "high school students", "[Bachelor's or Master's] students", "doctoral students" and "employed". The "students" category includes the two respondents who have graduated in the month before the recording. The "employed" category only includes the respondents in full-time employment.

Participant	**Age** (years)	**LOR** (years)	**Occupation**	**Residence**
Uk-F-1	21	3	student	Brno
Uk-F-2	23	4	employed	Prague
Uk-F-3	23	3	doctoral student	Prague
Uk-F-4	26	3	employed	Prague
Uk-F-5	24	1	doctoral student	Prague
Uk-F-6	24	5	student	Brno
Uk-F-7	25	5	doctoral student	Prague
Uk-F-8	19	7	high school student	Prague
Uk-F-9	27	3	employed	Prague
Uk-F-10	26	3	employed	Prague
Uk-F-11	40+	7	employed	Prague
Uk-M-1	23	2	student	Prague
Uk-M-2	24	1	student	Prague
Uk-M-3	27	3	doctoral student	Prague
Uk-M-4	21	4	student	Prague
Uk-M-5	21	4	student	Prague
Uk-M-6	23	2	doctoral student	Prague
Uk-M-7	22	5	student	Brno

Appendix B: Reading task („Milánek“)

Maminka se zeptala Milana:
„Milánku, už máš hotový úkol? Kdy ho budeš psát?“

Milan chvilku přemýšlel, a pak odpověděl:
„Já musím napsat pár souvětí na Říhovou, kde budou nějaké gramatické fígle. Například vztažné věty, čárky před ‚a‘ a podobně. Vlastně jsem tě chtěl poprosit, jestli mi s tím nepomůžeš.“

„Můžeme se na to mrknout klidně hned,“ řekla maminka. „Jen bych dala vařit vodu na čaj a podívám se, jestli máme citróny. Jak dlouho nám ten úkol zabere? Bude to těžké?“
„No, mají tam být i různé příklady na zastaralou a knižní slovní zásobu. Skoro půlku jsem už ve škole udělal, ale moc dobře mi to nešlo. Chtěl bych začít co nejdřív. Až budu hotov, došel bych ti do lékárny pro ten *PNEUMOCYT*.

A potom půjdu hrát fotbal. Včera jsem dal čtyři góly.
Nebýt Láďových faulů, mohlo jich být víc. Jen míč budu muset přifouknout. Neboj se, dám pozor na auta.“

Appendix C: Line plots

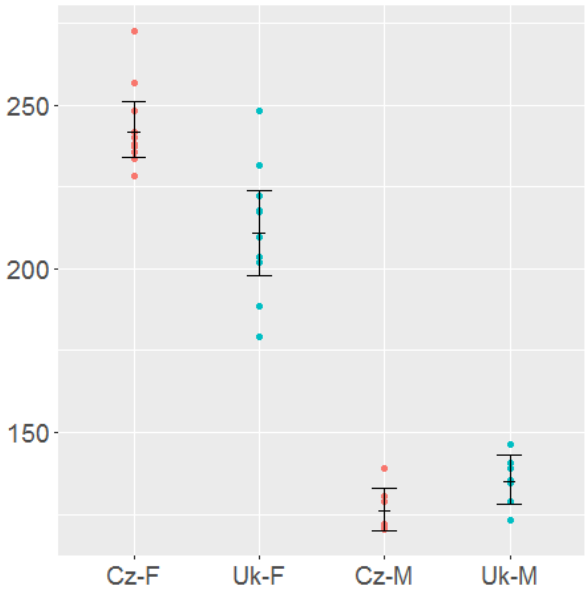

Figure 7.1. Mean F0 (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

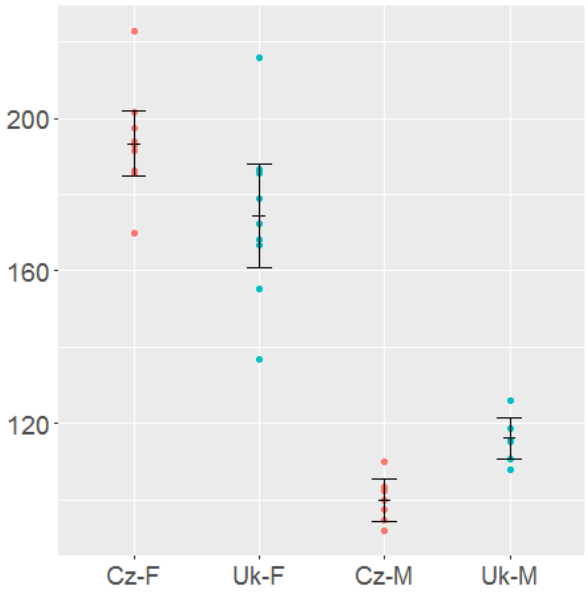

Figure 7.2. F0 baseline (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

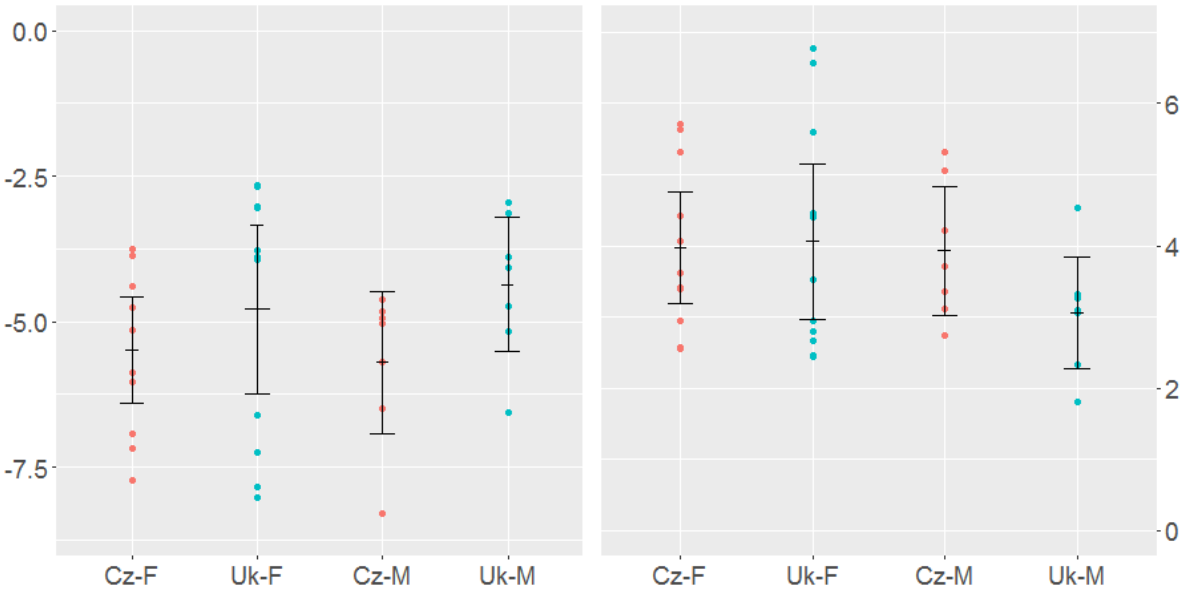

Figure 7.3. Minimal (left) and maximal (right) F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

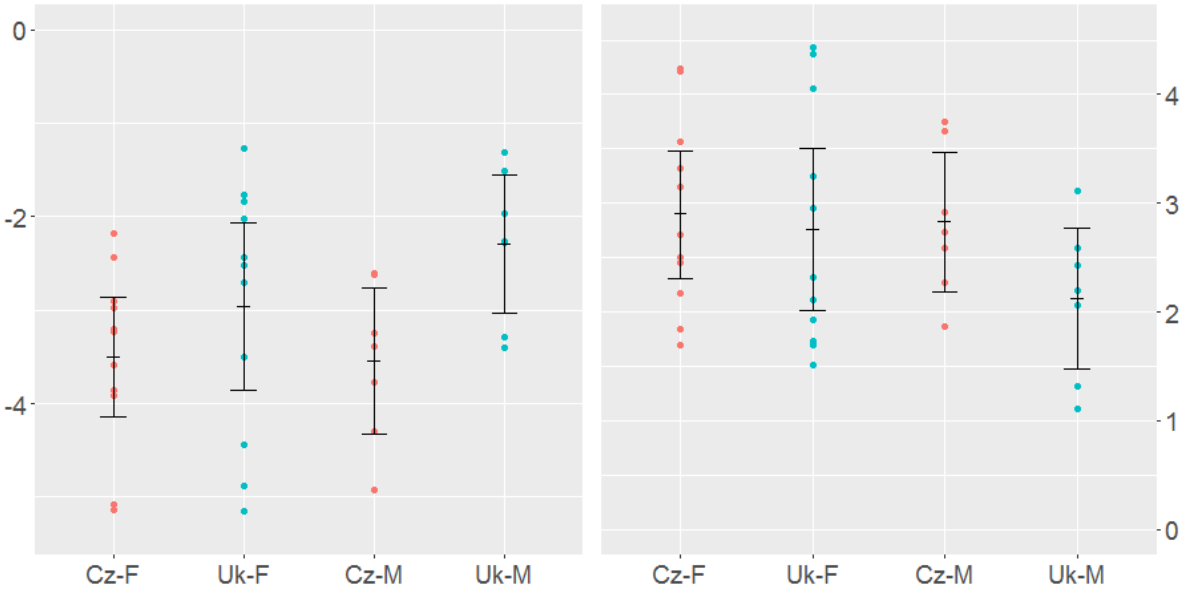

Figure 7.4. Average 10th (left) and 90th (right) percentile values of F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

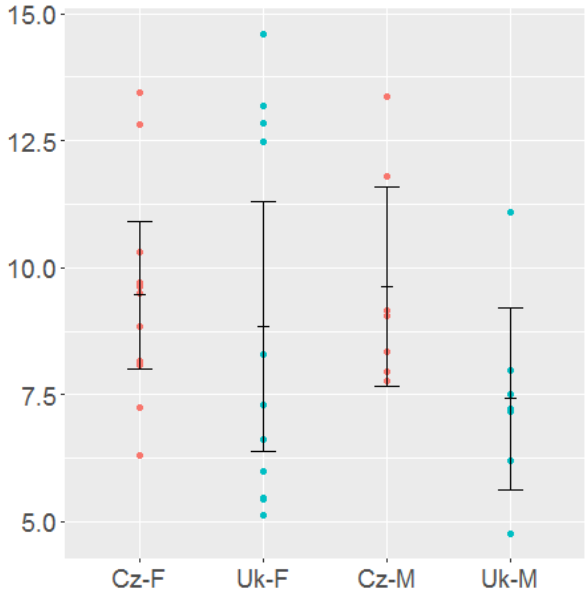

Figure 7.5. Average variation range of F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

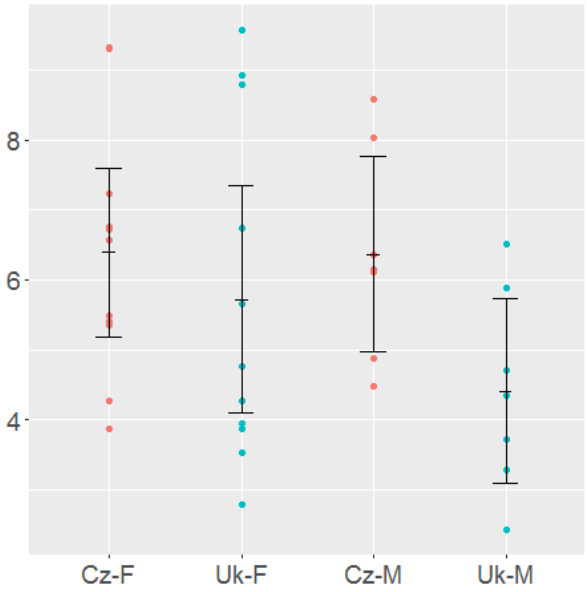

Figure 7.6. Mean values of the 80-percentile range of F0 (in semitones) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

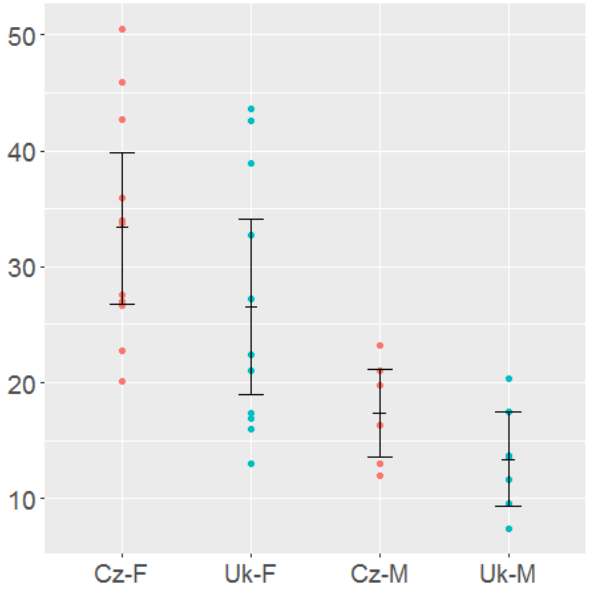

Figure 7.7. Standard deviation of F0 (in Hz) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

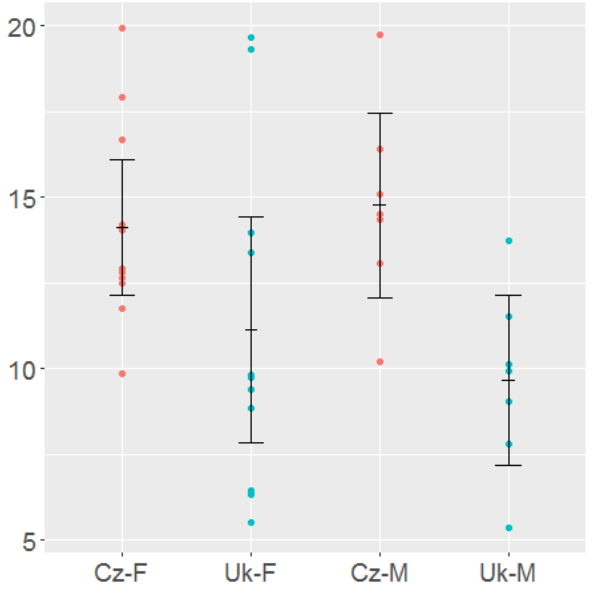

Figure 7.8. Cumulative slope index (in semitones per second) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.

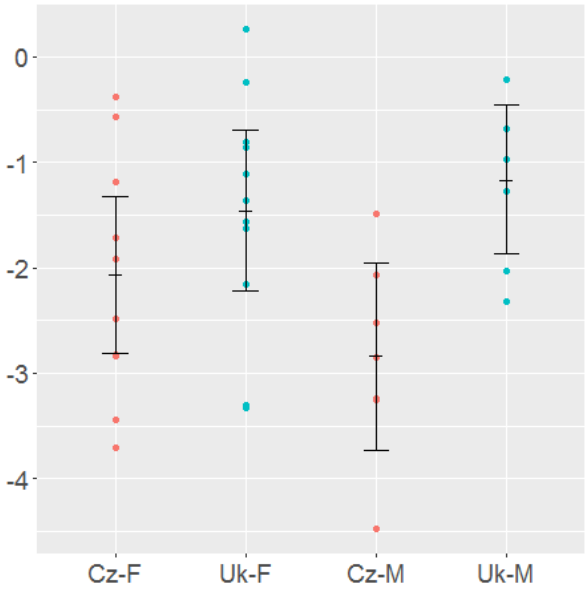

Figure 7.9. Average values of downtrend gradient (in semitones per second) for two genders (M – male, F – female) and languages (Cz – Czech, Uk – Ukrainian). Whiskers indicate 95% confidence intervals.