

Filosofická fakulta Univerzity Karlovy

Fonetický ústav

Michaela Šedivá

**Melodická variabilita řeči v závislosti na
afektivním ladění politické debaty**

Melodic variation and affective valence in political debates

Bakalářská práce

Vedoucí práce: Mgr. Pavel Šturm, Ph.D.

2020

Poděkování

Ráda bych podělovala doktoru Pavlu Šturmovi za trpělivost a ochotu, kterou po celou dobu psaní této práce prokazoval. Dále děkuji docentu Janu Volínovi za poskytnutí skriptu pro měření intonačního rozpětí a také doktoru Tomáši Bořilovi za laskavou konzultaci ohledně statistického testování. V neposlední řadě děkuji doktorkám Haně Kratochvílové a Anje Bunzel za podporu a Tomáši Svobodovi, že mi byl vždy oporou.

Prohlášení

Prohlašuji, že jsem bakalářskou práci na téma *Melodická variabilita řeči v závislosti na afektivní valenci politické debaty* vypracovala pod vedením vedoucího bakalářské práce samostatně za použití v práci uvedených pramenů a literatury. Dále prohlašuji, že tato bakalářská práce nebyla využita k získání jiného nebo stejného titulu.

V Praze dne 24.5.2020

Podpis

Abstrakt

Tato práce zkoumá vliv afektivní valence na intonační variabilitu spontánní češtiny tří účastníků politické debaty. Pro tento výzkumný úkol si stanovila tři parametry intonační variability, na jejichž vztah s afektivní valencí (jakožto jednou z dimenzí afektivního dimenzionálního modelu) se zaměřuje. Těmito parametry jsou intonační rozpětí, počet melodémů v nádechovém úseku a ukazatel kumulativní strmosti (který představuje nový parametr pro zkoumání celkové proměnlivosti intonační kontury). Kromě variability byl prozkoumána též hlasová poloha mluvčích (pomocí mediánu f_0). Práce v teoretické části seznamuje se svými psychologickými i fonetickými východisky a zmiňuje předchozí zjištění v daném oboru, na jejichž základě stanovuje své hypotézy.

Pro samotný výzkum byla použita dvacetiminutová spontánní nahrávka politické debaty mezi dvěma politiky a moderátorem. Specifika vybrané komunikační situace jsou v teoretické části rovněž upřesněna. Po manuálním zpracování materiálu a určení afektivní valence pro každý nádechový úsek byla tato nahrávka foneticky analyzována a výsledky prezentovány pro každého mluvčího zvlášť.

Výsledky práce přinesly několik dílčích zjištění, která nasvědčují tomu, že afektivní valence může intonační variabilitu ovlivňovat, a to především v ohledu na intonačního rozpětí, ale i celkovou proměnlivost kontury. Vliv afektivní valence na počet melodémů v nádechovém úseku se v této práci prokázat nepodařilo. Závěrem práce tudíž je, že má smysl podrobit téma důkladnějšímu zkoumání za použití většího množství dat. Právě jejich nedostatek byl v představeném výzkumu jednou z překážek proti významnějším závěrům. Práce také na základě svých zkušeností navrhuje některé změny oproti vlastní metodice pro další podobné výzkumy a zároveň přichází na základě vlastních výsledků s hypotézou, že existují různé strategie užívání intonačních změn při vyjadřování afektivních stavů.

Klíčová slova

intonace, melodie, intonační rozpětí, afektivní stavy, afektivní valence

Abstract

This thesis investigates the impact of affective valence on intonation variability within the context of spontaneous speech acts in the Czech language. It is based on an examination of speech patterns by three discussants in a political debate. For this purpose, we have chosen three analytical parameters: intonational span, number of melodemes in one breath group, and cumulative slope index (CSI), the last of which is a new parameter for the description of the overall variability of intonational contours. In addition, another parameter described changes of the median of f_0 .

The theoretical framework of this thesis draws on concepts stemming from phonetics and psychology. The research sample consists of a twenty-minute recording of a political TV debate between two politicians, chaired by a moderator. The specifics of this sample are discussed in the theoretical part of this thesis. In our own investigation, we first examined the material manually with regard to the affective valence of each breath group, followed by a phonetic analysis and a statistical analysis.

The execution and results of this study enable the following conclusions: our analysis has shown that there may be a correlation between affective valence and intonational variability, especially with respect to the intonational span and the overall variability presented by the CSI. We were unable to trace the impact of the affective valence on the number of melodemes per breath group. However, the lack of sufficient data has been a significant obstacle over the course of our examination. Furthermore, there may be additional strategies of using intonation changes when expressing affective states, which have not been considered at depth here. Therefore, we recommend that future studies on similar topics should implement larger data sets and some methodical changes.

Keywords

Intonation, melody, intonational span, affective states, affective valence

Obsah

Obsah	6
1. Úvod.....	8
2. Prozodický kontext práce.....	10
2.1. Intonace jako dimenze prozodie.....	11
2.1.1. Percepce intonace a její fyziologická podstata	12
2.1.2. Psychoakustický problém intonace	13
2.1.3. Intonační variabilita	14
2.2. Melodém a kadence	15
2.3. Lineární členění promluvy	16
3. Afektivní stavy a afektivní valence.....	19
3.1. Problematika termínu afektivní stavu.....	19
3.2. Schererovo pojetí afektivních stavů jako východisko pro práci.....	20
3.3. Dosavadní výzkumy projevů afektivních stavů v řeči	21
3.4. Afektivní valence v dimenzionálním modelu.....	24
3.4.1. Problematika neutrality afektivní valence při řeči	26
3.5. Specifika prostředí politické debaty při vyjadřování afektivních stavů	28
4. Hypotézy a cíle práce	30
5. Metoda	32
5.1. Materiál	32
5.1.1. Charakteristika použitého materiálu.....	32
5.1.2. Příprava materiálu pro analýzu.....	32
5.2. Problematika subjektivity při určování afektivní valence	34
5.3. Zvolené parametry intonační variability a jejich analýza	36
5.3.1. Hlasová poloha mluvčích	36
5.3.2. Intonační rozpětí.....	36
5.3.3. Parametr kumulativní strmosti.....	38
5.3.4. Analýza počtu melodémů v nádechovém úseku	39

5.4. Statistické testování výsledků	39
6. Výsledky.....	40
6.1. Ukazatel hlasové polohy	40
6.2. Percentilové intonační rozpětí.....	41
6.3. Parametr kumulativní strmosti.....	43
6.4. Počet melodémů v nádechovém úseku.....	44
6.5. Korelace mezi intonačním rozpětím a CSI	45
7. Diskuse.....	46
Seznam zdrojů.....	50

1. Úvod

Dosud provedené studie ohledně změn v oblasti prozodických vlastností řeči v důsledku projevu emocí se shodují na existenci vlivů psychických stavů na fyziologické změny, které ústí v percepčně vnímatelné a při komunikaci mluvčími rozeznatelné posuny. Pro taková zjištění však nemusíme chodit do akademických vod – změny na této rovině jsme totiž schopni vnímat zcela bezděčně a kromě toho je umíme i interpretovat. Vzpomeňme při této příležitosti na film Miloše Formana *Černý Petr* (1963) a na tomto faktu vystavěnou (a do dneška notoricky známou) scénu, kde si Vladimír Pucholt (coby Čenda) všímá intonačního rozdílu při pronesení pozdravu „ahoj“ v situacích s různým afektivním zabarvením. Že rozeznávání emočního ladění mluvčího skrze prozodii není nadpřirozenou schopností, kterou by si scenárista vymyslel, ale běžnou lidskou způsobilostí, jež je mluvčím vlastní a v komunikaci nápomocná, dokazují ve svých studiích psychologové již dlouhá desetiletí.

Jako na fyziologicky projevenou psychologickou reakci se na emoci poprvé začal dívat koncem devatenáctého století americký psycholog William James. Od té doby jde o téma, které zajímá vědce nejen na poli psychologie, ale má rovněž přesah do přírodních věd. Existuje množství výzkumů (především neuropsychologických), které se v různé míře zabývají souvislostí emocí a jejich odrazem v mluveném projevu.

Je však na místě podotknout, že na poli zkoumání emočních projevů při řeči bylo dosud věnováno více pozornosti projevům neverbálním (například mimika), jak uvádí Laukka, Juslin & Bresin (2005), a těžiště zájmu se směrem ke zvukovým projevům začalo přesouvat až na přelomu nového tisíciletí. Tento fakt pravděpodobně souvisí i s příchodem techniky pro analýzu, kterou máme v současné době pro fonetický výzkum k dispozici, a můžeme tak laické předpoklady potvrdit exaktními metodami.

Je jisté, že vybrané téma představuje komplexní problematiku, která mnohanásobně přesahuje možnosti bakalářské práce. Tato práce si vytyčila za cíl přispět alespoň malou částí, zaměřenou pouze na jeden aspekt afektivních stavů – afektivní valenci, a to v souvislosti s jedinou dimenzí prozodického popisu řeči – intonací – a zejména pak její variabilitou. Tu budeme v práci reprezentovat třemi parametry, které jsme pro její zachycení předem stanovili. Jedná se o **intonační rozpětí, počet melodémů v nádechovém úseku a ukazatel kumulativní strmosti**. Cíl práce je čistě deskriptivní – chceme zjistit a popsat, jak (a zda vůbec) afektivní valence a její míra intonační variabilitu spontánního projevu ovlivňuje.

Práce je rozdělena na teoretický a praktický oddíl. V úvodní teoretické kapitole se zaměřujeme na prozodický kontext práce – vymezujeme související termíny, jako je intonace, melodém či nádechový úsek, a věnujeme pozornost těm aspektům, které jsou pro orientaci v práci klíčové.

V následující kapitole stavíme výzkum do kontextu psychologického. Osvětlujeme termín afektivní vztahy, předkládáme jejich dělení a seznamujeme s dimenzionálním modelem, který definuje afektivní valenci, jejíž vztah s intonační variabilitou budeme examínovat v druhé části práce. Dále interpretujeme výsledky některých výzkumů z této oblasti a uvažujeme o nich jako o možných východiscích pro formulaci hypotéz. V závěru kapitoly krátce referujeme o specifikách, která s sebou přináší prostředí politické debaty, jež představuje zkoumaný materiál v praktické části této práce.

Teoretická část je s částí praktickou propojena v krátké kapitole č. 4, v níž formulujeme hypotézy. Následuje metodologická kapitola, ve které seznamujeme s použitým materiálem a jeho úpravou ve fonetickém softwaru Praat. Následně více do hloubky představujeme všechny parametry, které jsme pro výpočet intonační variability zvolili, a popisujeme statistické metody zpracování získaných hodnot. Nakonec zobrazujeme výsledky, ke kterým jsme došli, interpretujeme je a uzavíráme celou práci diskusí nad nimi.

Za cíl si klademe přispět novým zjištěním ke zkoumání afektivně-prozodických charakteristik spontánní řeči, jelikož se jedná o oblast s velkým potenciálem vzhledem k technologickému pokroku na poli řečového inženýrství. Jedním z dílčích cílů je rovněž ověřit možnosti použití nahrávky politické debaty pro podobné výzkumy a po zkušenosti s prací s tímto materiálem referovat o jejích silných stránkách a případných omezeních.

2. Prozodický kontext práce

Než začneme hovořit o intonačním modulování řeči v závislosti na psychickém rozpoložení člověka, které je předmětem našeho zkoumání, musíme toto téma nutně zasadit do širšího teoretického kontextu. K tomu nám pomůže představení konceptu prozodie, jejíž je intonace nedílnou součástí.

Termín prozodie (z řeckého *prosódia*) má své kořeny již ve starověku. Původně označoval řeckou píseň s instrumentálním doprovodem, později byl spojován s versologií a uměním recitace. Dnešní fonetika chápe prozodii jako souhrn jevů v rámci suprasegmentální roviny řeči – jakákoli jednotka, která je tedy vyšší než samostatná hláska, bude touto rovinou ovlivněna. Ačkoliv byla prozodie tématem fonetiky dříve zanedbávaným (přednost dostávaly spíše jevy na úrovni segmentální), od poloviny 60. let 20. století jde o oblast atraktivní a vyhledávanou.

Prozodii je řečová komunikace prostoupěna ve velké míře. Hraje významnou roli v porozumění a má schopnost měnit významy řečových celků, jelikož je vysoce konvencionalizovaná, jazykově vázaná a zřetelně systematická. Pro příklad stačí uvést české věty oznamovací a tázací zjišťovací (otázka očekávající odpověď ano/ne), které od sebe bez přítomnosti kontextu není možné rozeznat, pokud mluvčí suprasegmentálních prvků (zde konkrétně intonace) nevyužije ve shodě s konvencemi.

Prozodie bývá popisována jako multidimenzionální jev, který má nejméně tři (nerozporuplné) základní dimenze. Tyto dimenze můžeme popsat jako produkční koreláty („to, co může mluvčí ovlivňovat“), akustické koreláty („to, co můžeme u takové změny měřit“) a percepční koreláty („způsob, jakým se změna projevuje při vnímání řeči“). K těmto korelátům můžeme přidat ještě informaci o tom, jak se reálně v řeči projevují.

Pro přehlednost tyto dimenze vynášíme do tabulky (Tab. 1), jelikož následující termíny bývají v některých popisech prozodie nesprávně kombinovány a zaměňovány. Tuto nepotěšující skutečnost glosují Jan Volín a kolegové slovy, že „*patří k malé fonetické násobilce držet tyto roviny oddělené, neboť vztahy mezi nimi jsou značně komplikované*“. (Volín et al., 2007, s. 261)

produkční korelát	akustický korelát	percepční korelát	Jazykové využití
rychlost kmitání hlasivek	frekvence	výška	melodie/intonace
doba držení traktu v urč. pozici	trvání	délka	tempo
různé vzájemně závislé artikulační a akustické jevy	intenzita	hlasitost (síla)	důraz

Tab. 1: Koreláty dimenzí prozodie.

Mezi základní dimenze prozodie zařazují někteří autoři i kvalitu hlasu a tento přístup vítá i česká fonetika. Metodiku její měřitelnosti přinášejí Campbell a Mokhtari v článku s příznačným názvem *Voice Quality: the 4th Prosodic Dimension* (Campbell & Mokhtari, 2003). K tabulce základních dimenzí bychom tak mohli přidat ještě jeden řádek, ve kterém by za produkční změnu bylo označeno nastavení vokálního traktu - jako akustický korelát bychom mohli označit spektrální vlastnosti. Percepčním jevem by potom byla právě kvalita hlasu. Na principu kombinace těchto dimenzí fungují jevy, jako jsou například přízvuk či rytmus.

Prozodie je ve své podstatě jevem multifunkčním. K jejím základním funkcím patří funkce lexikální (se kterou se setkáváme například u tónových jazyků, kde má melodická realizace schopnost měnit významy slov), funkce gramatická (například již zmíněný rozdíl mezi otázkami zjišťovacími a větami oznamovacími; dále realizace větného členění; subordinace), funkce diskurzivní (řízení rozhovoru, turn-taking), funkce sociofonetická (například specifické zacházení s prozodií napříč republikou – třeba plzeňské „zpívání“ či specifická realizace otázek obyvateli Hané) a nakonec funkce afektivní (zrcadlí se v ní náš aktuální psychický stav), která bude pro tuto práci stěžejní.

Afektivní funkce je velice významnou složkou komunikace, neboť samotný komunikovaný obsah neumožňuje v úplnosti porozumět komunikační situaci. K tomu je často potřeba právě informace o rozpoložení mluvčího, při jejíž absenci může docházet k neúplnému pochopení sdělovaných informací.

2.1. Intonace jako dimenze prozodie

Pro pozdější orientaci v tématu je potřeba provést malý exkurz do intonologického světa a rovněž do jeho terminologie – mladost a složitost disciplíny s sebou totiž nese jistá terminologická úskalí. Neklademe si za cíl tyto problémy vyřešit, ale považujeme za důležité osvětlit, jak se k následujícím pojmům stavíme, a také tato stanoviska podložit.

Skarnitzl, Šturm & Volín (2016) hovoří o melodii jako o tónových průbězích tvořících rozpoznatelné celky v hudbě, autoři nicméně připouštějí i přesah tohoto termínu do terminologie

lingvistické, která však podle nich zavádí pro své potřeby kromě melodie další termín – intonace. Ten je tradičně možné interpretovat dvěma způsoby:

- a) intonace „v širším slova smyslu“
- b) intonace „v užším slova smyslu“

Pro širší záběr tohoto termínu autoři doporučují užívat právě termínu *prozodie*, který zahrnuje i další suprasegmentální jevy, které moduluji řetězec slabik a mají za úkol tvořit celky, pomoci odhalit jejich význam a provést posluchače po promluvě (viz kap 2. této práce). Snadná záměna těchto dvou termínů může být způsobena tím, že právě intonace je vnímána jako nejsilnější složka prozodie. Intonaci v této práci však budeme vnímat „v užším slova smyslu“.

Palková (1994, s. 161) v rozsáhlé kapitole o suprasegmentální rovině řečové komunikace uvádí celou podkapitulu, která se zabývá terminologickou problematikou ohledně melodie a intonace. Její pohled je následující:

Intonace je chápána jako komplexní zvuková kvalita (zahrnující především tónovou a silovou modulaci), termínem melodie se označuje výhradně průběh výšky základního tónu. V tom smyslu tvoří tedy melodie (základní) složku intonace. Místo termínu melodie užívají někteří autoři ještě konkrétnějšího označení „průběh základního tónu (f_0)“.

Zde je však nutné podotknout, že v citovaném textu autorka vůbec neuvádí výrazu *prozodie*. V širším kontextu je tedy nasnadě se domnívat, že její pojetí intonace je v citované práci totožné s pojetím intonace „v širším slova smyslu“. Některé novější práce (např. již zmiňovaní Skarnitzl et al., 2016) preferují označovat tento celistvý koncept právě za prozodii, jež je ve své podstatě multidimenzionálním jevem a intonace „v užším slova smyslu“ je jednou z těchto dimenzí, jak již bylo nastíněno v předchozím oddílu této práce.

Melodií je pak v pojetí Palkové chápána intonace „v užším slova smyslu“. Co je však důležité pro její definici našeho pojetí, je v citovaném úseku zmíněný termín **průběh základního tónu (f_0)**, který vystihuje použití termínu intonace v této práci a jde podle našeho názoru o nejuniverzálnější a ke zmatení nejméně náchylný termín, ačkoli se zde proti použití termínu „melodie řeči“ v témže významu nevymezujeme.

2.1.1. Percepce intonace a její fyziologická podstata

Bylo by chybou se domnívat, že pro intonaci je nosný úplně každý segment v řečovém proudu. Intonaci jsme definovali jako průběh základního tónu (f_0), přesněji řečeno se však jedná o jeho percepční korelát. Aby percept intonace mohl vzniknout, je nutné zapojit činnost hlasivek, které

svým kmitáním tónovost udávají (a různou frekvencí kmitání pak různou kvalitou těchto tónů – tomu odpovídá termín „výška“). Ačkoli je výsledný percept člověkem vnímán jako spojitý, výšku určující segmenty jsou pouze ty, které označujeme jako znělé, jelikož právě při nich dochází ke kmitání hlasivek, jež tón vytváří. Pozoruhodnou schopností lidského zpracování těchto ne vždy spojitých tónů je interpolace (překlenutí) neznělých „děr“ v proudu řeči tak, že vjem ve výsledku nabude spojitosti, ačkoli z čistě fyzikálního pohledu se o spojitou křivku nejedná.

Při této příležitosti je ještě záhodno zmínit termín *mikrointonace*. Dosud jsme o intonaci hovořili jako o něčem, co „je vnímáno“ (viz Tab. 1 s koreláty). Vnímáno ovšem může být pouze to, co leží nad hranicí vnímatelnosti. Mikrointonace je však jevem, který je lidským mozkem postaven pod tuto hranici – jde o mluvčím nezamýšlené výkyvy F_0 , které jsou způsobeny fyziologickými možnostmi mluvního ústrojí a nejsou tudíž žádným způsobem významotvorné. Jejich vznik je tedy čistě „mechanickou“ a „provozní“ záležitostí, která nemá při řeči relevanci, a může tak být lidským percepčním aparátem zanedbána. Tento jev zmiňujeme z důvodu, že při programové extrakci kontury f_0 pro intonologické výzkumy se přistupuje k tzv. stylizaci kontury, jak blíže uvedeme v metodické kapitole. Tento proces má za cíl programem přesně naměřená data (která tedy zahrnují i mikrotonační jevy) zjednodušit a těmito úpravami vlastně modelovat její skutečné lidské vnímání. Informace o něm jsou považovány za cennější než počítačem naměřená exaktní data, která bez zasazení do kontextu lidské percepce zůstanou pouhými čísly. Za tímto směrem orientovaný přístup vděčí světová fonetika především holandské škole IPO (*Instituut voor Perceptie Onderzoek* – institut pro výzkum percepce, Eindhoven), která již v devadesátých letech minulého století publikovala výsledky svého výzkumu (Hart, Collier & Cohen, 1990), jenž experimentálně doložil, že lidská percepce intonace je oproti naměřeným datům opravdu zjednodušená a ne všechny prvky jsou pro ni relevantní. Intonační modelování by tak podle této školy mělo být orientováno na to, co dokážeme rozlišit, nikoli na to, co dokáže naměřit počítač.

2.1.2. Psychoakustický problém intonace

Skrze stylizaci kontury, ke které nás nutí rozdíl mezi měřenými a vnímanými hodnotami, se dostáváme k další zásadní problematice, jež je sice spíše charakteru metodického, nicméně i na ní závisí celistvost informací o intonaci, které je nutné zmínit již v teoretické části práce. Tímto problémem je fakt, že lidský sluch nevnímá intonační změny lineárně. Ukazuje se, že objektivně stejně velké frekvenční rozdíly v různých částech pásma slyšitelnosti jsou vnímány jako různé velké (Skarnitzl et al., 2016). Roli v tomto subjektivním vnímání hraje výška – určitý rozdíl bude vnímán jako markantnější v nízkých frekvencích. Ačkoli jsme tedy zvyklí objektivní změny ve frekvencích

vyjadřovat v lineárních jednotkách, tedy v hertzech (Hz), musíme pro analýzu intonačních jevů nalézt jednotku, která je (stejně jako naše vnímání) ne lineární, ale logaritmická. Tohoto úkolu se zhostil Patterson (2000), když na základě svého výzkumu v rámci doktorské práce označil za nejvhodnější půltónovou stupnici (s jednotkou půltón, ST, původně koncept využívaný v hudební vědě), která je logaritmická a nejlépe se vedle dalších examinovaných jednotek (Hz a ERB) podle jeho výzkumu hodí pro popis intonačního rozpětí. Ke stejnému výsledku jako Patterson došel později i Nolan (2003). Tuto jednotku budeme používat i v našem výzkumu.

2.1.3. Intonační variabilita

Intonační variabilitou míníme míru proměnlivosti intonace v určitém úseku promluvy. Tento úsek bude v práci vymezen částí determinovanou nezměněnou afektivní valencí v rámci nádechového úseku (více kap. 3.3). Zajímá nás, jak se bude intonační variabilita vyvíjet v závislosti na momentálním naladění mluvčího.

Provést správný výběr parametrů pro hodnocení intonační variability není jednoduché. Vycházíme z předpokladů, které komplexně formulují Hruška & Bořil (2017, s. 36):

In analysing the intonation of an utterance, we usually look for pitch events – movements and patterns that are relevant for the perception of intonation. It is still hard to determine what is and what is not a pitch event and there are different approaches to this issue. A pitch event may be represented by a turning point (a peak, a valley or an inflection point) in the pitch contour, by a simple rising, falling or level pitch movement or by a compound pitch movement such as a rise-fall or fall-rise. We can either observe individual pitch events and patterns and their alignment in relation to the segmental layer, or we can observe overall characteristics of the intonational contour, such as pitch range or the rate at which pitch movements occur.

Souhrnně tedy můžeme říci, že pro analýzu intonace (a potažmo její variability) můžeme použít výše popsané „pitch events“ (melodické události jako body obratu, melodické pohyby) a jejich míru výskytu, případně rozpětí f_0 , jaké je při jejich realizaci využito.

Pro tuto práci jsme pro podchycení variability zvolili tři parametry, které detailněji popíšeme v metodické části práce, seznámíme se s nimi však již zde, abychom je mohli zasadit do kontextu.

Budeme zde pracovat s intonačním rozpětím, jelikož některé výzkumy naznačují souvislost jeho rozšíření se zvýšenou emocionalitou (Palková, 1997). Dále pro analýzu použijeme poměrně nový ukazatel, kterým je parametr kumulativní strmosti (Hruška, 2016), jenž byl navržen pro celkovou kvantifikaci intonační variability v určitém úseku. Posledním parametrem, který jsme pro výzkum zvolili, je počet melodémů v nádechovém úseku. Tento parametr sice není vhodný pro komplexní popis intonační variability, ale může přinést jinou zajímavou informaci, totiž počet intonačních

frází (či promluvvých úseků) v jednom nádechovém úseku. Tyto dosud neuvedené termíny představujeme v následujících podkapitolách.

2.2. Melodém a kadence

Jedním z klíčových pojmů, které dosud nebyly vysvětleny, je melodém. Jejich počet v nádechovém úseku bude předmětem zkoumání v praktické části práce a tento parametr bude sloužit jako doplnění informací o intonační variabilitě v závislosti na afektivním ladění.

Melodém patří k základním termínům suprasegmentálního popisu řeči. Kolektiv pražského Fonetického ústavu ve svém terminologicky založeném článku popisuje melodém následovně: „Termín melodém označuje funkční melodické schéma (jako množinu variant) pro charakteristickou konturu určující konec výpovědního celku nebo konec promluvvého úseku.“ (Palková, Veroňková, Volín & Skarnitzl, 2004)

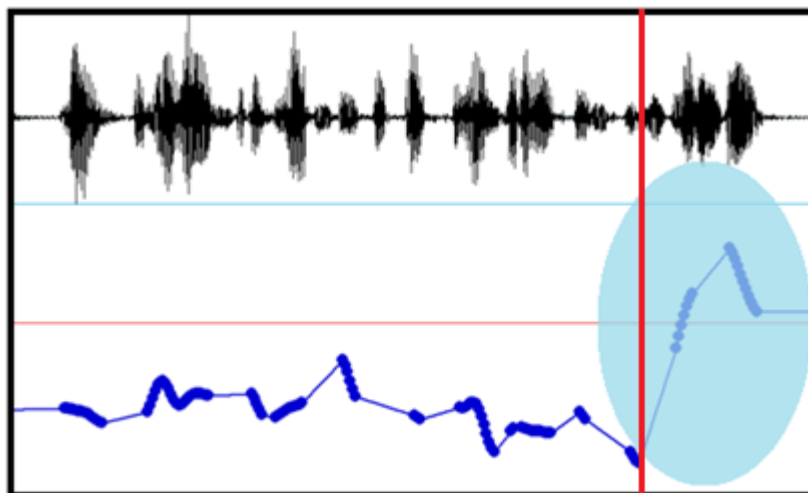
To tedy znamená, že melodém najdeme pouze na konci určitého celku, zpravidla začíná na posledním přízvuku ve větě a pokračuje až na konec tohoto celku. Výpovědní celek či promluvvý úsek však není při řeči závislý na nádechovém úseku, což znamená, že se melodémů v rámci jednoho nádechového úseku může vyskytovat i více, pokud takovýchto celků více obsahuje. Výše citovaní autoři dále osvětlují potenciálně matoucí pojem *cadence*, který podle nich označuje různé, reálně fyzicky uskutečněné varianty melodémů. Můžeme tedy říci, že zatímco *melodém* je jakýmsi abstraktním funkčním prvkem, *cadence* odpovídá jeho fyzické realizaci, která sama o sobě nese význam. Různé *cadence* jsou reprezentovány v různých mluvních stylech. (viz příhodné srovnání s problematikou hlásky a fonému ve Skarnitzl et al. 2016)

V popisu melodického inventáře českého jazyka se na základě českých prací setkáváme s následujícími základními typy melodémů:

1. melodém ukončující klesavý (MUK; M1) – vyskytuje se ve větách oznamovacích, rozkazovacích a u doplňovacích otázek, jeho realizace (*cadence*) nabývá různých podob
2. melodém ukončující stoupavý (MUS; M2) - typický pro otázky zjišťovací (s odpovědí ano/ne) – vzhledem k identické gramatické stavbě otázky zjišťovací a věty oznamovací je zde melodie fonologicky funkční (Palková, 1994)
3. melodém neukončující (MN; M3) – signalizace pokračování souvětí po skončení jeho dílčí věty

Tyto melodémy však mohou mít různou podobu při realizaci a jejich názvy jsou pouze orientační, konkrétnější a detailnější popis melodie nám poskytuje právě *cadence*, jak můžeme demonstrovat

na obrázku Praatem extrahovaného objektu *Pitch tier* níže (Obr. 1), kde můžeme zřetelně vidět celý nádechový úsek s jednou prozodickou frází a jedním ukončujícím melodémem stoupavým (zaznačen modrou elipsou), který je však realizován s kadencí stoupavo-klesavou.



Obr. 1: Melodém ukončující stoupavý s kadencí stoupavo-klesavou na konci věty tážací zjišťovací z úst moderátora Jana Punčocháře ve znění „Mají podle vás regulační poplatky ve zdravotnictví smysl?“ Svislá červená čára demonstruje začátek iniciační hlásky [s] ve slově „smysl“.

2.3. Lineární členění promluvy

Nemalou částí prozodie je otázka členění promluvy. Ve výzkumné části této práce budeme materiál dělit a analyzovat na základě nádechových úseků. To ale není zdaleka jediná možnost, jak se s členěním materiálu vypořádat. V tomto případě jsme k nádechovému úseku dospěli jako k řešení nevhodnějšímu. Proč?

Jak již bylo v této kapitole nastíněno, prozodie je obor, který trpí terminologickými nejasnostmi. Co se dělení proudu řeči na různé jednotky, platí tento problém dvojnásob – a to v tomto ohledu vycházíme pouze z česky psané literatury.

Obecně je známo, že čím větší jednotku chceme vymežit, tím větší potíže tuto snahu doprovázejí. Na segmentální úrovni to obvykle nebývá problém, s přechodem do suprasegmentální roviny se však začínají projevovat určité nejasnosti.

Nejmenší takovou jednotkou (pomineme-li slabiku) je jednotka na úrovni jednoho slova. Jednotce na úrovni fonologického slova odpovídá v prozodickém popisu tzv. mluvní takt. Tímto termínem byl v české tradici nahrazen starší termín přízvukový takt, který byl označen za nevhodný (Palková

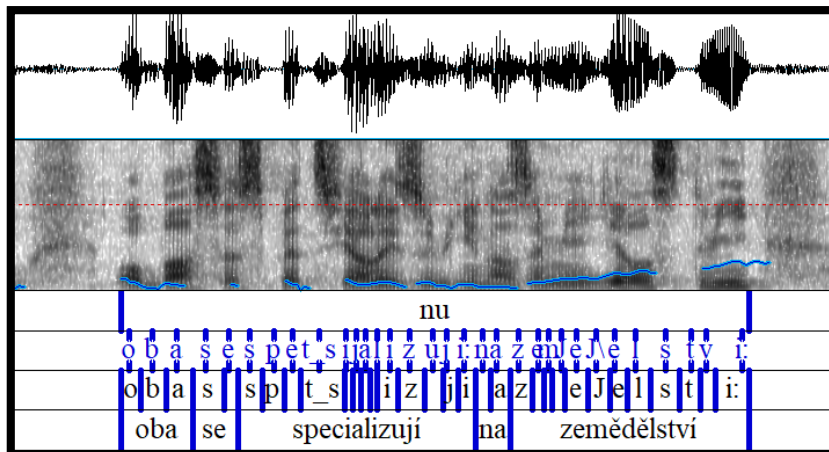
et al., 2004). Takto malá jednotka není pro intonologicky zaměřený výzkum vhodná, její představení je však nutné pro pochopení dalších konceptů.

Pokud se vydáme v této hierarchii výš, řekněme na strukturní úroveň odpovídající jednoduché větě (klauzi), začíná se terminologie dostávat do potíží. Palková (1994) na tuto úroveň staví tzv. promluvový úsek, jenž v metrických popisech odpovídá termínu kolón. Popisuje jej jako jeden takt či jejich skupinu, která je pocit'ována jako jeden sounáležitý intonační celek. V tomto kontextu se můžeme setkat i s termíny intonační či prozodická fráze (Skarnitzl et al., 2016). Vyšší jednotkou je potom promluva či výpověď – ta má v ideálním případě odpovídat jednotce na úrovni věty. Promluva se skládá z více taktů, může však obsahovat pouze jediný.

Když se podíváme na předchozí stručný popis, narážíme na problém, který nám zabraňuje v našem výzkumu některou z představených jednotek pro dělení materiálu využít. Pokud platí, že promluvový úsek může v určitém případě odpovídat jednomu taktu a zároveň promluva může odpovídat jednomu promluvovému úseku, musí také platit, že jedna promluva může být tvořena pouze jedním taktem, ale stejně tak několika promluvovými úseky. Z toho vyplývá, že kdybychom se rozhodli například pro promluvu, kterou bychom vymezovali percepčními předěly, získali bychom materiál rozkouskovaný na části o nesourodých délkách. A to ani nezmiňujeme problém, který se týká subjektivity určování hranic představených celků.

Z výše popsanych důvodů jsme se rozhodli pro náš výzkum rozdělit materiál do úseků, které jsou z našeho pohledu výhodné především z důvodu nespornosti vymezení. Jedná se o **nádechový úsek**.

Nádechový úsek (*breath-group*) je jednotka, u které můžeme definicí jasně určit její hranice. Jde totiž o tu část projevu, která není rozdělena nádechem. To znamená, že na hranicích každého nádechového úseku stojí nádech. Tato fyziologicky motivovaná skutečnost nám intenzivně pomáhá zbavit se nejasností při umíst'ování hranic, jelikož nádech je jev percepčně výrazný, zvláště v případě materiálu dobré kvality. A nejedná se pouze o sluchovou percepci, kontrola je v tomto případě možná i zraková ve spektrogramu, jak demonstruje následující obrázek (Obr. 2).



Obr. 2: Demonstrace snadného a intuitivního rozpoznání nádechového úseku (nu) v programu Praat. Z obou stran úseku (úroveň volných intervalů) jsou zřetelně vidět nádechy.

A je to také právě fyziologická motivovanost, která hraje ve prospěch užití nádechového úseku jako lineární jednotky v této práci – zajímá nás totiž, jak se který typ zažívaného afektivního stavu projevuje na frekvenci potřeby provedení nádechu. Nádechový úsek znamená jednotku přirozenou, kterou není potřeba nijak složitě definovat či přejmenovávat. Z tohoto důvodu je v moderních intonologických výzkumech často používán (např. Volín, Galeone & Johnson 2017; Volín & Bartůňková, 2015), a to navzdory tomu, na základě nádechů mohou vzniknout úseky o rozličných délkách.

3. Afektivní stavy a afektivní valence

Téma afektivních stavů, emocí a jejich dělení je velice široké a komplexní; jeho alespoň částečné pokrytí přesahuje do velké míry možnosti bakalářské práce a odpovídá často nekoordinované a rozrůzněné snaze psychologů, která trvá již dlouhé desítky let. Přesto, představujeme-li práci, která se de facto zabývá vlivem změn psychických stavů člověka (ať je autoři klasifikují, popisují či pojmenovávají jakkoli) na intonační variabilitu jeho projevu, je nutné toto téma alespoň ve stručnosti otevřít, objasnit několik základních problematických termínů a představit dřívější výzkum v této oblasti.

3.1. Problematika termínu afektivní stavy

Termín „afektivní stavy“ může být považován v mnoha ohledech za problematický. Vinna je vysoká nejednotnost pojmosloví emoční psychologie, pod kterou jejich zkoumání spadá, ale také mnoho přístupů a disciplín, které s nimi pro své účely zacházejí velice různě a často beze snahy vytvořit jednotnou terminologii.

V této práci budeme s afektivními stavy pracovat poněkud navzdory tradici české psychologie. Ta totiž, jak se zdá po povrchově provedené rešerši, přímý ekvivalent anglického *affective states* vůbec neužívá a přízvisko „afektivní“ můžeme nejčastěji nalézt v souvislosti s poruchami nálad, které jsou známé také pod názvem afektivní poruchy.

Afekt sám o sobě v české psychologii silně konotuje specifický druh (často spíše negativně chápané) emoce. Podle Encyklopedie obecné psychologie (Nakonečný, 1997) jsou afekty silné emoce, které mívají výrazné vegetativní i motorické projevy. Afektivní stav potom může být v této souvislosti chápán jako stav člověka, který se nachází pod vlivem této silné emoce. Často se tak v laické a popularizační formě psychologie setkáváme s tímto termínem v pozici jakéhosi nevladatelného emočního návalu (nejčastěji u dětí), který neuropsychologie skutečně zná, avšak označuje ho odborně za afektivní záchvaty (anglicky však *breath-holding spells* – tento termín lépe vystihuje, k čemu při takovém záchvatu dochází).

Pro nás je v tomto ohledu směrodatný význam, který navrhují Scherer et al. (2000) a dále rozvíjí Scherer (2003), kteří výrazem „*affective states*“ míní psychické stavy člověka, a to na nejobecnější možné rovině. Takto chápané afektivní stavy lze dále rozdělovat podle různých kritérií, jak uvádíme dále.

V kontrastu k Schererovi se s afektivními stavy vypořádává respektovaná Psychologie Atkinsonové a Hilgarda (Nolen-Hoeksema, 2012) a její český překlad. Zde jsou afektivní stavy položeny na roveň pocitům, které doprovázejí emoce při subjektivním prožitku, jenž je podle této příručky pouze jednou ze šesti složek emočního procesu. Ve slovníkové části publikace se však výklad afektivních stavů s chápáním Scherera (2003) do jisté míry shoduje, když jsou zde jako „méně určité afektivní stavy“ definovány nálady. Je však nutné dodat, že v totožné slovníkové části, jen o pár stran dále, termínem „afektivní stav“ autoři vysvětlují význam subjektivního prožívání. Tato nesrovnalost nicméně dobře dokládá značnou komplikovanost celé emocionální psychologie, zejména pak ošemetnost související terminologie – v některých pracích bývají afektivní stavy a emoce volně zaměňovány a užívány zcela synonymně, v jiných pojetích je naopak jejich nezáměnnost zcela esenciální. Ani jeden z pojmů tak není možné jasně definovat.

3.2. Schererovo pojetí afektivních stavů jako východisko pro práci

Emoce a afektivní stav jako nezáměnné pojmy prezentuje Scherer (2003), jehož výkladu je možné se v terminologické rozrůzněnosti (možná do jisté míry způsobené některými ztrátami v překladu) přidržet. Tento švýcarský psycholog nakládá s afektivními stavy jako s množinou psychických stavů, které je možné navzájem vymezit za pomoci modelu sedmi charakteristik, jimiž jsou jejich intenzita, trvání, synchronizace (fyzické změny v těle), zaměření, rychlost změny, dopad na okolí a dopad na chování. Tato kritéria mu umožňují vytvořit klasifikaci afektivních stavů do pěti skupin, které v překladu (Skarnitzl et al., 2016) vyznívají následovně:

- a) emoce
- b) nálady
- c) postoje
- d) aktuální ladění
- e) osobnostní vlastnosti

Dalším nesnadným úkolem zůstává určit, který z těchto afektivních stavů nejlépe odpovídá stavům, které do své řeči promítá účastník politické debaty, jež je materiálem pro praktickou část této práce. Pokud vezmeme v úvahu Schererovu klasifikaci, můžeme začít zvažovat například emoce, jejichž vlivy na řeč bývají často zkoumány, jak uvádíme v následující podkapitole. Jsme však při řeči vystaveni nepřetržitě emocím, které jsou podle jeho vymezení intenzivní, krátkodobé a mají velký dopad na změny jak v těle, tak v chování? Bezesporu se v řeči také odrážejí, ale nejde o stavy, které

bychom zažívali nepřetržitě. Je však nezpochybnitelné, že určitým stavům jsme při komunikaci vystaveni neustále.

Za nejlépe odpovídající afektivní stavy pro komunikační situace považují Skarnitzl et al. (2016) aktuální ladění (*attitudes*), které ze svého vymezení v Schererově modelu vychází jako určitý krátkodobý postoj mluvčího obsahu sdělení či k osobě, která toto sdělení přináší. Aktuální ladění je stav, který lidskou komunikaci prostupuje ve všech jejích momentech. V určitých situacích však může být mluvčím vědomě potlačováno, či naopak intenzifikováno.

3.3. Dosavadní výzkumy projevů afektivních stavů v řeči

Jak již bylo nastíněno v úvodu, výzkumy vlivu psychických stavů na řečový projev byly v minulosti zastíněny zájmem o jejich projevy faciální. Přesto však můžeme zájem o toto téma pozorovat již ve starověku. Vývoj zkoumání shrnuje přehledně opět Scherer (2003). Na základě tohoto přehledu můžeme zájem o projevy emocí v řeči rozdělit do několika historických etap. Pozoruhodné je, že v každé z etap zajímalo toto téma vědce z jiných oborů – jde často o významná jména světové vědy. Tento fakt dokládá širokou interdisciplinaritu vědní oblasti, do které jsme se ponořili.

Ve starověku byla afektivní řeč předmětem zájmu autorů, kteří byli fascinováni především uměním rétoriky, jako například Aristoteles či Cicero. Poté zájem na dlouhou dobu utichl; znovu se začal projevovat až v devatenáctém století na poli zkoumání evolučních biologů (Ch. Darwin, G. Bell) a pokračoval na začátku století dvacátého, kdy se jeho těžiště přesunulo do rukou psychiatrů (např. B. F. Skinner).

Jak však uvádí vzápětí Scherer (2003), systematický výzkum souvislosti emocí s řečovými projevy začal probíhat až v šedesátých letech dvacátého století. Opět šlo o iniciativu psychiatrů, kteří se pokoušeli skrze ně diagnostikovat různé stavy. Během šedesátých a sedmdesátých let došlo také k tzv. komunikačně-pragmatickému obratu v lingvistice, který k diskutované problematice navedl také jazykovědce. A v posledních desetiletích se poznatků z této oblasti využívá zejména v řečovém inženýrství, kde napomáhají například při redukci nepřirozenosti řečové syntézy, jelikož prozodie je všeobecně považována za složku, která je pro její přijatelnost stěžejní a zároveň se jí stále nedaří uspokojivě nasimulovat (Mozziconacci, 2002).

Právě současné výzkumy pak rozděluje Laukka et al. (2005) do dvou hlavních přístupů. První můžeme nazvat jako „syntetizující“, jenž spočívá v systematické manipulaci akustických charakteristik řeči a následném hodnocení vzorků posluchači na škálách emočních dimenzí. Druhý přístup se liší pouze tím, že posluchači nehodnotí manipulovanou řeč, ale řeč s (většinou) uměle

vyjádřenou emocí. Cílem obou zmíněných postupů je zjistit, jak které akustické změny vnímají mluvčí (a nejde jenom o intonační charakteristiky, ale také například o temporální či intenzitní, čili „hlasitostní“) a s jakým aktuálním laděním či emoční zabarveností je spojují. Zatímco tyto výzkumy můžeme označit za „dekódovací“ (*decoding* – Scherer, 2003), zmiňme ještě přístup opačný, tedy „kódovací“ (*encoding*). Spočívá v examinaci akustických profilů vzorku expresivní řeči – jde tedy o zkoumání, jež nazírá na stejnou problematiku, pouze z odlišné perspektivy. Tyto studie můžeme rozdělit do tří skupin podle povahy vzorku: přirozená expresivní řeč, řeč s navozenými emocemi (za použití psychoaktivních látek) a řeč se simulovanými emocemi (často profesionálními herci). Právě tím, že bývají pro podobné výzkumy nejčastěji používána pouze jakási nespontánní vyobrazení emocí, je jejich přínos poměrně limitován. Při uměle projevovaných expresích mají mluvčí tendenci projevy v závislosti na kulturních stereotypch a ve snaze vyhovět zadání zintenzivňovat a schematizovat. Mezi důvody, proč výzkumy nespontánní emotivní řeči nemohou dosahovat přínosů spontánních výzkumů bychom mohli zařadit i důvod, že posluchači jsou globálně schopni simulované emoce a reálné emoce v řeči rozlišit (Audibert, Aubergé, & Rilliard, 2008; Sauter, & Fischer, 2018).

Přesto však pro představu zmiňme některé dílčí výsledky takovýchto studií, především ty, které se týkají zjištění o intonačních jevech.

Pittam & Scherer (1996) shrnují dosavadní výsledky předchozích kódovacích studií pěti základních emocí – vzteku, znechucení, strachu, radosti a smutku. V souvislosti s intonací se vztek v mnoha v nich projevuje zvýšenou základní frekvencí, zvýšenou intonační variabilitou a také zvýšeným intonačním rozpětím, stejně jako například radost. Zvýšenou intonační variabilitu a rozpětí vykazují v mnoha případech také projevy strachu. Zdá se, že opačné tendence mají projevy smutku, kde je průměrná frekvence nižší a intonační rozpětí užší. U znechucení dosud nebyly žádné konzistentní výsledky doloženy. Celkově se zdá, že pouze za pomoci zmíněných parametrů není možné od sebe emoce rozeznávat. Banse & Scherer (1996) tak navrhují do budoucích výzkumů zapojit co nejvyšší počet akustických parametrů. Ve vlastním výzkumu se zaměřili rovnou na 14 různých emocí, které byly pro co nejvyšší autenticitu předvedeny profesionálními herci na scénářích vybraných ze speciálního korpusu emočních situací. Každá realizace pak byla pečlivě ohodnocena a vybraný počet těch nejlépe rozpoznatelných a nejautentičtějších byl použit pro akustickou analýzu. Z intonačních parametrů byla pro každou položku měřena průměrná f_0 , standardní odchylka a 25. a 75. percentil hodnoty f_0 – později autoři zjistili, že tyto parametry vzájemně víceméně silně korelují. Nejvyšší průměrná f_0 byla naměřena u těch emocí z výběru, které autoři považují za „nejintenzivnější“ – žhavý hněv, panický strach a euforie. Nejnižší f_0 z měřených emocí vykazovala nuda, ostatní se pohybovaly v specifikované blízkosti středu.

Laukka et al. (2005) provedli komplexní výzkum nespontánních projevů pěti základních emocí v řeči: vztek, znechucení, strach, radost a smutek. Tyto emoce byly nahrány a následně percepčně ohodnoceny studenty a experty na škálách aktivace, valence, intenzity a potence – tedy na škálách dimenzionálního modelu. V dalším kroku byla provedena analýza korelace hodnot těchto dimenzí s různými akustickými parametry, kromě jiného také s těmi intonačními. Ve výsledcích se korelace intonačních parametrů projevily v souvislosti s vysokou aktivací, potencí a intenzitou napříč zkoumanými emocemi. Vysoké hodnoty na těchto škálách korelovaly shodně s vysokou variabilitou f_0 . Vysoká aktivace a intenzita emocí rovněž vykazovala vysoké hodnoty průměrné f_0 a jejího vysokého maxima. Z těchto zjištění lze usuzovat, že celkově emocionálnější řeč bude spíše obecně intonačně výraznější a variabilnější.

Jako poslední zmiňme studii Rodero (2011), která byla zaměřena na zkoumání využívání specifických konturových intonačních typů a výšku hlasu jako strategie při vyjadřování emocí a na jejich percepční důležitost. Pro své účely vybral autor čtyři emoce, které uznal jako vhodné vzájemné protějšky na valenční a aktivační ose – radost, úzkost, klid (*calmness*) a smutek. Vychází z teorie, že vyšší průměrná frekvence je typická pro emoce s vysokou aktivací, jako jsou radost, úzkost či strach, nízkou průměrnou frekvenci spojuje s emocemi s aktivací nízkou – smutek, klid, pocit bezpečí. Emoce předvedli herci na totožné větě a následně realizace ohodnotila stovka studentů na škále, která zachycovala jejich mínění o typičnosti dané realizace pro konkrétní emoci. Nahrávky s průměrným nejvyšším skóre (tedy nejbližší k zamýšlené emoci) byly potom porovnány s ostatními a byly popsány charakteristiky, které se pravděpodobně o jejich vysoké skóre zasloužily. Nejtypičtější radost se podle nich vyznačovala velkou variabilitou kontury a vysokou průměrnou f_0 z dalšího porovnávání vyšlo najevo, že u této emoce byla pro posluchače důležitější variabilita kontury než vysoký průměr frekvence. Stejný model byl nalezen i pro úzkost, kde však výsledky naznačují, že variabilita kontury hraje ještě větší roli než u radosti. U smutku získala nejvyšší skóre performance s nejnižší průměrnou frekvencí, ale také s nejnižší variabilitou. Hodnocení dalších nahrávek i zde naznačují větší důležitost variability kontury než průměrné výšky f_0 . Pro klid (*calmness*) byla nejtypičtější nahrávka, která sice neměla ze všech z poskytnutých nejnižší průměrnou frekvenci, ale vyznačovala se nejmenší variabilitou kontury, což je opět informace, která podporuje dosud pozorované zjištění, že variabilita kontury je pro percepci důležitější než průměrná výška f_0 , což je jedním z hlavních přínosů této studie. Dalším přínosem je i vytvoření charakteristik těchto čtyř emocí za pomoci uvedených parametrů, které by jistě bylo zajímavé testovat i na dalších emocích.

Studíí, které se zaměřují pouze na zkoumání emocí v řeči spontánní, je i navzdory obecně vnímanému rozdílu mezi přirozeným a nepřirozeným projevem stále poměrně málo. Mozziconacci (2002) se k důvodům obtížnosti práce s přirozeně emočně zabarvenou řečí vyjadřuje následovně: „*The recording of genuine emotion raises ethic issues, as well as difficulties with emotion labeling, and control of the recording situation.*“ Je pravda, že pro výzkumy diskretních emocí není jednoduché všechny tyto problémy odstranit. Opravdu emotivní spontánní projevy mohou být velice citlivým materiálem a představa, že by někdo dobrovolně takové situace nahrával a poskytoval k výzkumu, je zřejmě nereálná. I s touto těžkostí si však někteří vědci dokáží kreativně poradit – například Palacios-Alonso et al. (2015) navrhli metodologii sběru dat při kooperativní videohře (multiplayer). Situace při videohře může být opravdu velice emotivní a zároveň vysoce spontánní, tudíž jde o velice chytrý způsob získání velkého množství emotivního materiálu, ačkoli je nutné počítat s tím, že distribuce zastoupených emocí nebude rovnoměrná.

Jak jsme však uvedli v předchozí podkapitole, nejsou to zřejmě emoce, jejichž projevy jsou v řeči nejvíce zastoupené, i když je pochopitelné, že jsou díky rozšířené představě jakési ohraničenosti (alespoň v rámci uvažování teorie diskretních emocí) na výzkumném poli afektivních stavů atraktivním objektem zájmu, dokonce takovým, že na základě nahraných emocí vznikají i databáze, mezi nejznámější patří například Dánská databáze emocionální řeči (DES) a Berlínská databáze emocionální řeči (BES).

3.4. Afektivní valence v dimenzionálním modelu

Od afektivních stavů se přesuňme k pojmu afektivní valence (*affective valence*), která je zásadním termínem pro práci a jeho význam považujeme za nutné zde vymežit. Briesemeister et al. (2012), stejně jako další články, které se dotýkají studia emocí a jejich promítání v mluveném projevu, pracuje s valencí jako s jednou z dimenzí emocí v dimenzionálním modelu. Dimenzionálních přístupů (které v psychologické teorii stojí v opozici k tzv. diskretním přístupům) bylo dosud navrženo více a za jejich otce můžeme označit německého interdisciplinárního vědce Wilhelma Wundta, jenž navrhl první dvoudimenzionální (2D) model. Ten měl umožnit odlišit od sebe různé druhy afektivních stavů, ne však ve smyslu, v jakém jejich dělení navrhl Scherer (s jehož pojetím jsme se seznámili v kap. 3.2), ale spíše na kvalitativní rovině. Postupně byly navrhovány další modely, některé pracovaly se dvěma, některé až se čtyřmi dimenzemi. Tou nejméně rozporovanou a nejčastěji zastoupenou dimenzí byla vždy právě afektivní valence.

Napříč literaturou je valence popisována jako jakýsi ukazatel **příjemnosti** daného afektivního stavu (s pólem *pleasantness* na jedné straně a protipólem *unpleasantness* na straně druhé), která je považována za jeden z „distinktivních rysů“ při popisu konkrétních emocí.

Laukka et al. (2005) v již představeném výzkumu pracují také s dimenzionální modelem (celkem o čtyřech dimenzích). Ohledně valence však autoři zmiňují, že na rozdíl od ostatních dimenzí se nepodařilo pouze na základě změřených akustických parametrů dopředu předpovědět hodnocení posluchačů. Přičítají to buď tomu, že nezachytili pro valenci ty pravé určující parametry, nebo pracují s variantou, že je možná valence v hlase reprezentována méně výrazně než dimenze ostatní. Studie však prokázala, že přece jen existují určité vzorce, které poukazují na souvislost některých jevů a exprese emocí s pozitivní valencí. Pro naši práci je z citovaného článku relevantní zmínka o „velké variabilitě f_0 “ (*large f_0 variability*) v souvislosti s pozitivní valencí, o které ostatně vypovídají i výsledky dalších studií (např. Scherer & Oshinsky, 1977). Bohužel toto zjištění o pozitivní valenci zde není uvedeno do kontrastu s valencí negativní a ze stručného popisu metodiky se ani nedozvídáme o tom, co vlastně „*large f_0 variability*“ v pojetí autorů znamená.

Teorii o souvislosti pozitivní valence a širšího intonačního rozpětí podporují i výsledky Recklinga & Küglera (2011). Jejich výzkum je tomu našemu v mnoha ohledech blízký. Tito autoři zkoumali vliv pozitivity či negativity v řeči na intonační rozpětí. Vycházejí při tom z hypotéz (položených na základech teorie biologických kódů), že promluvy s pozitivním obsahem budou vykazovat širší intonační rozpětí proti promluvám konotovaným negativně. Tito výzkumníci provedli tři experimenty (dva kódovací a jeden dekódovací), na nichž tuto hypotézu testovali. V rámci prvního experimentu šlo o čtení valenčně odlišných replik dialogů. Výsledkem (nutno podotknout, že autoři nepracovali se pultónovou stupnicí, ale s Hz a omezili se na počítání maximálního rozpětí) bylo celkové potvrzení jejich hypotézy – negativně konotované věty měly průměrně celkem o 52 Hz užší intonační rozpětí než ty pozitivně konotované.

Druhý experiment byl podobný, valenčně odlišné věty však byly participanty pouze čteny v kombinaci s výplňkovými větami (které pravděpodobně předpokládaly jakési neutrální sdělení – k této problematice více v sekci 3.4.1.). Výsledky zkoumání širě intonačního rozpětí byly v této části srovnatelné s předchozím experimentem – autoři naměřili průměrně o 56 Hz širší rozpětí u pozitivních promluv.

Poslední experiment byl percepční – respondenti měli za úkol hodnotit věty z prvního zmíněného experimentu na škálách *polite/ inappropriate*. Vybrány byly věty, které představovaly „dobré“ i „špatné“ (tedy vhodné a nevhodné) zástupce svých kategorií na základě průměrného rozpětí. Výsledky ukázaly, že „dobré“ pozitivní věty (tedy se širším rozpětím) získaly od respondentů

pozitivnější hodnocení než „špatné“ pozitivní a „dobré“ negativní (redukovanější rozpětí) získaly negativnější hodnocení než „špatné negativní“.

Tato zajímavá sada experimentů tedy hypotézu autorů prokazuje. Je na ní sympatické, že nepracuje přímo s předstíranými emocemi, ale využívá přirozeného dojmu, který z předložené věty mluvčí získá a podle toho ji interpretuje, aniž by dostal další externí instrukce. Tato situace o něco lépe simuluje běžnou řeč, jelikož nevyžaduje zapojení emocí, na které je u ostatních výzkumů kladen důraz, ale pracuje s něčím, co je blíž právě aktuálnímu ladění.

S kritikou tradičního (a v literatuře nejrespektovanějšího) modelu, který valenci zobrazuje na kontinuu se dvěma protilehlými póly, přichází Briesemeister et al. (2012). Ten představuje valenci na evaluativním modelu (*evaluative space model, ESM*), jenž odmítá tradiční teoretické východisko, že emoce jsou ze své povahy buď pozitivní, nebo negativní. Proto staví své experimenty na modelu o dvou valenčních dimenzích – jedné pro pozitivní a jedné pro negativní valenci. V práci se mu podařilo dokázat, že se respondenti nezděraňují popsat některé emoce pomocí různých měř na obou těchto škálách. Je to přístup jistě opodstatněný, jelikož bezpochyby existují emoce, které například Nakonečný (2000) označuje za smíšené – řadí k nim například nostalgii či lítost. Jelikož se však v této práci nechystáme zkoumat diskrétní emoce, ale pouze aktuální ladění ve spontánní řeči (kde předpokládáme, že se bude percepčně jedna z případných dvou valenčních složek projevat výrazněji), budeme se držet přístupu tradičního.

Je to i aktuální ladění, u kterého se afektivní valence projevuje. Například Skarnitzl et al. (2016) těmto stavům přiřazují dvě dimenze – dimenzi přijetí/odmítnutí a dimenzi míry či intenzity. Z jejich popisu vyplývá, že první ze zmíněných koresponduje s obecnou představou afektivní valence, tudíž je možné ji zkoumat i u materiálu, který není primárně emotivní, jen je potřeba jeho pečlivější examinace a důkladné seznámení s mluvčími a celou komunikační situací.

3.4.1. Problematika neutrality afektivní valence při řeči

Ačkoli se v této práci spoléháme na tradiční pojetí afektivní valence, považujeme za nutné vyvázat se z obvyklého chápání tohoto modelu s ohledem na potenciální „neutrální bod“ mezi negativní a pozitivní částí bipolárního kontinua. V dimenzionálních modelech, alespoň jak je popisuje Briesemeister, Kuchinke & Jacobs. (2012), bývá totiž brán za samozřejmý, aniž by bylo přihlédnuto k faktu, že neutralita je v řeči jen těžko pozorovatelným jevem.

O spornosti neutrality řeči referují například Fónagy (1983) či Skarnitzl et al. (2016). K tomuto názoru se v rámci naší práce rovněž přikláníme, ačkoli přiznáváme, že není často jednoduché

valenci v komunikaci určit (viz kap. 5.2). Primárně je tomu tak logicky v případech, kdy není člověk sám bezprostředním účastníkem konkrétní komunikační situace, ale stojí v roli pozorujícího, který ji má zkoumat. V takové pozici je velice jednoduché při těžko identifikovatelných situacích sklouznout k řešení, které se zdá být nejjednodušší – tedy prohlásit promluvu s ne zcela zřejmým afektivním laděním za neutrální.

Je nutné si uvědomit, že právě komunikační situace je tím, co afektivní ladění při řeči primárně ovlivňuje. Při verbální komunikaci (například v kontextu zkoumané politické debaty) jakožto mluvčí vždy **reagujeme** na situaci, které dává vzniknout předchozí vývoj rozhovoru (participant se například ohrazuje, dává najevo souhlas/nesouhlas, pobavení a podobně – vždy v reakci na předchozí repliky či obecný vývoj rozhovoru), nebo na situaci, do které je mluvčí vsazen ostatními okolnostmi (například moderátor, který na začátku diskusi uvádí – ačkoli nemůže být ovlivněn přímo předchozím hovorem, protože zatím žádný neproběl, je jeho výstup je nutně valenčně definován ostatními okolnostmi – potřebou působit na diváky přívětivě, pozitivně, přísně a podobně).

I přes legitimitu této úvahy se však můžeme setkat také s termíny, jako je „neutrální sdělení“. V literatuře se můžeme například dočíst o experimentech, které se snažily neutralitu pro své potřeby uměle nasimulovat (Vlčková Mejvaldová, 2006) – zde ovšem nemůže být uvažováno o přirozené a spontánní řeči, jejíž charakteristiky se snažíme mapovat. Kromě toho si také musíme uvědomit, že i snaha o produkci neutrálního sdělení bude nutně ovlivněna přístupem, který k takovému úkolu mluvčí zaujímá. Vlčková Mejvaldová (2006) například pro účely svých experimentů (porovnávání prozodických charakteristik různých jazyků a jejich vzájemné porozumění) instruuje své mluvčí, aby vyřkli zcela bezpříznakově větu „*Kláro, stalo se mi něco nečekaného*“. Problém s podobným zadáním je hned několikanásobný – mluvčí se nemůže vyvázat z postoje, který vůči úkolu či vůči sémantickému vyznění věty zastává, a i kdyby to v experimentálních podmínkách dokázal, důvod zkoumání takového stavu vlastně neexistuje, jelikož ve spontánní řeči už z její podstaty podobná situace nemůže nastat. Zmiňovaná autorka si je zřejmě kontroverze neutrality vědoma, proto záměr jejího zkoumání přejmenovává na bezpříznakovost, jež podle ní znamená, že se mluvčí nesnaží o vědomé vyjádření nějakého postoje. I to je ale situace umělá a pro celkové poznání nepřínosná. Z výše popsaných důvodů možnost neutrality z naší metodiky vynecháváme.

Hodnotu afektivní valence budeme v našem materiálu určovat v každém z nádechových úseků. Pro detailnější sledování trendů ve výsledcích jsme se však rozhodli hodnotit i její míru (výraznost) – jinak řečeno, projev může být nepříjemný či příjemný „malinko“, „trochu“, „docela dost“, „hodně“ a podobně. V práci budeme rozlišovat výraznost valence na dvou hladinách, a to na nižší (méně výrazné) a vyšší (výraznější).

Co se určování valence týká, mohli bychom (vcelku oprávněně) čelit námitkám, že neexistuje spolehlivý způsob, jak objektivně zjistit její hodnotu. Řešení této problematiky se více věnujeme v kapitole věnované metodě výzkumu (5.2).

3.5. Specifika prostředí politické debaty při vyjadřování afektivních stavů

Doposud jsme hovořili o obecných zjištěních v oblasti vyjadřování afektivních stavů při řeči. Okrajově jsme zmínili i jejich závislost na různých komunikačních situacích a vyjádřili jsme přesvědčení, že jsou v určité míře přítomny ve všech každodenních interakcích.

Materiálem pro představovanou práci je záznam řeči z vysoce specifického prostředí, které představuje televizní politická debata. Tento žánr se sice vyznačuje spontaneitou, na druhou stranu ale klade na zúčastněné poměrně vysoký tlak, který pochází ze snahy předvést sebe (jakožto momentálního zástupce politické strany či zastávce určitých názorů) v tom nejlepším světle, tudíž je nutné vzít v úvahu, že se aktéři nebudou chovat zcela přirozeně – možná právě kvůli tomu, že se budou snažit co nejpřirozeněji působit.

Všichni zúčastnění se nacházejí v pozici, která jim umožňuje v dané situaci mnoho získat, ale také mnoho ztratit. Často nezáleží pouze na obsahu prezentovaných názorů, ale rovněž na způsobu jejich podání. Pokud politik podá v televizní debatě „kvalitní“ a sebejistý výkon, může na svou stranu přiklonit mnohé nerozhodné potenciální voliče, působí-li ve vysílání nerozhodně, může dojít i k opačnému efektu.

Ne zcela rozdílný úkol má také samotný moderátor debaty – místo o potenciální přívržence politické strany bojuje o příznivce svého pořadu, případně celého televizního kanálu, proto rovněž záleží na tom, jak se v diskusi prezentuje. Rozhodný, připravený a nesmlouvavý moderátor se může v dobře vedené debatě zviditelnit a stát se pro předvedené kvality velice respektovaným a diváky vyžadovaným – tak jako se to podle mnoha médií i politiků podařilo například moderátorce České televize Světlaně Witowské v poslední předvolební politické debatě v roce 2018 mezi prezidentskými kandidáty Jiřím Drahošem a Milošem Zemanem.

Kromě toho má však moderátor ještě další roli: teorie médií konstatuje, že zajímavost debaty je dána především její vyhroceností, kterou právě moderátor může svým konáním ovlivnit. Jeho úlohou je podněcovat roztržky a akcentovat především takové aspekty diskutované situace, na které se přítomní politici neshodnou (Čmejrková et al., 2013). Proto je na místě od pořadu podobného formátu očekávat spíš celkové negativnější ladění.

Pokud jsou si všichni účastníci vědomi potenciálu debaty, budou se snažit chovat tak, aby se divákům co nejvíce zalíbili. Podle Čmejkové je důležité, aby politik v napjatých a často kontroverzních situacích, do kterých jej moderátor dostává, byl schopen udržet si svou tvář. To znamená, že musí čelit nejenom protivníkovi a nenechat se jím zahanbit, ale musí být ostražitý také vůči moderátorovým často manipulativním otázkám. Musí si také umět ubránit slovo a své území v podobě časového úseku, který mu moderátor nabízí k reakci, aby se prezentoval jako silný zástupce.

Navzdory všem uvedeným zákonitostem je však cílem zúčastněných rovněž prokázat jistou racionalitu, chladnou hlavu a kooperaci. Proto dochází ke konstantnímu střetu cílů (Čmejková et al., 2013), které ústí ve snahu účastníků i přes zachování vlastní tváře prokázat, že jsou za určitých podmínek schopni kultivované diskuse a kompromisů, že jsou si sice prezentovanými názory jistí, ale ne zaslepení, což přispívá k dojmu jisté moudrosti a vyzrálosti zúčastněného (nejen) politika.

Z výše představeného konceptu politické debaty a jejích cílů je nasnadě vzhledem k užitému materiálu učinit následující východisko:

Ačkoli jde o žánr, který má divákům přinést vedle informovanosti také zábavu v podobě vyhocenějších situací (*infotainment* – spojení informací a zábavy), je zároveň v zájmu samotných účastníků, aby se nenechali vyprovokovat k příliš emotivním reakcím, které by mohly být interpretovány jako zbytečně hysterické, a ztratit tvář. Lze tudíž očekávat, že aktuální ladění zúčastněných osob bude během debaty spíš negativní, nikoli však nutně přespříliš. Jak jsme již uvedli v předchozí podkapitole (3.4.1), afektivní stavy ve spontánní komunikaci jsou ovlivněny především situacemi, do kterých se mluvčí dostává. Říkáme-li, že se moderátor snaží tyto situace udělat nekomfortní, dá se (možná poněkud potlačená) negativita v průměrné politické debatě jistě očekávat spíš než pozitivita.

4. Hypotézy a cíle práce

V teoretické části práce jsme se nejenom seznámili se základními koncepty a pojmy, ale představili jsme také dosavadní výzkumy, jejichž znalost nám usnadní orientaci v tématu a umožní lépe definovat výzkumný úkol a formulovat naše hypotézy. Využijme proto tohoto prostoru pro krátkou rekapitulaci získaných poznatků a jejich propojení s praktickým úkolem této práce.

Při rešerši jsme zjistili, že dosavadní práce na poli promítání afektivních stavů do prozodických aspektů řeči je orientována více na diskrétní emoce a na řeč nespontánní, jelikož práce se spontánní řečí naráží na mnoho metodologických problémů a žádoucí silné emoce se v běžně dostupných zdrojích přirozené řeči vyskytuje spíše pomálu, což značně komplikuje jakýkoli kvantitativní výzkum. Dostupné zdroje spontánní řeči (například politická debata) oproti tomu poskytuje spíše projevy aktuálního ladění, které je v porovnání s emocemi slabší, zato však stabilně přítomné, tudíž vhodné pro výzkum. Z těchto důvodů se domníváme, že dosud zjištěné tendence projevů emocí se v našem výzkumu mohou u příslušných typů afektivního ladění potvrzovat, ale v méně intenzivní míře, jelikož se dá očekávat, že slabší afektivní stav se bude i slaběji projevovat.

Afektivní valence, jejíž vlivy jsme se rozhodli zkoumat, je parametrem ve výzkumech velice populárním. Její vliv na intonační variabilitu však nebyl dosud konzistentně prokázán ani u uměle předváděných emocí, ačkoli některé výsledky více nezávislých výzkumů spíše spojují pozitivitu s vyšším intonačním rozpětím (např. Reckling & Kügler, 2011), jiné výsledky však naznačují, že se intonační charakteristiky emocí ležících v protikladu na valenční ose příliš neliší, naopak se zdá, že jsou v tomto ohledu poměrně blízké, jak konstatovali například Banse & Scherer (1996) u dvou tak protikladných emocí, jako je vztek a euforie. Obecně je také vyšší intonační rozpětí dáváno do souvislosti s intenzivnější mírou emocionality a celkové excitace (například Laukka et al., 2005; Palková 1997).

Z výše popsaných, spíše kusých poznatků máme důvod se domnívat, že intonační variabilita bude vykazovat závislost především na síle (či míře) valence promluvy než na příslušnosti k té či oné lateralitě valenční osy, i když

Domníváme se, že extrémnější valenční hodnoty budou korelovat se zvýšenou intonační variabilitou, zatímco „základní“ valenční hodnoty budou oproti nim vykazovat variabilitu nižší, a to především v intonačním rozpětí. Zároveň ze stručného představení konceptu politické debaty, která slouží jako výzkumný materiál pro tuto práci, je zřejmé, že extrémnější valenční hodnoty budou zastoupeny řidčeji než ty základní, což by mohlo vést k nerovnoměrnému zastoupení jednotlivých stupňů a potížím při zobecňování tendencí.

Také je nutné uvést, že počítáme s nutnou individualitou mluvčích, které máme pro náš výzkum k dispozici pouze tři, tudíž bude intonační variabilita každého z nich analyzována zvlášť. Jedním z mluvčích je navíc moderátor debaty, který se objevuje v jiné roli než zbylí dva mluvčí. Bude proto zajímavé výsledky mluvčích z použitého materiálu porovnat vůči sobě (a to i s ohledem na roli, ve které se v situaci politické debaty nacházejí), identifikovat individuální rysy, případně velice nezávisle poukázat na tendence, jež by byly pro všechny zúčastněné mluvčí společné. Za další cíl si klademe kriticky zhodnotit navrženou metodologii a na základě jejích problémů navrhnout v návaznosti na tuto práci jejich řešení pro potenciální následující výzkum v této oblasti.

5. Metoda

5.1. Materiál

Zvukový materiál pro potřeby této bakalářské práce byl vybrán z portfolia materiálů poskytnutých pro účely Fonetického ústavu (FÚ) českou mediální společností FTV Prima spol. s r. o. Jedná se tedy o velice kvalitní a profesionálním televizním štábem pořízený materiál, který vyhovuje nejen požadavkům pro uskutečnění stanoveného výzkumného úkolu této práce, ale může být též zařazen do korpusu Fonetického ústavu a použit k dalším účelům.

5.1.1. Charakteristika použitého materiálu

Výběr materiálu proběhl nahodile ze série poskytnutých dílů televizního politického diskusního pořadu Nedělní partie. Tento pořad svým formátem vyhovuje účelu práce – debata je žánrem, jež se vyznačuje spontaneitou a dochází v ní k častým interakcím, které ovlivňují afektivní valenci řeči u všech mluvčích.

Jde o pořad s jedním moderátorem (Jan Punčochář) a dvěma hosty. Konkrétním materiálem použitým pro práci je zvukový záznam debaty vysílané na FTV Prima dne 7. 2. 2010. Název dílu je „*Boj o zdravotnické poplatky pokračuje*“ a hosty jsou političtí protivníci Michal Hašek, tehdejší hejtmán Jihomoravského kraje (reprezentující názory ČSSD), a Petr Gandalovič, tehdejší ministr zemědělství (reprezentující názory ODS). Díky obsazení debaty zástupci dvou nejsilnějších politických stran té doby dochází v rámci materiálu k živým (nicméně stále kultivovaným) názorovým přestřelkám. Tématem je především politika a její nejpálčivější problémy z doby období pořízení nahrávky: zavedení zdravotnických poplatků, stavba silnic a podobně.

5.1.2. Příprava materiálu pro analýzu

Před analýzou musely být na materiálu provedeny úpravy, které detailněji popíšeme v této kapitole. Materiál ve formátu .wav byl nejprve nařezán v programu Adobe Audition na části o délce přibližně 20-30 sekund. V této práci jsme analyzovali záznam o celkové délce dvacet minut, která odpovídá předepsané délce nahrávek pro bakalářskou práci.

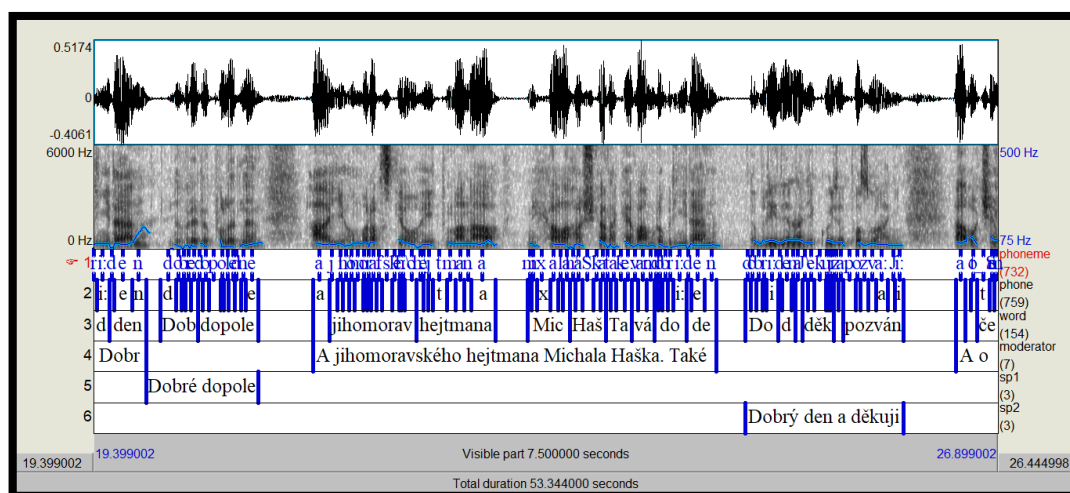
Pro každý úsek byl ve fonetickém programu Praat (Boersma & Weenink, 2020) vytvořen samostatný soubor TextGrid (TG), který ve fonetickém výzkumu používáme pro vytvoření

korespondence psaného textu a řečové nahrávky (anotace zvuku, segmentace, labelování hlásek). Do tohoto souboru byl nakopírován příslušný přepis řeči a po poslechové kontrole a korekci byl programem pro automatickou segmentaci *Prague labeller* (Pollák et al., 2007) text nasegmentován do tří intervalových vrstev (interval tiers) – *phrase* (text celé části), *word* (text rozdělený na jednotlivá slova), *phone* (realizovaná hláska) a jedné bodové vrstvy (point tier) *phoneme* (foném, jemuž odpovídá hláska ve vrstvě *phone*).

TG byl následně doplněn vrstvami, do kterých byl rozdělen a nakopírován text podle konkrétních mluvčích (označené jako „*moderator*“ = Jan Punčochář; „*sp1*“ = Petr Gandalovič; „*sp2*“ = Michal Hašek), vrstva *phrase* mohla být po tomto kroku odstraněna. Automatická segmentace pak vyžadovala manuální úpravy v několika ohledech:

1. úprava hranic segmentů (konzultováno s titulem Fonetická segmentace hlásek – Machač & Skarnitzl, 2009). Důležité bylo především umístění vokálů a jejich hranice – řídili jsme se pravidlem plné formantové struktury, jehož snahou je zachytit vokál v co „nejčistší“ podobě
2. přidání realizovaných nefonémických jevů (realizace tvrdého hlasového začátku slov začínajících na samohlásku – ráz, realizace začátku slov za pomoci schwa, hezitační zvuky)
3. vyznačení hlasových překryvů ({OLP}) jakožto nepoužitelné části materiálu, uzpůsobení ostatních vrstev TG tomuto faktu
4. vyznačení všech disfluencí ({DSFL}) a zkreslení signálu ({SSC}), které by mohly ovlivňovat zkoumané parametry

Výstupem celého tohoto časově náročného procesu je soubor TG se strukturou, kterou demonstruje Obr. 3.



Obr. 3: Ukázka struktury Textgridu během přípravy materiálu.

V tomto stadiu jsme přikročili k rozřezání materiálu na části určené nádechovými úseky. Samotné sekvence s pouhými nádechy jsme odstranili. Úseky, které byly kratší než 1,2 sekundy, jsme po vzoru Volína et al (2017) sloučili s jedním z přilehlých úseků, volili jsme vždy ten kratší.

Rovněž jsme z materiálu vyřadili výrazné překryvy, tedy takové, kde skutečně několik sekund mluvili dva či všichni tři lidé dohromady, části s ojedinělými přerušeními jsme ponechali. Po tomto procesu jsme získali dohromady 307 nádechových úseků s distribucí vzhledem k mluvčím, kterou demonstruje následující tabulka (Tab. 2).

mluvčí	počet položek
moderator	65
sp1	137
sp2	105

Tab. 2: Rozdělení 307 nádechových úseků mezi zúčastněné mluvčí.

Pro každý z úseků jsme nechali vytvořit skriptem objekt pitch-tier, který sám o sobě bude sloužit pro analýzu intonačního rozpětí i parametru kumulativní strmosti. Tyto objekty pak musely projít manuální kontrolou, aby se odstranily chyby způsobené extrakcí f_0 ze signálu (oktávové skoky v důsledku rázů, třepené fonace apod.).

V tomto momentě byl materiál po technické stránce připraven k analýze intonačního rozpětí i kumulativní strmosti, ještě předtím však musela být každému nádechovému úseku přidělena hodnota afektivní valence.

5.2. Problematika subjektivity při určování afektivní valence

Úkol určování afektivních valencí byl v rámci celé práce tím nejohroženějším, jelikož je doprovázen velkou mírou subjektivity. Tu by bylo možné eliminovat pouze rozsáhlým percepčním testem, což však by představovalo samostatný výzkum hodný celé bakalářské práce. Problematika se subjektivitou v otázce určování valence se nám jinak jeví v současné situaci jako neřešitelná, jelikož objektivní metody dost dobře dosáhnout nelze. Dobrou zprávou je pro nás nicméně zjištění Banse & Scherera (1996), kteří došli k závěrům, že při hodnocení nahraných afektů mají hodnotící posluchači mnohem častěji tendenci zaměňovat různé pozitivně laděné stavy mezi sebou, než aby je nesprávně považovali za negativní. Ačkoli takovéto zjištění musí být bráno s rezervou (problém uměle nahraných emocí jsme diskutovali v kapitole 3.3 o dosavadních výzkumech), dostává se nám skrze něj do rukou argument, který do jisté míry naši metodiku ospravedlňuje.

Problém jsme dále řešili připuštěním určité míry subjektivity jako nevyhnutelné součásti výzkumu takového ražení. Některé sporné či nejasné případy však byly pro snížení nejistoty konzultovány a diskutovány s další kompetentní osobou.¹

Zásadou, kterou jsme se při tom snažili dodržovat, byla snaha o konzistenci hodnocení na úrovni jednotlivých mluvčích, což žádalo pečlivé seznámení s charakteristikami projevu jednotlivých mluvčích a zejména kontextu replik (na co jednotliví mluvčí reagují). K tomu spolehlivě posloužil čas strávený přípravou materiálu pro analýzu, obě činnosti provedla autorka práce. Dá se proto spoléhat alespoň na tato opatření, ačkoli samozřejmě zůstává k diskusi, zda je takovýto postup ideální a jak se případně vyhnout tímto způsobem zkresleným výsledkům.

S vědomím tohoto problému byla určena afektivní valence u každého z 307 položek. Jak již bylo uvedeno v předchozí části, pracovali jsme při určování se škálou o čtyřech bodech – dvou pro základní valenci (tj. pozitivní a negativní) a dvou pro výraznější či extrémnější ladění.

Hodnoty jsme v materiálu značili písmenky pro určenou valenci (p/n), úseky s extrémnějším laděním byly k tomu doplněny znaménky – pro pozitivní valenci „+“ a pro negativní „-“. Výrazně pozitivní část promluvy bude tedy označena jako **p+**, méně výrazně pozitivní jako **p**, výrazně negativní jako **n-** a méně výrazně negativní jako **n**.

Distribuci přiřazených valencí k úsekům jednotlivých mluvčích demonstruje následující tabulka (Tab. 3).

	p+	p	n	n-
moderator	0	47	17	1
sp1	2	73	52	10
sp2	0	23	52	30

Tab. 3: Distribuce valenčních hodnot u mluvčích.

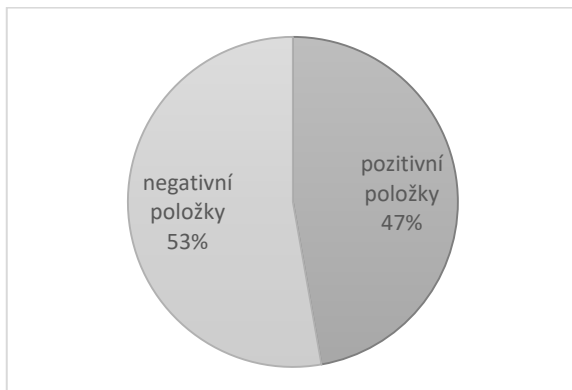
Z tabulky č. 3 je zřejmé, že se v materiálu téměř vůbec nevyskytovaly položky s výrazně pozitivní valencí², proto nebyly pro analýzu použity, stejně jako pouze jediná (a tudíž zcela neprůkazná) výrazně negativní položka u moderátora.

Je zřejmé, že problém malého vzorku se týká i dalších kategorií (n u moderátora, n- u sp1) a na tento fakt je nutné myslet při interpretaci výsledků.

¹ Mgr. Pavel Šturm, Ph.D., vedoucí práce

² Takto byly vyhodnoceny pouze dva za sebou následující úseky, kdy se mluvčí sp1 upřímně zasmál vlastnímu přerěknutí.

Spíše pro zajímavost uvedme, že dříve vyjádřený předpoklad o převažování negativního ladění celého materiálu nebyl přesvědčivě potvrzen. Ačkoli v jednotkách negativní úseky spíše převažují, jak je zřejmé z následujícího obrázku (Obr. 4), provedený binomický test statisticky významnou odlišnost mezi počty nepotvrdil.



Obr. 4: V debatě mírně převažovaly položky označené jako negativní, významnost počtu se ale nepotvrdila.

5.3. Zvolené parametry intonační variability a jejich analýza

5.3.1. Hlasová poloha mluvčích

Pro lepší představu o mluvčích a jejich práci s hlasem jsme se rozhodli k dosud představeným parametrům intonační variability přidat ještě informaci o změně polohy hlasu každého z mluvčích v závislosti na afektivní valenci. Tu zjistíme za pomoci výpočtu mediánů f_0 pro každou z uvedených valencí.

5.3.2. Intonační rozpětí

Pro analýzu intonačního rozpětí jsme se řídili postupem, který je popsán v článku Volín et al. (2017) – melodickou konturu (průběh f_0) jsme extrahovali v programu Praat skriptem, který zároveň výslednou konturu vyhladil (stylizoval – více v kap. 2.1.1) 10 Hz filtrem a neznělé úseky mezi znělými body proložil interpolační křivkou. Tento skript byl laskavě poskytnut jeho autorem doc. J. Volínem. Pro analýzu dat bylo dále potřeba v následujícím kroku manuálně a pomocí sluchové percepce zarovnat chybně vyznačené body na intonační křivce (více v sekci 5.1.2.).

Jak jsme se dozvěděli z další literatury, měření intonačního rozpětí není tak přímočarou záležitostí, jak by se mohlo na první pohled zdát. Se statistického pohledu totiž není úplně šikovné spoléhat se na krajní (tudíž extrémní a v každém úseku pouze jednou zastoupené) hodnoty.

Prostor pro uvážení nejlepšího způsobu měření intonačního rozpětí nám poskytují Volín J. & Bartůňková, H. (2015), kteří své metody rozvádějí a vysvětlují nevýhodu tohoto nejjednoduššího zavedeného parametru. Ve svém článku pracují se třemi typy měření rozpětí, která mají potenciál zachytit míru frekvenční variability:

- I. *min-max range* – neboli variační rozpětí (Volín, 2007) je již zmíněným rozdílem mezi nejvyšší a nejnižší naměřenou frekvencí v konkrétním úseku. Výstupem variačního rozpětí pro každý nádechový úsek budou dvě číselné hodnoty – jedna odpovídající nejnižší naměřené hodnotě a jedna pro hodnotu nejvyšší. Autoři toto měření považují za riskantní, náchylné k chybám a málo informativní, od jeho použití tedy spíš ustupují a jeho výsledky ve studii ani nepublikují. Volín (2007) dokonce uvádí, že je tento parametr většinou intonologů odmítán, jelikož ani uživatelé jazyka při odhadu intonačního rozpětí extrémny nevyužívají. Dále k tomuto parametru píše, že by měl být využíván co nejméně a pouze v případech, kdy si to situace vyloženě žádá.
- II. *percentil range* – (percentilové rozpětí) spočívá v seřazení všech naměřených hodnot a následném „odřiznutí“ určitého procenta extrémních hodnot z obou krajů. V citovaném výzkumu se pro určení rozpětí využívá 10. a 90. percentil. Volín (2007) hovoří o tom, že toto rozpětí je v intonologických výzkumech výhodné, jelikož při percepci se posluchači příliš nesoustředí na samotné extrémní hodnoty, ale na krajní hodnoty, které jsou v konkrétním perceptu více frekventované.
- III. *Inter-quartile range* – je vlastně typem percentilového rozpětí, které využívá pouze středních 50 % hodnot, nejvyšších a nejnižších 25 % hodnot (první a čtvrtý kvartil) se zbavujeme.

Autoři následně výsledky prezentují pouze grafem s percentilovým rozpětím. Tuto skutečnost odůvodňují tím, že grafy mezikvartilního rozpětí a standardní odchylky (kterou rovněž v této studii měřili) vycházely s výše definovaným mezikvartilním rozpětím téměř identicky. *Min-max range* neprezentují vůbec, znovu zdůrazňují jeho nespolehlivost, která plyne ze závislosti na extrémních hodnotách, ačkoli jde o ukazatel zdánlivě nejjednodušší a nejpřímochařejší. My se budeme v této práci držet jejich metodiky a budeme proto měřit rozpětí percentilové, přesněji mezi 10. a 90. percentilem.

Pro každého mluvčího budeme intonační rozpětí měřit zvlášť a průměry všech typů valencí vyneseme do krabicového grafu.

5.3.3. Parametr kumulativní strmosti

Dalším ze zkoumaných parametrů v této práci je takzvaný ukazatel/parametr kumulativní strmosti (původně *cumulative steepness*, později *cumulative slope index*, *CS*). Tento parametr zavedl pro účely své bakalářské práce Robin Hruška (2016). Na základě jeho práce byl o rok později autorskou dvojicí Hruška & Bořil (2017) vydán článek v časopise *AUC Philologica* (3/2017), kde byl parametr poprvé prezentován v kontextu vědecké literatury. Jedná se tedy o termín ve fonetickém světě vesměs nový.

Tento parametr byl navržen k souhrnné kvantifikaci celkové variability melodické kontury a jeho funkci Hruška a Bořil (2017) komentují následovně:

„Navrhli jsme ukazatel kumulativní strmosti, tzv. cumulative slope (CS) index, který jedním souhrnným číslem vyhodnocuje míru proměnlivosti kontury základní frekvence řečového signálu (f_0) vzhledem k celkovému trvání analyzovaného úseku.“

Ukazatel kumulativní strmosti tak můžeme označit za jakýsi index obecné míry intonační variability, jehož hodnota ukazuje proměnlivost kontury v půltónech (důvody měření intonační variability v půltónech uvádíme v kap. 2.1.2) za sekundu. Tento ukazatel je definován následující rovnicí:

$$CumSlope = \frac{1}{T_{tot}} \sum_{n=2}^N |f(n) - f(n-1)|$$

Kde:

T_{tot} = celkové trvání hlasové aktivity (v sekundách)

N = počet diskrétních bodů na intonační křivce

$f(n)$ = frekvence v semitónech n -tého bodu

Z představené definice vyplývá, že čím vyšší hodnota kumulativní strmosti bude pro daný úsek vypočítána, tím by měla být jeho intonační křivka proměnlivější (vykazovat více vrcholků a sedel, ale i větší intonační rozpětí) a řeč v tomto ohledu variabilnější.

5.3.4. Analýza počtu melodémů v nádechovém úseku

Analýza počtu melodémů v nádechovém úseku byla u všech nádechových úseků provedena percepčně. To je v současné situaci samozřejmě jediný možný způsob, jak ji provést. Celou analýzu prováděla autorka práce, některé problematické části byly s odstupem času znovu kontrolovány.

Z důvodu subjektivity není možné vyloučit jistou chybovost, jelikož vzhledem k množství materiálu nebylo reálné konzultovat každý nejednoznačný případ s další kompetentní osobou. Počet identifikovaných melodémů by měl tedy alespoň přibližně zastupovat počet intonačních frází v nádechovém úseku.

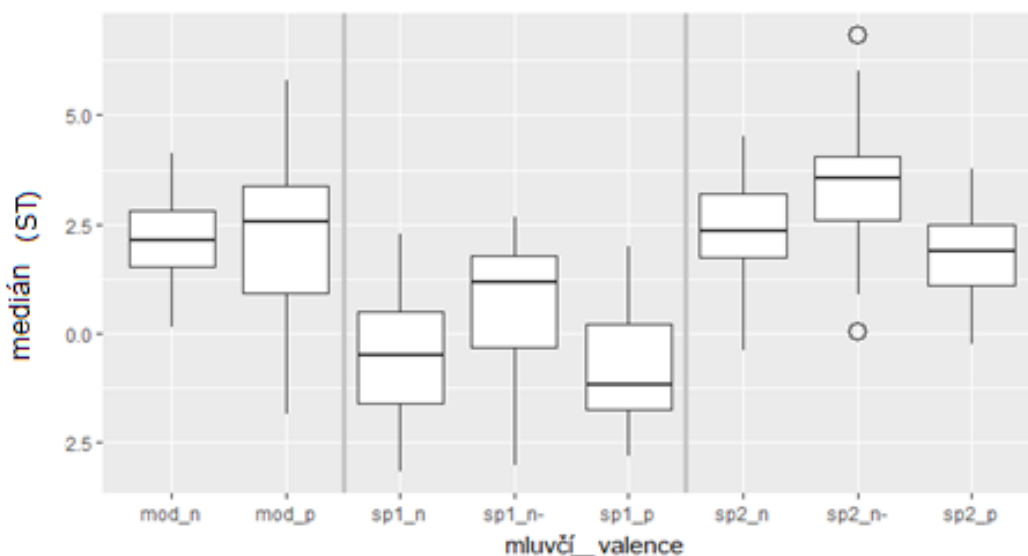
5.4. Statistické testování výsledků

Rozdělení všech naměřených dat se pohybovala v rámci logaritmicko-normálního rozdělení, proto byl pro určení statistické významnosti u všech zkoumaných parametrů použit dvoustranný nepárový t-test s hladinou významnosti $\alpha = 0,05$. Pro mluvčí, u kterých byly zastoupeny tři valenční kategorie, jsme tuto hladinu korigovali Bonferroniho korekcí, která pro tři dílčí testy znamená hladinu $\alpha = 0,016667$.

6. Výsledky

6.1. Ukazatel hlasové polohy

Zavedení tohoto parametru plyne z předpokladu, že možná existuje rozdíl v průměrné hlasové poloze v závislosti na afektivní valenci. Výsledky byly získány naměřením mediánů f_0 v úsecích stejného valenčního označení pro každého mluvčího.

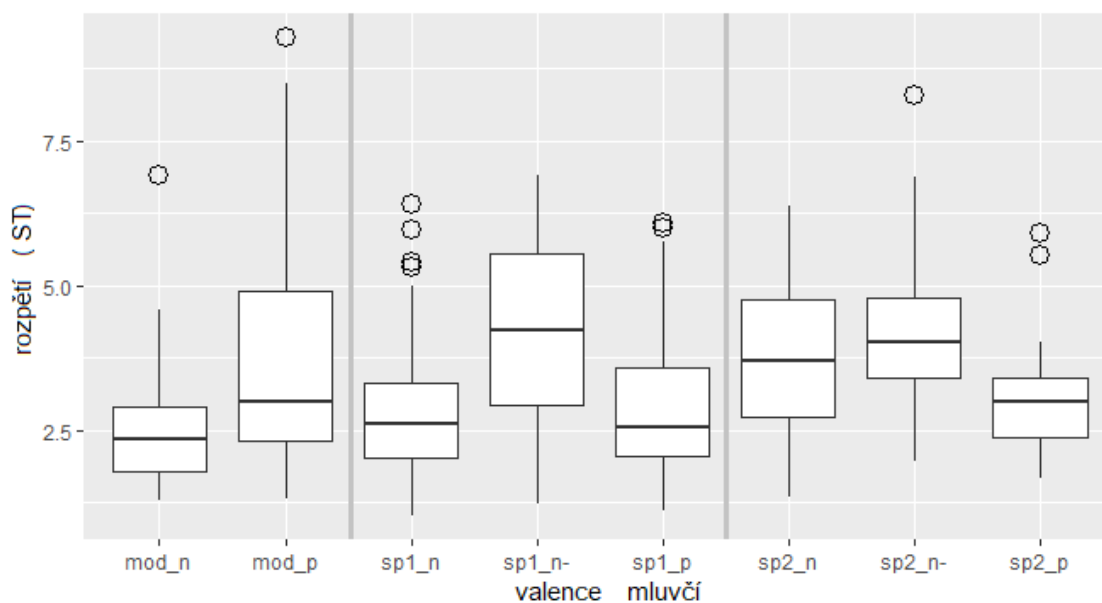


Obr. 5: Mediány f_0 všech mluvčích v závislosti na afektivní valenci.

Z krabicového grafu (Obr. 5) je zřejmé, že mediány f_0 mají u konkrétních mluvčích určité tendence. Zatímco u mluvčího mod (moderátor Jan Punčochář) je tato tendence spíše nevýznamná (statistická významnost nebyla t-testem potvrzena), u dalších dvou (sp1 a sp2) je možné vidět v hladinách f_0 pro odlišné valence jisté rozdíly. Zajímavá je především tendence, kdy oba mluvčí dosahují nejvyšší průměrné frekvence u excitovaných promluv. Statistická významnost však byla potvrzena pouze u mluvčího sp2, a to jak pro pár negativní – výrazně negativní ($t(51)=2$, $p = 0,007$), tak pro pár pozitivní – výrazně negativní ($t(51)=2$, $p = 0,000015$), z čehož vyplývá, že afektivní valence může mít na celkovou polohu hlasu v některých případech určitý vliv. Rozdíl je v obou případech vyšší než 1 ST, jedná se tedy i o rozdíl foneticky významný.

6.2. Percentilové intonační rozpětí

Pro analýzu percentilového rozpětí jsme zvolili percentilové rozpětí mezi 10. a 90. percentilem, které dále budeme označovat pouze jako „intonační rozpětí“. Rozdíly mezi naměřenými hodnotami (medián, prostor výskytu 95 % dat a odlehlé hodnoty) vynášíme do jednoho krabicového grafu pro všechny mluvčí. Statistické testování bylo vzhledem k povaze dat provedeno stejným způsobem jako u předchozího parametru.



Obr. 6: Percentilové intonační rozpětí v závislosti na afektivní valenci pro všechny mluvčí.

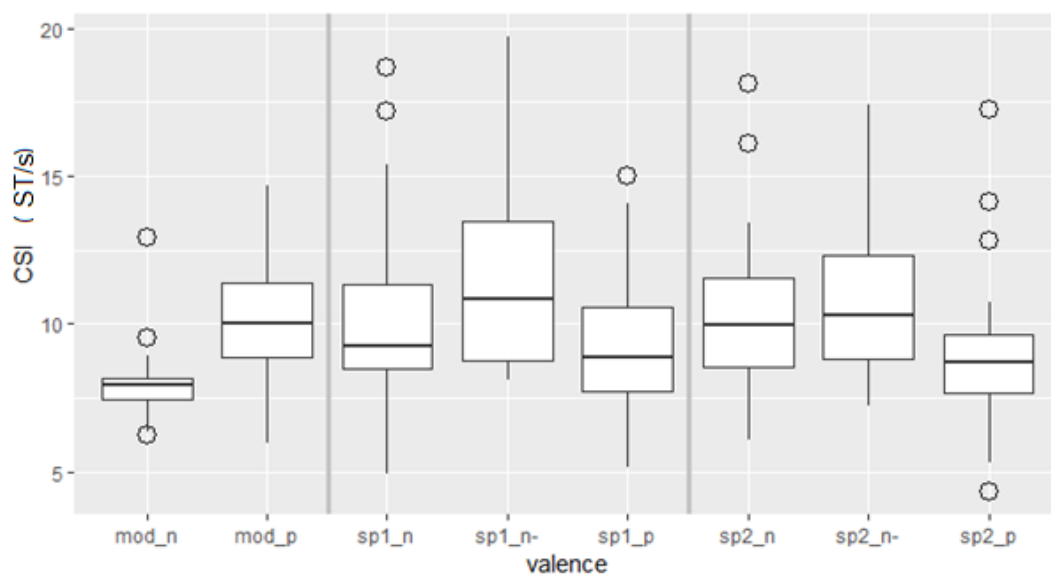
Krabicový graf (Obr. 6) zobrazuje naměřené hodnoty v krabicovém grafu pro moderátora Jana Punčocháře, u kterého byly reprezentativně zastoupeny pouze základní hodnoty negativní (n) a pozitivní (p), jejichž počty byly značně nevyvážené (17:47). Extrémněji se u jeho promluv projevila afektivní valence pouze v jednom případě, který byl tudíž vyřazen. Graf naznačuje, že mediánem určená šíře intonačního rozpětí u tohoto mluvčího byla o více než jeden půltón širší u pozitivně označených částí jeho promluvy než u těch negativních. Vzhledem k absenci extrémních hodnot jsme statistickou významnost testovali pouze t-testem, který předpokládanou významnost rozdílu potvrdil ($t(45)=2,01$, $p = 0,016$). Percepční významnost byla následně potvrzena poslechem některých částí a na povrch vyšlo zajímavé pozorování, které odhalilo mezi pozitivně laděnými částmi poměrně velké množství intonačně výrazně realizovaných otázek, což by mohlo být faktorem, který tento fakt mohl ovlivnit.

Poněkud odlišné tendence můžeme pozorovat u naměřených hodnot intonačního rozpětí promluvího Petra Gandaloviče (sp1). Počet extrémně negativních promluv nám umožnil na rozdíl od předchozího mluvčího zahrnout i promluvy označené jako n-, ačkoli by se dalo hovořit o tom, zda je jejich počet (10) dostatečně reprezentativní. Samotný rozdíl mezi pozitivními a negativními promluvami se u tohoto mluvčího na intonačním rozpětí projevil v mnohem menší míře než u předchozího mluvčího. Medián zvoleného percentilového rozpětí výrazně negativních promluv (n-) u P. Gandaloviče dosáhl o více než půltón širšího rozpětí v kontrastu se základními pozitivními i negativními promluvami. Provedený t-test po Bonferroniho korekci však statistickou významnost nepotvrdil u žádné z dvojic negativní – pozitivní, výrazně negativní – pozitivní, negativní – výrazně negativní.

Nakonec okomentujme výsledky závislosti intonačního rozpětí na afektivním ladění druhého hosta pořadu, Michala Haška. Tento mluvčí sice ve shodě s mluvčím sp1 vykazuje nejširší rozpětí u průměrů extrémně negativních úseků (n-), rozdíl vůči základním negativním hodnotám (n) je ale v tomto případě výrazně menší a t-test s Bonferroniho korekcí významný rozdíl rovněž nepotvrdil. Naopak střední hodnota (medián) rozpětí pozitivně (p) hodnocených úseků je u tohoto mluvčího o méně než tři čtvrtiny nižší, než střední hodnota rozpětí úseků negativních (n). V tomto případě stejný test významnost potvrdil ($t(53)=2, p = 0,017$). Je ale zřejmé, že tento výsledek stojí přímo v opozici proti též statisticky významnému výsledku u mluvčího Jana Punčocháře – rozpětí je v případě sp2 širší pro negativní úseky než pro pozitivní. Se stejnou tendencí byl jako významný potvrzen rozdíl u dvojice pozitivní – výrazně negativní ($t(51)=2, p = 0,001$).

6.3. Parametr kumulativní strmosti

Parametr kumulativní strmosti jsme vypočítali pomocí skriptu poskytnutého FÚ FFUK. Jednotkou tohoto parametru jsou púlťóny za sekundu, tedy ST/s. Výsledky pro každého mluvčího prezentujeme v krabicovém grafu. Vzhledem k povaze dat bylo testování statistické významnosti ošetřeno stejnými prostředky jako u intonačního rozpětí.



Obr 7: Hodnoty kumulativní strmosti (CSI) v závislosti na afektivní valenci pro všechny mluvčí.

Z grafu (Obr. 7) je zřejmý významný rozdíl v hodnotě CSI reprezentovaných valenčních hodnot pro mluvčího Jana Punčocháře (mod), který naznačuje výrazný rozdíl mezi celkovou intonační variabilitou negativně a pozitivně ohodnocených úseků u tohoto mluvčího, přičemž pozitivní vykazují variabilitu podstatně vyšší. Významnost rozdílu byla potvrzena t-testem ($t(37)=2$, $p = 0,00006$). Tento výsledek podle očekávání koreluje s výsledkem stejného mluvčího pro intonační rozpětí.

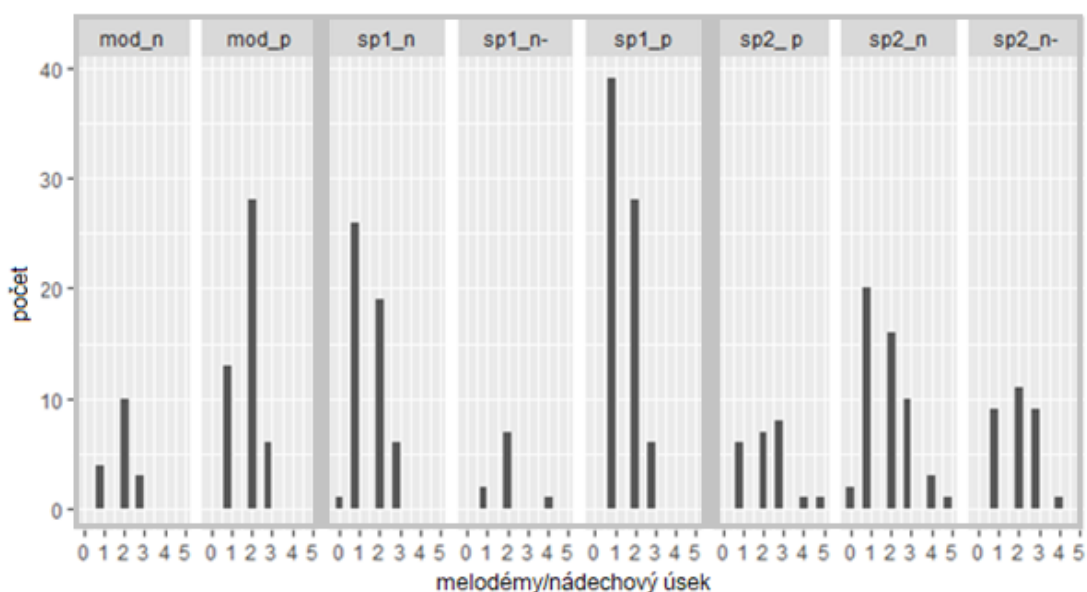
Opět jinak se projevoval mluvčí Petr Gandalovič (sp1). Ačkoli můžeme pozorovat podobný vzorec jako u intonačního rozpětí stejného mluvčího (tedy největší intonační variabilitu pro úseky n- a nejnižší pro úseky p, stejně jako u intonačního rozpětí), v tomto případě nebyl významný statistický rozdíl zvoleným testovacím procesem potvrzen ani pro jednu z dvojic.

U mluvčího Michala Haška (sp2) je v grafu možné pozorovat nejnižší variabilitu u pozitivních úseků a nejvyšší u výrazně negativních úseků, ale provedený test statistické významnosti se všemi

náležitostmi (podobně jako u předchozího mluvčího) neprokázal signifikantní rozdíl u žádné z valenčních dvojic.

6.4. Počet melodémů v nádechovém úseku

Dále jsme provedli analýzu počtu melodémů v každém nádechovém úseku. Distribuci melodémů pro všechny mluvčí ukazuje histogram (Obr. 8)

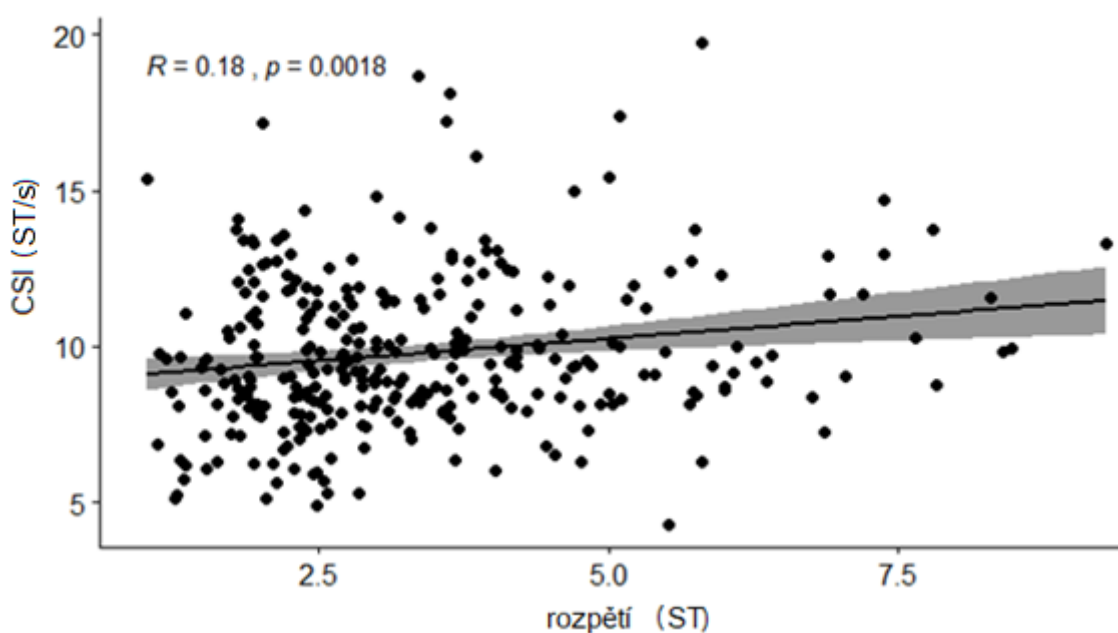


Obr. 8: Distribuce počtů melodémů v nádechovém úseku pro jednotlivé mluvčí a afektivní valence.

Distribuce počtů melodémů v nádechovém úseku (která měla reprezentovat přibližný počet prozodických frází) se v tomto výzkumu neprojevila jako proměnná závislá na hodnotě afektivní valence. Jak je vidět na histogramu (Obr. 8), zdá se, že souvislost mezi těmi to dvěma proměnnými na zkoumaných datech neexistuje. Provedené testy statistické významnosti (opět jsme použili t-test s případnou korekturou) signifikantní rozdíly nepotvrdily ani v jediném z případů.

6.5. Korelace mezi intonačním rozpětím a CSI

Vzhledem k výsledkům intonačního rozpětí a parametru kumulativní strmosti bylo nasnadě očekávat vysokou korelaci mezi těmito dvěma parametry. Podrobili jsme je korelačnímu testu a vynesli do korelačního grafu (Obr. 9). Výsledný Pearsonův korelační koeficient ($r = 0,18$) však indikoval pouze velmi nízkou pozitivní korelaci. Z toho vyplývá, že pro výzkum intonační variability má zřejmě smysl podrobovat data oběma těmito měřeními.



Obr. 9: Korelační graf pro parametr kumulativní strmosti a intonační rozpětí.

7. Diskuse

Cílem této práce bylo pokusit se za pomoci několika parametrů přispět ke zjištění, zda je možné mapovat změny intonační variability v závislosti na afektivní valenci aktuálního ladění mluvčího. Výstupem praktické části jsou měření tří zadaných parametrů, které jsme doplnili ukazatelem hlasové polohy jakožto další hlavní složky intonačního popisu.

Ukazatel hlasové polohy odhalil u obou hostů (sp1 a sp2) tendenci zvýšení hlasové polohy u excitovanějšího z měřených stavů, což koresponduje s předchozími zjištěními, která přinesli například Banse & Scherer (2003) či Laukka et al. (2005).

Výsledky měření intonačního rozpětí pro každého mluvčího přináší poměrně rozporuplná zjištění, která však do jisté míry potvrzují dílčí hypotézy, které jsme představili v kapitole 4. U mluvčího Jana Punčocháře (moderátor) jsme pracovali pouze s hodnotami **n** a **p**, jelikož extrémnější nebyly v dostatečné míře zastoupeny (tato a jiné problematiky výzkumu budou diskutovány níže). Fakt, že se intonační rozpětí u tohoto mluvčího pohybovalo v širším pásmu u pozitivně laděných promluv než u negativně laděných, podporuje teorii, za kterou stojí například zmiňovaní Reckling & Kügler (2011) či Scherer & Oshinsky (1977). Přímo proti ní ale stojí hodnoty naměřené u mluvčího Michala Haška (sp2), které (s ověřenou statistickou významností) ukazují vyšší rozpětí naopak pro negativně hodnocené promluvy. Je tedy namístě přemýšlet nad možností, zda neexistují při řeči různé strategie využívání intonačního rozpětí.

Další z původních hypotéz, tedy že intonační rozpětí spíše ovlivňuje celková míra excitace než její příslušnost k jedné ze dvou stran valenční osy (jak nadnesli například Cruttenden, 1986; Palková 1997; Laukka, 2005), se nepodařilo prokázat. Ačkoli data naměřená u mluvčího P. Gandaloviče jí zdánlivě nasvědčovala – rozdíl nedosáhl statistické významnosti. Tuto teorii navíc nebylo možné ověřit na páru výrazně pozitivní a negativní, jelikož s výrazně pozitivními hodnoceními jsme se v rámci celého materiálu téměř nesetkali, což považujeme za nejslabší místo potenciálu politické debaty pro účely zkoumání afektivních stavů, jak uvedeme v části věnované problémům tohoto výzkumu.

Aplikování parametru kumulativní strmosti (CSI), jenž představuje výpočet komplexní intonační variability, přineslo rovněž zajímavé výsledky, které mohly být na první pohled podobné výsledkům intonačního rozpětí. Následně provedený korelační test ale významnou korelaci nepotvrdil. Tento fakt ovšem přináší zjištění, že je vskutku namístě do výzkumů zahrnovat oba tyto parametry.

Ověření významného rozdílu (nejen z hlediska statistického, ale i lingvistického) jsme zaznamenali pouze u mluvčího Jana Punčocháře, kde se projevíly pozitivně laděné úseky jako variabilnější než

ty negativní. U mluvčího Petra Gandaloviče se jako nejvariabilnější projevíly výrazně negativní úseky, srovnání s ostatními skupinami ale neprokázalo statisticky významný rozdíl. Tendence podobného typu se projevovaly i u mluvčího Michala Haška.

Za povšimnutí stojí rozdíl v tendencích CSI i intonačního rozpětí vzhledem k pozicím zúčastněných mluvčích v rámci politické debaty. Vzhledem k opačným tendencím moderátora a obou hostů bychom mohli uvažovat nad tím, zda tato realita nehraje ve způsobu užívání prozodických prostředků roli. Vzhledem k absenci výrazně negativních promluv u moderátora však není možné na tuto teorii dále navázat a bylo by potřeba ji ověřit na dalším materiálu. Po opětovném poslechu materiálu ale můžeme říci, že vyšší variabilita (vč. rozpětí) mohla být u moderátora způsobena množstvím otázek (realizovaných poměrně výraznou intonací), které byly pokládány především v přívětivém stylu, tudíž byly hodnoceny jako pozitivní. Negativně hodnocené části byly naopak poslechově plošší, byla z nich slyšet snaha o nadhled a jakousi profesionální strohost či suchost při oponování hostům. U obou hostů se podle poslechu intonačně výrazněji projevovaly úseky označené jako extrémně negativní. Často šlo o percepčně výrazná stoupání (signalizující intenzivní podráždění, vyvedení z míry). Promluvy těchto mluvčích hodnocené jako základně pozitivní či negativní by se percepčně daly popsat jako intonačně nevýrazné a tento dojem se kvantitativním výsledkům nevymyká. Oba hosté při vyjadřování jiné než excitované valence zní velice klidně, přičemž můžeme vzpomenout na citovaný výzkum Rodero (2011), podle kterého je pro percept klidné promluvy nízká variabilita celkové kontury určujícím faktorem.

Ze zkoumaných parametrů se jako nejméně závislý na afektivní valenci v našem výzkumu projevil počet melodémů (tedy intonačních frází) v nádechovém úseku. Ani u jednoho z mluvčích nebyla prokázána žádná tendence, která by o vlivu valenčního ladění na tento parametr poukazovala. Než se ale souvislost kategoricky vyloučí, měl by proběhnout detailnější výzkum – zrádností politické debaty je v tomto ohledu i fakt, že mluvčí hovoří pod tlakem vyhrazeného času a jsou často přerušováni, čili člení svou promluvu do nádechových úseků nepřírozeně a častou jsou nuceni se vyjadřovat rozkouskovaně (Čmejková, 2013).

Pozornost diskuse zasluhují i četné metodické obtíže, které se při procesu práce projevíly. Jedním z nejzávažnějších problémů je z pohledu autorky samotné určování afektivní valence ve spontánním projevu. Ačkoli byla provedena pečlivě a dodatečně kontrolována, není při prezentovaném způsobu jejího určování nikdy možné zaručit správné zařazení. Důvodem je subjektivita vnímání valence v řeči. Pokud je afektivní valence prezentována jako nezávislá proměnná, od které se odvíjí celý výzkum, máme za to, že by měla pro tyto účely být určena spolehlivěji. Je sice pravda, že afektivní stavy jsou tématem, u kterého je subjektivita nevyhnutelná,

jelikož do jisté míry vyplývá ze samotné jejich definice, domníváme se však, že by bylo vhodnější nespolehat se na určení jedinou osobou. Jako řešení navrhuje buď samostatný percepční výzkum pro celý objem dat, který by ukázal obecné tendence v posuzování mnoha lidmi, nebo alespoň využití pomoci několika asistentů. Při přistoupení k percepčnímu testu se však jeden problém zdánlivě řeší, nicméně další vyvstává. I takové množství dat, které jsme představili v tomto výzkumu a které není z pohledu fonetických výzkumů nijak nadprůměrné, by znamenalo pro posluchače velice dlouhý a úmorný proces hodnocení. Pro budoucí práce, které by se chtěly tímto tématem zabývat, zvážení alespoň nějakého opatření pro objektivizaci určování afektivní valence výrazně doporučujeme.

Samotné množství zkoumaných dat (a zejména jejich distribuce) v této práci je také poněkud neuspokojivé – některé kategorie u konkrétních mluvčích čítaly příliš málo hodnot na to, aby mohly být vůbec použity, jiné jsme naopak použili i s vědomím, že množství pro reprezentativní výsledek není zcela dostačující. Při současných zjištěních by ale bylo zajímavé tento výzkum dále rozvést a jeho výsledky ověřit na větším množství dat. Z našich výsledků totiž vyplývá, že průměrná hodnota f_0 , intonační rozpětí i hodnota vyčíslená parametrem kumulativní strmosti může na afektivní valenci záviset. Některé projevované tendence se však nepodařilo statisticky potvrdit, což by ale nemělo být signálem, že již není potřeba výzkum melodické variability v závislosti na afektivní valenci dále rozvádět, ba naopak. Toto téma je jistě nosné a zaslouží další pozornost, jejíž součástí by měla být možnost zkoumat rozsáhlejší vzorek dat.

V posledním bodě diskuse bychom rádi zhodnotili potenciál využití politické debaty jako materiálu pro výzkumy afektivních stavů. Nespornou výhodou tohoto materiálu je jeho zvuková kvalita, která umožňuje poměrně precizní zpracování v porovnání s ostatními dostupnými zdroji spontánního hovoru. Je však nutno přiznat, že spontaneita může být v tomto typu materiálu poněkud omezena – hosté se mohou snažit se „hlídat“, bez dostupné obrazové složky také může být nesehnáno odfiltrovat části, kde například moderátor své repliky čte a podobně. Dále zmiňme již výše uvedený problém skákání do řeči a „rozkouskovanosti“, který politickou debatu diskvalifikuje pro zkoumání parametru, jako je počet melodémů v nádechovém úseku.

Je ale namístě podotknout, že i přes tyto nástrahy je politická debata vhodným materiálem pro výzkum afektivních stavů, jelikož při jejím konceptu se nedá vyvarovat vyhroceným situacím, které nám mohou skrze kvalitní zvukový záznam zprostředkovat mnoho informací o afektivních stavech. Z naší zkušenosti vyplývá, že je nutné počítat s poněkud jednostrannější distribucí dat, poněvadž se v použitém materiálu neosvědčilo očekávat výrazně pozitivní projevy. Nelze samozřejmě vyloučit existenci politické debaty, která bude takové úseky obsahovat hojněji,

z představení principu politických debat (sekce 3.5 této práce) je ale zřejmé, že není cílem tohoto typu pořadu podobné stavy vyvolávat, ba naopak.

Pro zakončení práce shrňme nejdůležitější závěry, které z ní vyplynuly:

V práci nebyla možnost pracovat se zcela uspokojivým objemem dat, pokud však předpokládáme správnost použité metodiky a přistoupíme k představené studii spíš jako k pilotnímu projektu, pak se z jejích výsledků zdá, **že afektivní valence může mít vliv na polohu hlasu, intonační rozpětí i celkovou intonační variabilitu** (reprezentující CSI). Představená studie naopak nepotvrdila souvislost mezi afektivní valencí a počtem melodémů (tedy intonačních frází) v nádechovém úseku. Získané výsledky však nedovolují přiklonit se k žádné z představených hypotéz, ačkoli dílčí tendence některým z nich nasvědčují. Je tedy namístě pro další zkoumání na tomto poli ověřit hypotézu novou, kterou je možné při současných výsledcích nastolit – totiž že existují určité **strategie**, které individuální mluvčí využívají ve svém projevu, a zda je nějaká z nich pro některou z rolí (moderátor vs. host) typická, například vyšší variabilita u pozitivních promluv u moderátora, tak jako u moderátora z této studie.

Pokud by se takovou hypotézu podařilo potvrdit, bylo by dále zajímavé zkoumat, zda je například některá s případných strategií pro posluchače přijatelnější. Také jsme zjistili, že politická debata může být v určitých případech pro své vlastnosti velice žádoucím materiálem pro výzkum projevů afektivních stavů v řeči, je ale nutné dopředu počítat s výše popsány limity, které se s jejím použitím vážou, proto rozhodně není jednoznačně vhodná pro všechny typy výzkumů.

Na úplný závěr uvedme, že tato práce svým zpracováním mnohem více otázek otevřela, než kolik jich vyřešila. Ukázalo se, že práce se spontánní řečí v souvislosti s afektivní prozodíí je v porovnání se zkoumáním studiově nahraných afektů o dost náročnější disciplínou, která svými překážkami může vědce do určité míry odrazovat.

Výzkum přirozených projevů afektivních stavů v řeči vyžaduje mnoho úsilí a výsledky nemusí být vždy jednoznačné. Věříme ale tomu, že námaha spojená s výzkumem afektivní prozodie spontánní řeči je námahou vynaloženou správným směrem. Projevování afektů je totiž tím, co zůstává až do současnosti technologiemi nenapodobitelnou lidskou schopností, což ji činí jedním ze zbývajících velkých tajemství řeči. A to je něco, co stojí za to zkoumat.

Seznam zdrojů

AUDIBERT, N., AUBERGÉ, V., & RILLIARD, A. (2008). How we are not equally competent for discriminating acted from spontaneous expressive speech.

BANSE R. & SCHERER K. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of personality and social psychology*. 70. 614-36. 10.1037/0022-3514.70.3.614.

BRIESEMEISTER, B. B., KUCHINKE, L., & JACOBS, A. M. (2012). Emotional Valence: A Bipolar Continuum or Two Independent Dimensions?

CAMPBELL, N. & MOKHTARI, P. (2003). Voice quality: The 4th prosodic dimension. *Proc 15th ICPhS*.

ČMEJRKOVÁ, S. (2013). *Styl mediálních dialogů*. Praha: Academia

FÓNAGY, I. (1983, a). *La vive voix*. Paris, Payot

HART, J., COLLIER, R., & COHEN, A. (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge. Cambridge University Press

HRUŠKA, R. (2016). *Vlastnosti kontur základní frekvence s ohledem na segmentální složení promluv*. Bakalářská práce. Dostupné z <https://dspace.cuni.cz/handle/20.500.11956/74263>

HRUŠKA, R. & BOŘIL, T. (2017). Temporal variability of fundamental frequency contours. *Acta Universitatis Carolinae – Philologica 3, Phonetica Pragensia XIV*, 35–44.

LAUKKA, P., JUSLIN, P. & BRESIN, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition and Emotion*. 19(5). 633-653.

MACHAČ, P. & SKARNITZL R. (2009). Fonetická segmentace hlásek. Praha: Epoque

MOZZICONACCI, S. (2002). Prosody and Emotions. Aix-en Provence, France: Speech
Prosody

NAKONEČNÝ, M. (1997). Encyklopedie obecné psychologie. Praha: Academia

NAKONEČNÝ, M. (2000). Lidské emoce. Praha: Academia

NOLAN, F. (2003). International equivalence: An experimental avaluation of pitch scales. In:
Proceedings of 15th ICPHS

NOLEN-HOEKSEMA, S. (2012). Psychologie Atkinsonové a Hilgarda. Praha: Portál, 2012

PALACIOS-ALONZO, D., BIARGE, M. V., LLUIOS, V.N., & VILDA, P.G. (2015).
Spontaneous Emotional Speech Recordings through a Cooperative Online Video Game.
CoSECivi.

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
INT-1997, 267-270.

PALKOVÁ, Z., VERONĚKOVÁ, J., VOLÍN, J. & SKARNITZL, R. (2004). Stabilizace některých
termínů pro fonetický popis češtiny v závislosti na nových výsledcích výzkumu. In: Duběda, T.
(Ed.), Konference česko-slovenské pobočky ISPhS 2004, 65–74. Praha: FF UK.

PATTERSON, D. (2000). A linguistic approach to pitch range modelling. (Unpublished doctoral dissertation). University of Edinburgh

PITTAM, J., & SCHERER, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (p. 185–197). Guilford Press.

POLLÁK, P., VOLÍN, J. & SKARNITZL, R. (2007). HMM-based phonetic segmentation in Praat environment. In: *Proceedings of the XIIth International Conference “Speech and computer – SPECOM 2007”*, pp. 537–541.

RECKLING, F. & KÜGLER, Frank. (2011). Pitch range in positive and negative connoted statements of german

RODERO E. (2011). Intonation and emotion: influence of pitch levels and contour type on creating emotions. *Journal of voice: official journal of the Voice Foundation*, 25(1), e25–e34. <https://doi.org/10.1016/j.jvoice.2010.02.002>

SAUTER, D. A. & FISCHER, A. H. (2018). Can perceivers recognise emotions from spontaneous expressions? *Cognition and Emotion*, 32:3, 504-515, DOI: 10.1080/02699931.2017.1320978

SCHERER, K. R., & OSHINSKY, J. S. (1977). Cue Utilization in Emotion Attribution from Auditory Stimuli. *Motivation and Emotion*, 1(4), 331–346.

SCHERER, K. R., JOHNSTONE, T., KLASMEYER, G., Banziger, T., (2000). Can automatic speaker verification be improved by training the algorithms on emotional speech? In: *Proc. ICSLP2000*, Beijing, China.

SCHERER, K. R. (2003). Vocal Communication of Emotion: A Review of Research Paradigms. *Speech Communication* 40, 227-256. *Speech Communication*. 40. 227-256. 10.1016/S0167-6393(02)00084-5.

SKARNITZL, R., ŠTURM P. & Volín, J. (2016). *Zvuková báze řečové komunikace: fonetický a fonologický popis řeči*. Praha: Karolinum

VLČKOVÁ-MEJVALDOVÁ, J. (2006). *Prozodie, cesta i mříž porozumění: experimentální srovnání příznakové prozodie různých jazyků*. Praha: Karolinum

VOLÍN, J., GALEONE, D. & JOHNSON, W. (2017). Temporal variability of fundamental frequency contours. *Acta Universitatis Carolinae – Philologica* 3, *Phonetica Pragensia* XIV, 35–44.
Pitch range of intonation contours in Czech English

VOLÍN, J. & BARTUŇKOVÁ, H. (2015). Assets and liabilities of simple descriptors of fundamental frequency tracks. In: Niebuhr, O. & Skarnitzl, R. (Eds.), *Tackling the Complexity in Speech*, 147–161. Praha: Filozofická fakulta, Univerzita Karlova v Praze.

VOLÍN, J. (2007) *Statistické metody ve fonetickém výzkumu*. Praha: Epoque

VOLÍN, J., HEDVÁBNÁ B., JANOUŠKOVÁ J., MACHAČ P., PALKOVÁ Z., SKARNITZL R., STUDENOVSKÝ D., VERONŇKOVÁ J., (2007). Poznání za mřížemi. *Naše řeč*, 90(5), 261 - 266

Použitý software:

R Core Team. (2018). *R: A language and environment for statistical computing* (Verze 3.2.4). Vienna: R Foundation for Statistical Computing. Získáno z <http://www.R-project.org/>

Boersma, P., & Weenink, D. (2020). *Praat - doing phonetics by computer* (Verze 6.1.09). Získáno z <http://www.praat.org>