

Univerzita Karlova v Praze  
Filozofická fakulta

Fonetický ústav

*Studijní program Filologie  
Studijní obor Fonetika*

*Disertační práce*

Lenka Weingartová

**Identifikace mluvčího v temporální doméně řeči**

**Speaker identification in the temporal domain of speech**

školitel: doc. PhDr. Jan Volín, Ph.D.

2015

**Prohlášení**

Prohlašuji, že jsem disertační práci napsala samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 1. 4. 2015

Podpis: .....

## **Poděkování**

Tato práce by nemohla vzniknout bez zásadní podpory mnoha osob. Nejdůležitější z nich je školitel, Jan Volín, který mi byl kromě nesmírně širokých znalostí a zkušeností také neustálým zdrojem motivace a inspirace. Vštípil mi nejen lásku k fonetické práci a vědeckému bádání vůbec, ale také do mě vtloukl zásady vědecké etiky, ukázal mi potřebnost pokory a výhody týmové spolupráce.

Děkuji také všem kolegům, přátelům, pomvědům a studentům z Fonetického ústavu za to, že dokázali vytvořit intelektuálně stimulující a přátelské prostředí, ve kterém byla radost pracovat. Speciální díky patří Tomáši Bořilovi za pomoc při automatické anotaci korpusu, Hance Bartůňkové, Elišce Churaňové, Veronice Pojarové a Ditě Hývlové za pomoc s ruční segmentací korpusu a jejich podporu v osobní i odborné rovině.

Františku Vlasákovi děkuji za rady při práci s Excelem a trpělivou pomoc při konstrukci temporálního modelu. Janu Harvalíkovi patří upřímný dík za to, že mě první dva roky doktorandského studia živil, abych se mohla plně věnovat vědě.

Všem svým přátelům děkuji za pravidelný přísun podpory, zajímavých informací a za trpělivost, kterou se mnou měli v průběhu mého studia. Neméně velký dík patří také mým rodičům za to, že ze mě vychovali zvědavou a pochybující osobu.

## **Abstrakt**

Tato práce si klade za cíl zevrubně popsat temporální charakteristiky mluvené češtiny prostřednictvím trvání hlásek a jejich změn pod vlivem několika prozodických i segmentálních faktorů, jako je pozice ve vyšší jednotce (slabice, slově či prozodické frázi), délka vyšší jednotky, hláskové okolí, struktura slabiky či frázové zpomalování. Řečový materiál pochází z korpusu semispontánních dialogů, který obsahuje 4046 promluv od 34 mluvčích. Deskripce jsou následně využity pro vytvoření temporálního modelu založeného na pravidlech, který slouží jako srovnávací báze pro analýzu kontur lokálního artikulačního tempa a jejich specifčnosti pro mluvčího. Výsledky naznačují, že systematické rozdíly mezi mluvčími se dají nalézt jak v segmentální doméně, tak i v temporálních konturách. Dále je také posouzen potenciál artikulačního tempa a globálních temporálních ukazatelů pro indentifikaci mluvčího.

**Klíčová slova:** temporální charakteristiky, temporální modelování, trvání hlásek, identifikace mluvčího, čeština

## **Abstract**

This thesis aims to thoroughly describe the temporal characteristics of spoken Czech by means of phone durations and their changes under the influence of several prosodic and segmental factors, such as position in a higher unit (syllable, word or prosodic phrase), length of the higher unit, segmental environment, structure of the syllable or phrase-final lengthening. The speech material comes from a semi-spontaneous corpus of scripted dialogues comprising 4046 utterances by 34 speakers. The descriptions are afterwards used for the creation of a rule-based temporal model, which provides a baseline for analysing local articulation rate contours and their speaker-specificity. The results indicate, that systematic speaker-specific differences can be found in the segmental domain, as well as in the temporal contours. Moreover, speaker identification potential of articulation rate and global temporal features is also assessed.

**Keywords:** temporal characteristics, temporal modelling, phone duration, speaker identification, Czech

# Obsah

1	Úvod.....	7
2	Výzkum rytmické struktury řeči .....	9
2.1	Globální temporální ukazatele .....	10
2.2	Modely spřažených oscilátorů.....	11
2.3	Experimenty s cyklickou řečí.....	12
2.4	Experimenty se synchronizací.....	12
2.5	Experimentální ověřování percepce rytmu .....	13
3	Temporální charakteristiky v rozpoznávání mluvčího .....	15
4	Modelování temporální struktury .....	17
4.1	Inherentní trvání hlásek.....	19
5	Temporální charakteristiky češtiny: přehled dosavadního výzkumu .....	20
5.1	Trvání českých hlásek .....	20
5.1.1	Vokály.....	22
5.1.2	Konsonanty .....	25
5.1.3	Obecné tendence .....	29
5.2	Faktory ovlivňující trvání hlásek .....	29
5.2.1	Artikulační tempo .....	29
5.2.2	Přízvuk.....	30
5.2.3	Temporální kompenzace.....	30
5.2.4	Vliv pozice v rámci vyšší jednotky.....	32
5.2.5	Vliv mluvního stylu .....	32
5.2.6	Vliv kategorie slova .....	33
6	Výzkumné otázky a cíle práce .....	34
7	Metoda a materiál .....	36
7.1	Korpus Minidialogy-H a jeho anotace .....	36
7.1.1	Vrstva hlásek.....	37
7.1.2	Vrstva fonémů.....	38
7.1.3	Vrstva typů hlásek .....	39
7.1.4	Vrstva slabik .....	40
7.1.5	Vrstva slov .....	41
7.1.6	Vrstva prozodických frází.....	42
7.2	Extrakce dat.....	42
7.3	Čištění dat.....	43
7.4	Statistická analýza dat .....	43
7.5	Artikulační tempo a globální temporální ukazatele .....	44
7.6	LAR.....	45
8	Deskripce trvání hlásek.....	46
8.1	Vokály .....	46
8.1.1	Vokály [a] a [a:].....	47
8.1.2	Vokály [e] a [e:].....	49
8.1.3	Vokály [i] a [i:] .....	50
8.1.4	Vokál [o].....	52

8.1.5	Vokály [u] a [u:] .....	52
8.1.6	Diftong [ou̯] .....	54
8.1.7	Vokál [o:].....	54
8.1.8	Vliv frázového zpomalování na trvání vokálů.....	55
8.1.9	Trvání vokálů po rázu .....	60
8.1.10	Vliv délky vyšší jednotky na trvání vokálů .....	61
8.1.11	Vliv slabičné struktury na trvání vokálů.....	65
8.2	Průměrné trvání vokálů .....	65
8.2.1	Dlouhé vs. krátké vokály .....	67
8.3	Konsonanty .....	67
8.3.1	Explozivy.....	68
8.3.2	Ráz .....	78
8.3.3	Frikativy.....	79
8.3.4	Afrikáty.....	88
8.3.5	Nazály .....	95
8.3.6	Vibranty a aproximanty .....	103
8.4	Průměrné trvání konsonantů .....	111
8.5	Diskuse k deskriptivním analýzám .....	113
9	Jevy specifické pro mluvčího .....	118
9.1	Trvání segmentů.....	118
9.2	Tempo řeči .....	123
9.3	Globální temporální ukazatele .....	126
9.4	Diskuse k forezním analýzám .....	128
10	Model temporálních charakteristik .....	130
10.1	Výchozí hodnoty .....	130
10.2	Koeficienty.....	132
10.3	Modelované faktory .....	133
10.4	Ladění modelu.....	135
10.5	Modelové kontury LAR a individuální rezidua .....	137
10.6	Diskuse k temporálnímu modelu .....	152
11	Závěr.....	157
	Reference .....	159
	Přílohy.....	166
	Příloha 1: Text korpusu Minidialogy-H .....	166
	Příloha 2: Fonémová transkripce korpusu Minidialogy-H .....	168
	Příloha 3: Sylabifikace textu korpusu Minidialogy-H.....	171
	Příloha 4: Tabulky deskriptivních statistik českých hlásek .....	174
	Příloha 5: Koeficienty temporálního modelu .....	177
	Příloha 6: Grafy reziduí lokálního artikulačního tempa .....	185

# 1 Úvod

Tato disertační práce má dva hlavní cíle: popsat temporální vlastnosti českých hlásek a faktory, které se podílejí na organizaci temporální struktury češtiny, a na tomto základě vytvořit model, který popisuje její referenční stav. Tento model bude následně sloužit jako srovnávací základ k prozkoumání individuálních odchylek jednotlivých mluvčích, které by se daly využít pro jejich identifikaci ze zvukového záznamu.

Temporální struktura řeči, tedy uspořádání řečových jevů a událostí v čase, úzce souvisí s několika dalšími problematikami, zejména s komplexním pojmem řečového rytmu. Ten dosud není uspokojivě popsán, a to ani v jiných jazycích světa. Obecně se ví, že různí lidé mluví různě rytmicky, různě překotně či různě rychle, také jinak v různých mluvních situacích či kontextech; temporální uspořádání se může měnit například i s náladou nebo afektivním stavem mluvčího. Fonetická podstata těchto změn ale není dosud zmapována. Do hry totiž vstupuje velké množství faktorů a zatím není jasné, jak je uchopit, tj. kvantifikovat či modelovat.

Podrobnější popis temporálních vlastností českých hlásek v deskripci češtiny prozatím chybí, často jsou uváděny informace zastaralé či zavádějící, a neexistuje ani srovnávací rámec, ke kterému by se mohly vztahovat nejen deskriptivní příručky, ale i odborné studie zabývající se touto problematikou. Poskytnutí reprezentativních hodnot trvání českých hlásek poslouží jako odrazový můstek k budoucímu zkoumání (například tempa řeči, akustických korelátů slovního přízvuku, cizineckého přízvuku v češtině apod.), které bude moci na těchto hodnotách stavět a porovnávat s nimi své výsledky. Tento srovnávací rámec bychom zde chtěli poskytnout.

Kromě popisu trvání českých hlásek chce tato studie také přispět k objasnění vlivu hlavních faktorů, segmentálních i prozodických, na temporální strukturu řeči. Mnohé vlivy jsou již známy z jiných jazyků, ale výsledky pro češtinu jsou prozatím kusé či těžko zobecnitelné. Temporální model, který bude na základě těchto výsledků vytvořen, by se následně mohl stát prvním krokem v popisu rytmu češtiny.

Předpokládáme, že výsledky této práce budou také přímo aplikovatelné – našim primárním cílem je postihnout individuální rozdíly, užitečné pro potřeby forenzní analýzy, sekundárně předpokládáme také využitelnost pro účely syntézy a rozpoznávání řeči. Mnoho studií se shoduje na tom, že je potřeba lépe modelovat trvání řečových segmentů vzhledem k ostatním prozodickým faktorům, aby se zlepšila úspěšnost rozpoznávačů řeči a zvýšila přirozenost řečové syntézy (např. Carlson, 1991: 281; O'Shaughnessy, 1995: 600; Lazaridis et al., 2010: 175). Také v systémech automatického rozpoznávání mluvčího je temporální doména zatím

využívána jen zřídka. Návrhy, jak a kde hledat individuální temporální rozdíly mezi mluvčími, by měly ulehčit práci forezním expertům, jimž tak přibude další biometrický parametr využitelný pro porovnávání mluvčích.

Práce je strukturována následovně: Ve druhé kapitole bude čtenář seznámen se zásadními otázkami současného výzkumu rytmické struktury řeči, kapitola 3 se zabývá využitím temporálních charakteristik při rozpoznávání mluvčího. V kapitole 4 jsou představeny přístupy k modelování temporálních charakteristik promluv. Kapitola 5 pak obsahuje zevrubné shrnutí studií na téma trvání českých hlásek a jejich výsledků. V kapitole 6 jsou nastoleny výzkumné otázky a hypotézy a kapitola 7 představuje využitou metodu a řečová data, na jejichž základě je celý výzkum proveden. Deskriptivní statistiky českých hlásek jsou obsahem kapitoly 8, kapitola 9 pak na základě těchto popisů shrnuje jevy specifické pro mluvčího. Modelu trvání českých hlásek, jeho výstavbě, ladění a aplikaci na reálné promluvy je věnována kapitola 10. Součástí kapitol 8–10 jsou vždy také diskuse k příslušným tématům. Kapitola 11 pak shrnuje závěry práce, její omezení a možné směry budoucího výzkumu.



## 2 Výzkum rytmické struktury řeči

Rytmus je jedním z jevů, které prostupují mnoha úrovněmi lidské existence. Tendence k pravidelně se opakujícím vzorcům nalézáme ve všech oblastech, od fyziologie přes astronomii až po hudbu. Rytmus řeči je pak komplexní fenomén, jenž podléhá mnoha vlivům, od fyziologických přes sémantické, syntaktické, pragmatické i extralingvistické (Klatt, 1976: 1210), a tyto vlivy mohou být v různých situacích a u různých mluvčích zcela odlišně hierarchizovány.

Kohler (2009a) ve svém vlivném článku o novém paradigmatu pro výzkum řečového rytmu podtrhuje zejména jeho komunikační funkci – rytmus řeči je produkován pro posluchače, aby mu ulehčil percepci a zpracování řečeného. Pravidelnosti napomáhají lidskému mozku snáze a rychleji zpracovávat vjemy, naopak nepravidelnosti vedou k potížím a vyššímu kognitivnímu zatížení. Pravděpodobně každý z nás má zkušenosti s obtížnou srozumitelností překotných či nepravidelných promluv – tento jev byl ověřen i empiricky (např. Huggins, 1979 nebo Quené & Port, 2005).

Rytmem se ovšem nemusí lišit pouze mluva jednotlivých osob. Vědci zabývající se řečí si všímali již v 18. století rozdílů mezi rytmem různých jazyků (v anglosaském kontextu především angličtiny a francouzštiny). V Pikeově knize *The Intonation of American English* (Pike, 1945) se tomuto percepčnímu rozdílu dostalo názvu slabičná vs. taktová izochronie (resp. *stress-timed* vs. *syllable-timed structure*), což poukazovalo na pravidelnost promluv vzhledem ke všem slabikám či pouze k přízvučným. Slabičně izochronní promluva má tendenci k tomu, aby všechny slabiky byly stejně dlouhé, naproti tomu taktově izochronní zachovává stejné vzdálenosti mezi přízvučnými slabikami. Již Pike však poukazoval na to, že jazyky neuplatňují výhradně jeden z těchto principů, ale liší se v tom, který upřednostňují. Kohler (2009a: 30) si všímá toho, že v pozdějších letech jako by tato poznámka byla zapomenuta, a izochronie byla dlouho považována za prostředek diskrétní kategorizace.

Myšlenka rozřadit jazyky světa do tříd podle jejich prozodických vlastností je však zcela jistě lákavá a stala se zdrojem mnoha podnětných experimentů, které obohatily naše znalosti o rytmu a temporální struktuře.

Nicméně Cummins (2002) přesvědčivě argumentuje, že bychom měli opustit myšlenku rytmu jako diskrétního parametru kategorizace jazyků a definovat jej spíše jako fenomén, který sjednocuje slabiky a jejich nadřazené jednotky v temporální struktuře promluvy. Toto spřažení vnořených prozodických jednotek je pak podle Cumminse proměnlivé a závislé na mluvčím, textu, stylu, a dalších faktorech.

Kohlerova novější definice rytmu, která spojuje percepční, produkční i komunikační stránku věci, zní takto:

„Constrained by the phonetic structures of the languages of the world, speech rhythm is the production, for a listener, of a regular recurrence of waxing and waning prominence profiles across syllable chains over time, with the communicative function of making speech understanding in various speaking styles more effective.“ (Kohler, 2009a: 41)

Obě definice se v žádném případě nevylučují, spíše doplňují. Zatímco Cummins podtrhává nutnost nahlížení na rytmus jako spojitý, hierarchický fenomén, Kohler přesunuje pozornost na posluchače a řečovou komunikaci.

Podotkněme zde, že Kohlerova formulace „vzrůstající a ubývající prominenční profily“ není zcela neproblematická a lépe by možná mohla být vyjádřena výrokem, že vzrůstající a ubývající prominence vytvářejí rytmický profil promluvy.

Klíčový je zde právě pojem prominence – v Kohlerově definici již nejde o binární vlastnost slabiky (má × nemá přízvuk), ale o průběžnou charakteristiku celé promluvy, jejíž pravidelnost je základním konstituentem rytmu. Za fyzikální koreláty prominenčních vzorců Kohler považuje trvání, základní frekvenci, intenzitu a spektrální dynamiku (Kohler, 2009a: 38). Jednou ze stále otevřených otázek je míra přispění těchto korelátů k percepci prominence v různých jazycích. Obzvláště v češtině, vyznačující se pevným slovním přízvukem bez přímého akustického značení, je tento problém stále naléhavější.

V dalších podkapitolách budou nastíněny hlavní současné experimentální přístupy k výzkumu rytmické struktury řeči.

## **2.1 Globální temporální ukazatele**

Přestože je myšlenka binárního (či ternárního) rozdělení jazyků do kategorií slabičné a taktové (příp. morové) izochronie dnes již považována za zjednodušující a můžeme snad říci i překonanou, stále se s ní i v současné době lze setkat v mnoha okrajovějších oblastech výzkumu řeči. Směr výzkumu, který se tyto rozdíly snažil zachytit, nicméně podnítil mnoho studií usilujících o empirickou kvantifikaci rytmu (např. Ramus et al., 1999; Low et al., 2000; Grabe & Low, 2002; Asu & Nolan, 2006; Dellwo, 2006 nebo Arvaniti, 2009).

Všechny tyto metody bývají souhrnně označovány jako „rytmické ukazatele“. Na základě trvání vokalických a konsonantických intervalů v řeči se snaží nějakým způsobem jedinou hodnotou postihnout globální temporální organizaci promluv (či celých jazyků). Výhodou

těchto ukazatelů je relativní snadnost a přímočarost jejich získávání ze zvukového materiálu. Kritika ale naráží především na to, že pouhým měřením časových intervalů není možné popsat rytmus řeči v celé jeho složitosti (Cummins, 2002; Kohler, 2009a: 33–34, Barry et al., 2009, Volín, 2010), nelze se vyhýbat vlivu ostatních prozodických faktorů zmíněných výše. Dalším problémem těchto globálních ukazatelů je nesnadnost jejich lingvistické interpretace. Nepravidelnosti zachycené stejnou hodnotou těchto metrik mohou mít velmi odlišné příčiny (viz např. Lowit, 2012, a její experiment s patologickou řečí), a naopak percepční pravidelnost nemusí být jimi vůbec zachytitelná (Kohler, 2009a: 39, Barry et al., 2009).

Globální temporální ukazatele Ramuse, Lowové, Grabové, Dellwa a ostatních zachycují jen jeden z jevů podílejících se na vnímaném rytmu – a to fonotaktickou strukturu jazyka a její realizaci danými mluvčími. Rytmus je tak zploštěn na konstrukt určený slabičnou strukturou a inventářem segmentů daného jazyka (Cummins, 2002). Na druhou stranu i tato doména nabízí prostor pro individuální odchylky, až do té míry, že mohou být využity pro rozpoznání mluvčích, viz např. Dellwo & Koreman (2008), Dellwo et al. (2012), nebo Leemann et al. (2014); pro češtinu Weingartová (2013).

Dalším bodem, který vyplývá z Kohlerovy definice a který globální temporální ukazatele nejsou schopny uchopit, je nutnost lokálního popisování rytmu. Proměnlivé prominenční profily vytvářené nejen trváním, ale i základní frekvencí, intenzitou a spektrální dynamikou jsou těžko zachytitelné jediným číslem, které nějak kvantifikuje jejich pravidelnost. Nicméně souhlasíme s Cumminsem (2002), jenž argumentuje, že i tak nám tyto ukazatele mohou poskytnout užitečné, ač dílčí informace.

## **2.2 Modely spřažených oscilátorů**

Spojité vztah mezi vnořenými prozodickými jednotkami, jak rytmus definuje Cummins (2002), se snaží modelovat teorie spřažených oscilátorů (např. Barbosa, 2002; O'Dell & Nieminen, 2008; Šimko & Cummins, 2010 nebo Large et al., 2010). V jejím rámci si můžeme představit například úroveň slabik jako jeden oscilátor a úroveň taktů jako druhý – oba jsou spřažené, takže své cykly navzájem ovlivňují. Převáží-li „rytmus“ oscilace slabik, vznikne slabičně izochronní promluva, ve druhém případě taktově izochronní. Tento přístup tedy umožňuje modelovat případy, kdy se v rámci jednoho jazyka (případně i v rámci jediné promluvy) objevují tendence jak ke slabičné, tak i k taktové izochronii.

Tyto modely také stejně jako globální temporální ukazatele popisují ve své podstatě pouze načasování segmentů. Vliv jiných prozodických faktorů je ale do jejich architektury možné zakomponovat.

### 2.3 Experimenty s cyklickou řečí

Ve své studii Cummins a Port (1998) využívají experimentální paradigma, které nazvali cyklickou řečí (*speech cycling*), k osvětlení vztahů mezi základními temporálními jednotkami rytmu. V tomto experimentálním paradigmatu jde konkrétně o produkci rytmických jednotek. Mluvčí angličtiny opakovali jednoduché fráze do rytmu metronomu a přízvukné slabiky měli synchronizovat s jednotlivými údery, jejichž relativní trvání vůči sobě bylo různě manipulováno. Převládající strategií bylo vnoření dvou nebo tří taktů do jednoho cyklu. To podle Cumminse (2002) dokazuje využívání přízvukového taktu mluvčími angličtiny pro vytvoření rytmického vzorce promluvy, přestože jde o extrémní případ vnucené rytmické organizace. Souvisejícím experimentem podpořil Tajima (1998) využívání podobné jednotky v japonštině (bimoraický takt).

Nicméně v případě aplikace úkolu na mluvčí italštiny a španělštiny toto experimentální paradigma selhalo. Ukázalo se, že mluvčí těchto jazyků nebyli schopni vůbec synchronizovat přízvukné slabiky s tóny metronomu. Cummins (2002) z toho vyvozuje, že italština ani španělština nemá k dispozici jednotky na úrovni přízvukového taktu, které by vnořila do opakující se fráze. Podobný způsob využily v poslední době také Chungová a Arvaniti (2012) k objasnění základních temporálně-rytmických jednotek v korejštině.

Ačkoliv mají promluvy produkované v rámci těchto experimentů jistě daleko k přirozené řeči, mohou osvětlit některé základní hluboko zakořeněné rytmické principy na základě temporálního uspořádání prominencí.

### 2.4 Experimenty se synchronizací

Kooperativní funkcí rytmu se zabývají experimenty se synchronizací řeči (např. Cummins, 2002, 2003 nebo 2009). Obvykle je úkolem mluvčích přečíst text současně s jiným mluvčím tak, aby byla jejich řeč co nejvíce synchronní. Experimentální podmínky mohou omezit vizuální kontakt, také se někdy zkoumá synchronizace s nahrávkou. V experimentu ze studie Cummins (2002) byla změřena temporální struktura jedné z vět tak, že mezi ručně označenými dobře definovanými body (exploze, onsety vokálů atd.) vznikly intervaly, které byly vyjádřeny nejprve absolutním trváním v milisekundách a následně jako poměrná část trvání nadřazeného intervalu. Každá věta každého mluvčího byla tedy vyjádřena dvěma vektory hodnot trvání. Posléze byla spočítána euklidovská vzdálenost těchto vektorů od všech mluvčích navzájem (absolutní trvání a poměrné trvání zvlášť) a promluvy byly podle vzdálenosti seřazeny. Rozdíl v pořadí dvou vět od stejného mluvčího a dvou vět, které byly

navzájem synchronizovány, posloužil jako index podobnosti temporální organizace daných promluv.

Výsledky ukazují zajímavý trend: porovnání absolutního trvání temporálních intervalů k sobě lépe přiřazuje dvě věty, které se mluvčí snažili synchronizovat, zatímco normalizované hodnoty trvání vyjádřené jako poměr z nadřazeného intervalu k sobě naopak lépe přiřazují dvě věty prosloušené stejným mluvčím, přestože obě věty byly proneseny v různých podmínkách (samostatné vs. synchronizované čtení).

Vidíme zde tedy nejen schopnost mluvčích synchronizovat temporální strukturu s jiným mluvčím, ale také fakt, že koordinace temporálních jednotek na nižší a vyšší úrovni je specifická pro mluvčího a mluvčí ji do velké míry nemění ani se změnou podmínek. To je v souladu s výsledky Wretlinga a Erikssona (1998) z experimentů s profesionálními imitátory. Tento výsledek podporuje jednak pohled na rytmickou strukturu jazyka jako důsledek jak slabičné struktury, tak i propojení vyšších a nižších prozodických jednotek, nicméně využitá metodologie je také relevantní pro identifikaci mluvčího – zdá se, že hypotéza individuálních rytmických vzorců může být ověřována vyjadřováním trvání segmentů jako poměrných částí jednotek prozodicky nadřazených.

Experimenty zabývající se synchronizací, nebo též akomodací či konvergencí mluvčích, nacházejí široké uplatnění. Jednak ve výzkumu cerebrálních periodicit (např. Ghitza & Greenberg, 2009), nebo se využívají též ve zkoumání diskurzního chování mluvčích při konverzaci (např. Beňuš, 2009). Kromě toho existuje též široká oblast výzkumu sensorimotorické synchronizace, tj. synchronizace vnímaných stimulů s pohybem (viz např. Repp, 2005).

## **2.5 Experimentální ověřování percepce rytmu**

Kohler (2009a: 35) vyzdvihuje a zdůrazňuje percepční roli rytmu, která bývá ve výzkumu rytmické struktury často zanedbávána. Podle jeho názoru by všechny výsledky týkající se produkce měly být důsledně ověřovány percepčně. Protože parametry, které nejsou percepčně relevantní, a posluchač je tedy nijak nevyužívá, nemá smysl do popisu rytmu zahrnovat (mohou mít ovšem dobrý smysl například při identifikaci mluvčího).

Již Hála (1962: 183) si všímá toho, že objektivní, artikulační trvání hlásek se nekryje s jejich trváním percepčním. Hlávky potřebují ke svému vyslovení různě dlouhý časový úsek – posluchači ale tyto rozdíly nevnímají jako změny tempa. Přestože je tedy z mnoha důvodů potřeba měřit objektivní trvání hlásek, do modelu rytmu řeči je zcela nezbytné zařadit i percepční rovinu a zachytit ty charakteristiky, které jsou relevantní pro posluchače.

Percepčních experimentů není mnoho, jednou z prvních byla Lehiste (1973, 1977, citováno z Lehiste, 1979), která etablovala izochronii jako subjektivní, nikoliv objektivní jev. Její výsledky naznačují, že posluchači nemají problém porovnávat trvání neřečových zvuků (v tomto případě šumu), ale velké potíže jim činí posoudit objektivní trvání zvuků řečových. Proto studie měřící izochronii selhaly – posluchači neslyší objektivní trvání akustických jevů a promluvy se jim zdají pravidelnější, než doopravdy jsou (Lehiste, 1979: 244). Tento závěr potvrdili také Donovan a Darwin (1979).

V poslední době se začínají objevovat další percepční experimenty, například Zhengová a Pierrehumbertová (2010) nacházejí vliv přízvučnosti slabiky na vnímání její délky, Kimová a McAuley (2012) ukazují, že odchýlné umístění jevu v pravidelné sekvenci vede ke změně percepce trvání tohoto jevu. Knightová a Cross (2012) porovnávali percepční rytmičnost řeči různých mluvních stylů od běžné konverzace přes politické proslovy po metrickou poezii a rovněž nechávali posluchače vytukávat rytmus. Ukázalo se, že posluchači jsou schopni poměrně konzistentně sledovat rytmus mluvího a že všechny zkoumané styly se vskutku signifikantně liší v rytmičnosti.

V této práci bude hlavní pozornost věnována kvantitativní analýze temporální struktury řeči, s plným vědomím toho, že nejde o rytmus jako takový. Nicméně načasování a trvání řečových jevů zcela jistě je základní a neodmyslitelnou součástí rytmu – a bez informací o temporální doméně nemůžeme na jeho komplexní popis vůbec aspirovat. Půjde tedy o nezbytný krok k uchopení rytmu češtiny. V souladu s Kohlerovým novým paradigmatem pro výzkum rytmu (Kohler, 2009a) by měl následovat další krok ve formě popisu řečové prominence, jejích fyzikálních korelátů a následně též percepce. Tímto směrem výzkum v pražském Fonetickém ústavu také postupuje, především v souvislosti s popisem české angličtiny (viz Volín & Weingartová, 2014, či Weingartová et al., 2014b), nicméně není předmětem této práce.

### 3 Temporální charakteristiky v rozpoznávání mluvího

V řeči existují rysy, které jsou pro každého mluvího specifické a umožňují nám více či méně spolehlivě rozpoznat osoby pouze podle jejich hlasu. Obor rozpoznávání mluvího se zabývá vyhledáváním a měřením těchto rysů a je jednou z hlavních aplikací forenzní fonetiky.

Většina výzkumu v oblasti identifikace mluvího se soustředí na spektrální informace v řečovém proudu, využití temporálních charakteristik je spíše vzácností. Nicméně jejich přitažlivost je založena na dvou hypotézách, za první mluvíci mohou mít svůj vlastní naučený rytmus řeči, za druhé, dynamika pohybů mluvidel je pro mluvího jedinečná (podobně jako například dynamika pohybů při chůzi, viz např. Bouchrika et al., 2011) díky fyziologickým odlišnostem. Jak již bylo řečeno výše, měření rytmu se obvykle omezuje na měření temporálních charakteristik, a nejinak je tomu i při rozpoznávání mluvího.

Temporální charakteristiky mají oproti charakteristikám spektrálním tu výhodu, že jsou odolné vůči distorzím signálu šumem nebo pásmovými filtry (například telefonního přenosu). Škodlivé jsou pro ně pouze reverbace neboli tzv. echo. Několik zdrojů také ověřilo, že mluvíci zásadním způsobem nemění své temporální charakteristiky, když chtějí maskovat svůj hlas (Eriksson & Wretling, 1997, Wretling & Eriksson, 1998 nebo Dellwo et al., 2009), což může být ve forenzní praxi velmi výhodné. Pokud mluvíci opravdu vědomě nemění některé své temporální charakteristiky, půjde o velmi užitečný biometrický ukazatel.

Ve své přehledové studii o forenzních praktikách ve světě uvádějí Goldová a French (2011: 753), že 93 % dotázaných forenzních expertů bere při rozpoznávání v úvahu tempo řeči. Z toho 80 % měří artikulační nebo mluvní tempo a 71 % se snaží nějakým způsobem postihnout rytmus řeči (bohužel článek neuvádí podrobnosti). Ve srovnání úspěšnosti fonetických parametrů pak tempo řeči jako úspěšné uvádí 21 % respondentů a rytmus 17 % (Gold & French, 2011: 754). Tato čísla by se mohla zdát nízká, nicméně autoři argumentují, že je to především kombinace různých fonetických rysů, která umožňuje úspěšné rozpoznání.

Hollien ve svém poloautomatickém rozpoznávacím systému SAUSI (*Semi-Automatic Speaker Identification*) také využívá některé temporální rysy, a to počet a délku pauz (intervalů ticha), poměr řeči a pauz, slabičné tempo, poměr řeči a celkového času promluvy, celkový čas promluvy, ve kterém je přítomna akustická energie, a další (Hollien, 2002: 166n.). Konstatuje, že tyto rysy, ač samy nejsou příliš úspěšné, vylepšují funkčnost systému v kombinaci s ostatními.

Nejčastěji využívaným ukazatelem identity mluvího je podle Goldové a Frenche (2011) artikulační tempo (trvání lingvistické jednotky, nejčastěji slabiky, za jednotku času, obvykle sekundu, s vyřazením pauz). Studie se shodují, že vykazuje poměrně malou variabilitu

v rámci mluvího, ale velkou mezi mluvími (všimla si toho např. už Goldman-Eisler, 1968; podrobněji se forenznímu využití věnují Künzel, 1997 nebo Jessen, 2007). Používá se také mluvní tempo, které zahrnuje pauzy, hezitace, dysfluence, apod.). Mluví se mohou vzájemně lišit právě těmito jevy, např. frekvencí či délkou pauz (O'Shaughnessy, 1995: 603, van Donzel & Koopmans-van Beinum, 1996; pro češtinu Weingartová et al., 2014a).

Také při výzkumu globálních temporálních ukazatelů pro kategorizaci jazyků do rytmických tříd (viz oddíl 2.1) naráželi autoři studií často na vysokou variabilitu mezi mluvími (viz např. Ramus, 2002; Asu & Nolan, 2006; White & Mattys, 2007; Arvaniti, 2009; Wiget et al., 2010). To vedlo k pokusům o využití těchto ukazatelů pro rozpoznávání mluvího (např. Dellwo & Koreman, 2008; Dellwo et al., 2012; Leemann et al., 2014 nebo Weingartová, 2013) s různou mírou úspěšnosti. Nicméně přestože zacházení s vokalickými a konsonantickými intervaly může obsahovat jistou idiosynkratičnost, čeština se svými konsonantickými shluky a lexikálně danou délkou vokálů se zdá poskytovat k individuálním odchylkám těchto ukazatelů méně prostoru.

Problémy s globálními metrikami již byly zmíněny. Zploštěním dynamického parametru, jakým je temporální organizace výpovědi, do jediného čísla přicházíme o velké množství informací, kódovaných právě lokálními změnami (zaznamenat chůzi člověka jediným číslem by jistě také bylo příliš zjednodušující). Proto je jedním z cílů této práce popsat temporální charakteristiky lokálně (tedy vytvářet a zkoumat temporální kontury promluv) a věnovat se především jejich změnám, odchylkám mluvích od srovnávacích hodnot či od sebe navzájem.



## 4 Modelování temporální struktury

Základem pro modelování temporální struktury je popis a model trvání segmentů (zejména hlásek). Jde o klíčový parametr pro kvalitní rozpoznávání i syntézu řeči (O’Shaughnessy, 1995: 600; Lazaridis et al., 2010: 175). U syntézy je totiž důležitá nejen srozumitelnost, ale též přirozenost, tj. co největší podobnost skutečné lidské řeči. Proto je třeba modelovat také prozodické charakteristiky promluv, zejména právě trvání segmentů, konturu F0 či intenzitu. Důležitým mezníkem v modelování trvání segmentů byl model Dennise Klatta (1976), který jednoduchým způsobem kombinuje inherentní (tj. průměrné nebo podle Klatta fonologické) a minimální, „nestlačitelné“ trvání (tj. nejmenší možná doba pro uspokojivé provedení artikulačního gesta, Klatt, 1976: 1215) jednotlivých hlásek s faktory, jež mají vliv na trvání výsledné.

Vzorec modelového trvání hlásky vypadá takto:

$$D_j = K(D_i - D_{\min}) + D_{\min},$$

kde  $D_i$  je inherentní trvání,  $D_{\min}$  minimální trvání dané hlásky,  $K$  je konstanta, která upravuje „stlačitelnou“ část hlásky a  $D_j$  je pak její výsledné trvání (Klatt, 1976: 1216). Klíčovou roli zde hraje  $K$ , které hlásku prodlužuje (když  $K > 1$ ) či zkracuje ( $K < 1$ ). Ve své studii podává Klatt (1976) shrnutí mnoha faktorů, jež mohou tuto konstantu ovlivnit. Podrobněji jim bude věnován prostor v kapitole 5.2.

Tento model byl v Klattově (poměrně malém) korpusu schopen predikovat trvání vokálů v různých pseudoslovech. Zda jsou predikce přijatelné i percepčně, ověřovali Carlson, Granström a Klatt (1979, citováno z Carlson, 1991) experimentem se syntetizovanou řečí, první dva autoři jej následně upravovali a rozšiřovali pro švédštinu. Na základě korpusu obsahujícího promluvy jediného mluvčího byli schopni predikovat trvání hlásek se směrodatnou odchylkou 20 ms (Carlson, 1991: 243).

Klattův model je klasickým příkladem modelu založeného na pravidlech (*rule-based*). Apriorním úsudkem nebo na základě analýzy dat formulujeme pravidla, která pak model vytvářejí. Ten je jen tak dobrý, jak přesná a vyčerpávající jsou námi stanovená pravidla a jaké množství faktorů ovlivňujících trvání jsme schopni postihnout. Takový druh výzkumu však může trvat velmi dlouho a vyžaduje fonetické znalosti a zkušenosti, což technicky založené studie považují za hlavní nevýhodu tohoto postupu. Technická větev temporálního modelování se proto uchyluje k modelům založeným na datech (*data-driven*). S pomocí statistického modelování jsou tyto systémy také schopny zachytit variabilitu v datech, někdy dokonce i úspěšněji. Nevýhodou pro naše účely je neprůhlednost těchto modelů – nedozvíme se, které řečové parametry a do jaké míry mají na trvání hlásek vliv (viz Kohler, 2009b: 5).

Při modelování trvání hlásek statistickými algoritmy se nejprve využívá selekce rysů (*feature selection*), ke zjištění, které parametry jsou relevantní pro model a které nikoliv, ke zvýšení úspěšnosti a omezení počítačích náročnosti (Lazaridis, et al., 2010: 176). To vlastně simuluje postup fonetika, který také hledá relevantní parametry mající vliv na trvání hlásek a irelevantní zamítá. Dále jsou tyto vybrané rysy využity jako vstupy do modelů.

Modelů řízených daty je mnoho a liší se především použitým algoritmem statistického modelování: např. rozhodovací stromy, algoritmy líného učení (*lazy learning*), algoritmy metaučení (*meta-learning*), lineární regrese, neuronové sítě, Bayesovské sítě, součty součinů (*sums-of-products*), atd. (viz Carlson, 1991; nověji též Lazaridis et al., 2010). Ovšem počítačově nové způsoby modelování se stále vynořují.

Nezbytným základem výše uvedených postupů jsou velké řečové korpusy – které je ale vzhledem k jejich obsáhlosti často nemožné manuálně zpracovávat, a tedy je nutné hranice hlásek stanovovat automaticky, což může vnést do modelu vyšší chybovost. Naopak modely založené na pravidlech jsou obvykle postaveny na materiálu ručně segmentovaném fonetiky, kterého je ale na druhou stranu řádově méně, protože jde o práci značně časově náročnou. Ideální cestou by jistě byl kompromis, nebo spíše spojení obou těchto přístupů a jejich vzájemné obohacení, k čemuž mnohé studie již léta vybízejí (např. Doddington, 1985: 1663, Carlson, 1991: 246 nebo Bourland et al., 1996). Algoritmy by bylo možné například trénovat na řečovém materiálu ručně anotovaném fonetiky – tento postup se začal úspěšně využívat již i v českém prostředí (Pollák et al., 2007).

Rysy nejčastěji využívané pro statistické modelování jsou identita dané hlásky, počet hlásek ve slabice, pozice hlásky ve slabice přízvuknost slabiky, pozice slabiky ve slově, pozice slova ve frázi, dále také vzdálenost od předcházejícího či následujícího přízvuku, vzdálenost od předcházejícího či následujícího předělu nebo pauzy, artikulační vlastnosti hlásky, identita okolních hlásek, atp. (Lazaridis et al., 2010: 179–180). Mnoho z těchto rysů má fonetické opodstatnění i v češtině (oddíl 5.2).

Statistické modely dosahují poměrně vysoké úspěšnosti, Lazaridis et al. (2010: 184) měli u 10 vyzkoušených modelů korelaci se skutečnými hodnotami neutrální a emocionálně zabarvené řeči pohybující se v rozmezí 0,6–0,8. Nutno ale podotknout, že jimi použitý korpus zahrnoval pouze řeč jedné jediné mluvčí, v případě více mluvčích lze očekávat snížení této korelace. Nejen tato, ale i další studie pak představují zahrnutí explicitního modelování trvání hlásek jako slibný směr dalšího výzkumu rozpoznávání a syntézy řeči (např. Russell & Cook, 1987; Bourland et al., 1996 nebo Pyllkkönen & Kurimo, 2004).

## 4.1 Inherentní trvání hlásek

Vlivy na změnu trvání hlásek v promluvě je podle jejich příčiny možné rozdělit na vnitřní a vnější. Vnitřní, inherentní specifikace trvání je výsledkem biomechanických principů (pohybů mluvidel nutných k dosažení artikulačního cíle), vnější specifikace pak bude zahrnovat prozodické, stylistické, kontextuální, afektivní a další faktory, které mluvčí ovládá vědomě. Při modelování trvání hlásek je důležité obě tyto skupiny zohlednit, a zejména je klíčové znát inherentní hodnoty trvání jednotlivých hlásek jako pevný bod, od kterého je pak možné odvozovat změny tohoto trvání.

Oddělit vnitřní trvání od vnějšího je však netriviální problém, nemusí být jednoduché zjistit, zda pozorovaná trvání jsou daným hláskám skutečně inherentní, nebo zda jsou nějakým způsobem mluvčím manipulována. Je například dobře známo, že kontrast v kvalitě vokálu je v mnoha jazycích světa doprovázen také kontrastem v trvání – otevřené vokály bývají delší než zavřené (pro přehled výsledků viz např. Lehiste, 1970: 18, pro češtinu viz Hála, 1962: 183 a kap. 4.1.1 této práce). Otázka je, zda tento kontrast je aktivně vytvářen mluvčími, nebo zda je pouze výsledkem biomechanických principů, a tedy vokálům inherentní – delší trvání otevřených vokálů by bylo v tomto případě způsobeno delším časem, který potřebují artikulátory (tj. dolní čelist a jazyk) k dosažení cílového nastavení.

Ve své disertační práci ukázala Caisse (1988), že mimovolní, nekontrolované procesy vyvolávají aditivní rozdíly v trvání, zatímco centrálně specifikované a mluvčím aktivně ovládané procesy vykazují multiplikativní efekt (Caisse, 1988: 135). Tuto hypotézu potvrzují i výsledky de Jonga (2004). Na těchto výsledcích staví také hypotéza Solé a Ohaly (2010): je-li nějaký rys ovládaný mluvčím, bude se přizpůsobovat globálnímu i lokálnímu načasování, tempu řeči, přízvuku, struktuře slabiky, apod. Automatické rysy, které souvisí s biomechanickými fyziologickými omezeními, budou zůstávat konstantní, nebo se budou přizpůsobovat pouze monotonicky, a tato změna bude předpověditelná z obecných fonetických principů (Solé & Ohala, 2010: 614). Tímto směrem se ubírají i studie Picketta et al. (1999) nebo Bouchera (2002), kteří nalézají podobné (aditivní vs. multiplikativní) vztahy při změnách tempa.

Všimněme si, že též Klattův vzorec výše je multiplikativní – čili počítá s mluvčím ovládanými změnami temporálních charakteristik. Inherentní trvání je zakomponováno v proměnné  $D_i$ . Velmi podobně bude postupováno i v této práci (viz oddíl 10.2).

## 5 Temporální charakteristiky češtiny: přehled dosavadního výzkumu

Temporálními vlastnostem češtiny se v průběhu historie fonetického výzkumu věnovala řada odborníků. Cílem tohoto oddílu bude jejich výsledky přehledně shrnout a okomentovat, a v další části práce zároveň přispět vlastními poznatky a spolehlivými průměrnými hodnotami trvání, které by odpovídaly současným požadavkům na kvantitativní experimentální výzkum.

### 5.1 Trvání českých hlásek

Prvním, kdo se pokusil o systematický popis trvání hlásek v češtině, byl průkopník české fonetiky, Josef Chlumský. Ve svém *Pokusu o měření českých zvuků a slabik v řeči souvislé* (1911) uvádí hodnoty trvání hlásek u dvou mluvčích češtiny, z nichž jedním byl on sám. Tyto poznatky dále rozvádí v rozsáhlejší práci *Kvantita, melodie a přízvuk* (1928), kdy již měl k dispozici mluvčí tři. Chlumský se zabýval jednotlivými mluvčími i temporálními jevy postupně a velmi podrobně. Jeho celkové výsledky přehledně shrnuje Hála v *Uvedení do fonetiky češtiny na obecně fonetickém základě* (1962). I přes stáří těchto studií (tehdy například nebylo možné měřit s přesností na milisekundy, jak je to obvyklé dnes) se stále jedná o počín zásadní, který snese srovnání s výsledky novějších kvantitativních studií. Nicméně jak konstatoval Hála, Chlumským naměřené hodnoty nemají obecnou platnost a bude potřeba je ověřit na větším počtu mluvčích a v různých mluvních situacích. Zaplnit tuto mezeru je také jedním z cílů této práce.

V šedesátých letech se trváním hlásek kromě drobné studie o českých a holandských vokálech L. Kaiserové s využitím výzkumu Přemysla Janoty (Kaiser, 1964) zabývala ještě práce Borovičkové a Maláče *The Spectral Analysis of Czech Sound Combinations* (1967), která se (pravděpodobně) stala východiskem též pro hodnoty trvání hlásek uváděné v *Mluvnici češtiny* (Petr et al., 1986). Přehledné shrnutí všech těchto výsledků lze nalézt v Palkové (1994).

Na tento směr výzkumu bylo po dlouhé odmlce navázáno teprve po obnovení pražského Fonetického ústavu v 90. letech – pro novější empirický výzkum trvání českých hlásek je třeba nahlédnout do kvalifikačních prací absolventů, plošná studie, která by replikovala Chlumského výsledky na bohatším materiálu, zatím provedena nebyla.

Trvání slabičných likvid se věnovala Vernerová (2006); Machač (2006) a Šimek (2010) popisovali trvání českých exploziv a Homolková (2009) frikativ. Relevantní výsledky můžeme najít též u Ondruškové (2011) zkoumající jednoslabičná slova. Hodnoty týkající se

vybraných vokalických elementů nalezneme v disertační práci Studenovského (2012) o českých diftonzích.

Geografickou výjimkou je výzkum V. J. Podlipského z Univerzity Palackého v Olomouci, který se zabývá kontrastem kvantity českých vokálů, také v kontextu osvojování angličtiny jako druhého jazyka (viz např. Podlipský, 2009 nebo Podlipský et al., 2009).

Jediným novějším zdrojem, kde je možné nalézt referenční hodnoty pro všechny české hlásky, je kniha J. Psutky a jeho spolupracovníků *Mluvíme s počítačem česky* (Psutka et al., 2006), kteří údaje o trvání hlásek extrahovali z rozsáhlého strojově zpracovaného korpusu pro potřeby řečových technologií, zejména syntézy řeči.

Co se týče typu využitého materiálu, je v těchto pracích zřetelný přechod od regulovaného laboratorního materiálu a pseudoslov přes čtené projevy až k semispontánním a spontánním promluvám. Povahu a také množství materiálu je nutno vzít v úvahu, neboť mohla mít na trvání segmentů zásadní vliv.

Chlumského (1928) mluvčí četli izolované věty do nálevky fonografu. Je otázka, do jaké míry bylo při tomto nastavení možné dosáhnout přirozené řeči, nicméně Chlumského výsledky obstojí i ve srovnání s moderními studii. Ve výzkumu Kaiserové a Janoty (1964) byla využita izolovaná slova čtená pěti mluvčími, Borovičková a Maláč (1967) zkoumali izolované logatomy čtyř mluvčích. Jejich výsledky se díky tomu (jak ještě uvidíme níže) značně liší od hodnot naměřených jinými odborníky. Údaje o metodě využitě pro získání hodnot z Mluvnice češtiny (Petr et al., 1986) nelze získat, autoři pouze uvádějí, že hodnoty jsou postaveny na výsledcích Borovičkové.

Co se týče moderních studií, většina je postavena na čtených a semispontánních projevech. Vernerová (2006) pracovala se souvislými čtenými texty šesti mluvčích, v doplňkovém materiálu pak zahrнула též pět profesionálních rozhlasových mluvčích. Machač (2006) využil nahrávky čtené řeči čtyř mluvčích, kromě toho též semispontánní projevy (vyprávění podle obrázků) blíže neurčeného počtu mluvčích, který se však pohybuje v desítkách. Podobný materiál využila i Homolková (2009), čtená řeč byla opět od čtyř osob, semispontánní vyprávění pak poskytlo 268 mluvčích. Zcela spontánní materiál k replikaci Machačova výzkumu měl k dispozici Šimek (2010), který zkoumal řeč šesti mluvčích v nahrávkách diskusních televizních pořadů. Ondrušková (2011) opět využila čtený text a následně spontánní rozhovor s experimentátorem, vše od šesti mluvčích. Čtené texty analyzoval i Studenovský (2012), a to celkem od 376 mluvčích. Ve studii Podlipského et al. (2009) byly využity také čtené texty, a to od šesti mluvčích. U všech citovaných prací se jedná o materiál sbíraný v rámci Pražského fonetického korpusu (Skarnitzl, 2010).

Psutka et al. (2006) použili rozsáhlý řečový korpus budovaný na katedře kybernetiky Západočeské univerzity v Plzni, který obsahoval 85 hodin spontánní řeči od 340 mluvčích a 75 hodin čtené řeči od 1090 mluvčích.

V případě měření temporálních charakteristik a především trvání je naprosto klíčová otázka segmentace. Spolehlivost údajů o trvání hlásek se nutně opírá o spolehlivost označení jejich hranic, což není problém triviální. Hlásky v řeči za sebou nenásledují zcela lineárně, nýbrž se do velké míry překrývají a často je nemožné určit, kde jedna z nich končí a druhá začíná. Při manuální segmentaci, která je využívána ve fonetickém výzkumu, může hrát velkou roli osoba anotátora, jeho zkušenosti, přesnost a konzistentnost – segmentace různých osob nemusí být přímo srovnatelné a mohly by do výsledků zavádět artefakty. Z důvodu potřeby jednotné koncepce vznikla v roce 2009 práce *Fonetická segmentace hlásek* (Machač & Skarnitzl, 2009), jež poskytuje souhrnná pravidla pro označování hranic segmentů. Novější studie z Fonetického ústavu a studie Podlipského a jeho kolegů (2009) se těmito pravidly řídí, a tedy můžeme s rozumnou mírou jistoty prohlásit tyto údaje za přímo srovnatelné.

Naproti tomu Psutka a kolegové (Psutka et al., 2006) vzhledem k velikosti korpusu určovali trvání hlásek automaticky na základě tzv. nuceného zarovnávání (*forced alignment*) textu ke zvuku, což vedlo k vysoké variabilitě hodnot trvání. Jak ještě uvidíme, směrodatné odchylky jejich trvání jsou řádově dvoj- až trojnásobné než u materiálu segmentovaného ručně.

Přestože automatická segmentace umožňuje zpracovat velké množství dat, vykazuje zároveň nesrovnatelně vyšší chybovost. Proto se v této práci pokusíme využít co největšího množství dat ručně segmentovaných.

V následujících podkapitolách budou shrnuty hodnoty trvání českých hlásek ze starších i novějších studií. Používáme českou transkripci IPA, v tabulkách a grafech jsou z praktických důvodů a kvůli technickým omezením u většiny hlásek ponechány alfabeticke znaky (znamená to, že v textu a v tabulkách se liší pouze [ɪ] – i a [x] – ch; také diftongy a afrikáty nejsou značeny s obloučkem).

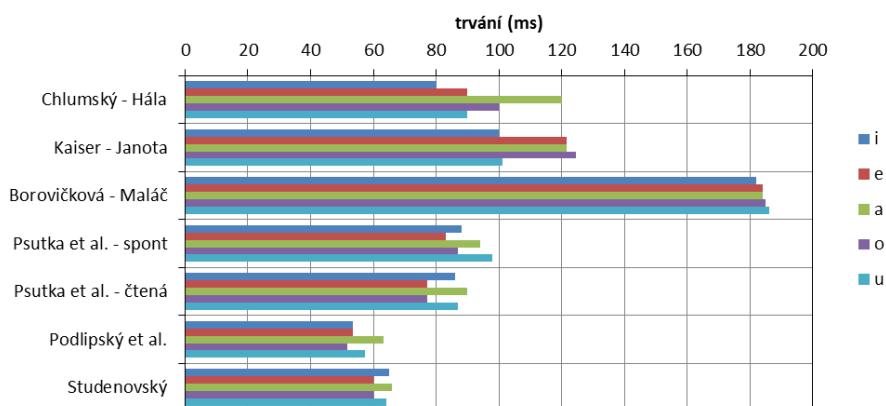
### 5.1.1 Vokály

Český vokalický systém obsahuje pět krátkých a pět dlouhých vokálů, které spolu navzájem stojí ve fonologickém kontrastu, a tři dvojhlasý. Tabulka 5.1 shrnuje průměrné hodnoty jejich trvání z dosavadní literatury, směrodatné odchylky nebo rozptyl.

Krátké	Chlumský - Hála		Kaiser - Janota	Borov. - Maláč	Psutka et al.				Podlipský et al.		Studenovský	
	Prům.	Rozptyl	Průměr	Průměr	Spontánní řeč		Čtená řeč		Průměr	s	Průměr	Rozptyl
í	80	50-100	100	182	88	64	86	70	53,5	19,9	65	24-100
e	90	60-120	121,6	184	83	71	77	50	53,5	16,2	60	33-105
a	120	90-160	121,7	184	94	66	90	68	63,1	20,9	66	30-100
o	100	70-130	124,6	185	87	57	77	46	51,7	15,4	60	50-150
u	90	60-120	101	186	98	71	87	56	57,3	22,9	64	29-110
Dlouhé												
í:	170	140-200	198		108	74	96	59	68,9	36,3	78	35-140
e:	190	160-230	211,6		125	82	122	66	91,8	37,1	118	80-250
a:	240	190-300	244		137	93	128	54	113	26,3	119	40-280
o:	200	160-250	202,6		201	122	132	58	89,3	17,7	109	80-280
u:	180	120-240			118	80	117	70	91,4	49,5	98	40-250
Díftongy												
eu							175	53			110,7	38-239
au					174	74	159	54			125,4	23-257
ou					147	92	127	78	102,8	26,4	109	35-280

Tabulka 5.1: Trvání českých vokálů podle Chlumského (1928, shrnuté v Hálovi, 1962: 190), Kaiserové (1964: 246), Borovičkové a Maláče (1967: 27), Psutky et al. (2006: 58), Podlipského et al. (2009: 134) a Studenovského (2012: 183). Směrodatná odchylka (pokud byla k dispozici) je označena jako s. Hodnoty jsou uvedeny v milisekundách.

Vizuální srovnání jednotlivých hodnot přináší obrázky 5.1a-c.

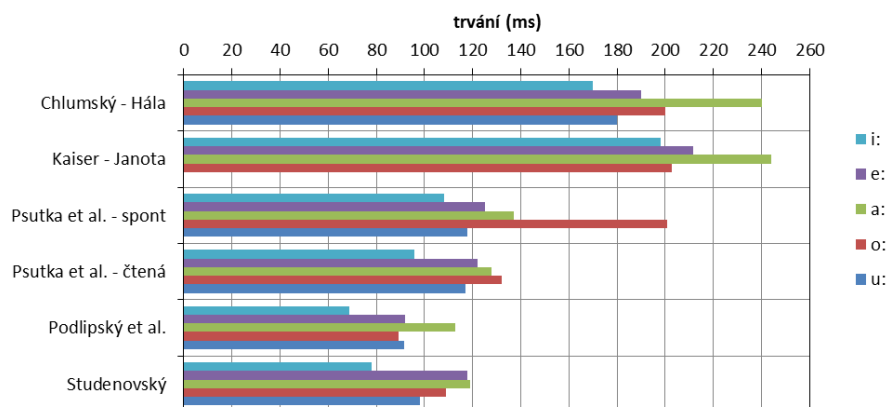


Obrázek 5.1a: Průměrné trvání krátkých českých vokálů podle Chlumského (1928) a Hály (1962: 190), Kaiserové (1964: 246), Borovičkové a Maláče (1967: 27), Psutky et al. (2006: 58), Podlipského et al. (2009: 134) a Studenovského (2012: 183).

Je zřetelné, že hodnoty Borovičkové a Maláče (1967) jsou pravděpodobně velmi ovlivněny typem materiálu (pomale čtení izolovaných logatomů) a ve světle ostatních výsledků se nezdají odpovídat přirozené řeči. Ostatně i sami autoři hypotetizují, že v přirozené řeči bude

trvání vokálů zhruba poloviční (Borovičková & Maláč, 1967: 61) – což skutečně víceméně souhlasí s ostatními výsledky.

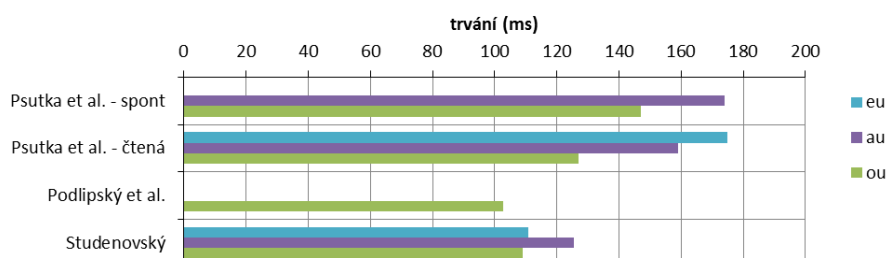
Zvláštní je, že u novějších studií výsledky neodpovídají hypotéze Hály (1962: 183) o inherentním trvání vokálů způsobeném pohyby artikulačního svalstva, kdy nejkratší bývají vysoké. U Psutky et al. (2006) i Studenovského (2012) jsou nejkratší středové. Zdá se však, že nízké [a] si skutečně nejdelsí trvání zachovává konzistentně. Bude zajímavé tyto hodnoty porovnat s výsledky předkládané práce.



Obrázek 5.1b: Průměrné trvání dlouhých českých vokálů podle Chlumského (1928) a Hály (1962: 190), Kaiserové (1964: 246), Psutky et al. (2006: 58), Podlipského et al. (2009: 134) a Studenovského (2012: 183).

Na obr. 5.1b můžeme ještě více než u krátkých pozorovat tendenci ke zkracování dlouhých vokálů směrem k novějším studiím. Zda je to skutečně diachronní trend v češtině, nebo zda je tato tendence pouze artefaktem změny přirozenosti materiálu (případně řečového tempa) u starších oproti novějším studiím, se můžeme bez dalšího výzkumu pouze dohadovat.

Poněkud zvláštní se může zdát hodnota [o:] u Psutky et al. (2006) ve spontánních projevech, nicméně příčinou budou pravděpodobně hezitační [no:], případně [jo:], kde se vokál neúměrně prodlužuje.



Obrázek 5.1c: Průměrné trvání českých diftongů podle Psutky et al. (2006: 58), Podlipského et al. (2009: 134) a Studenovského (2012: 183).

Českým diftongům nebylo dosud věnováno tolik pozornosti jako jednoduchým vokálům, přesto i zde můžeme nalézt referenční hodnoty (viz obr. 5.1c). Nesoulad mezi Studenovským



(2012) a Psutkou et al. (2006) je zapříčiněn zejména diftongem [e̯u]. Důvodem, proč je u Psutky o tolik delší, může být jeho ne zcela vzácná dvojslabičná realizace, kterou Studenovský důsledně vyřazoval, nicméně automatická segmentace ji není schopna zachytit. Podíváme-li se na poměr trvání dlouhých a krátkých vokálů do tabulky 5.2, vidíme, že Chlumským stanovený poměr 2:1 (Chlumský, 1928: 28 nn.) se v současnosti pohybuje spíše kolem 1,7:1, nejmenší přitom konzistentně nalézají novější studie u [ɪ] × [i:], kde je temporální rozdíl podepřen také rozdílem v kvalitě. Podobným procesem by v budoucnu mohla projít také dvojice [u] × [u:] (a poměry trvání u novějších studií tuto hypotézu podporují), kdy krátké se otevírá a přibližuje středovým (Podlipský et al., 2012).

	Chlumský - Hála	Kaiser - Janota	Psutka et al. - spont	Psutka et al. - čtená	Podlipský et al.	Studenovský
i:/i	2,1	2,0	1,2	1,1	1,3	1,2
e:/e	2,1	1,7	1,5	1,6	1,7	2,0
a:/a	2,0	2,0	1,5	1,4	1,8	1,8
o:/o	2,0	1,6	2,3	1,7	1,7	1,8
u:/u	2,0	0,0	1,2	1,3	1,6	1,5

Tabulka 5.2: Poměr trvání dlouhých a krátkých českých vokálů spočítaný z hodnot Chlumského (1928) a Hály (1962: 190), Kaiserové (1964: 246), Psutky et al. (2006: 58), Podlipského et al. (2009: 134) a Studenovského (2012: 183).

Chlumský (1928) zkoumal trvání [a], [a:] v prodlužujících se taktech a konstatoval, že trvání obou hlásek se může zkrátit až na polovinu. Hála (1962: 193) si všímá toho, že dlouhé [a:] v několikaslabičném slově bylo dokonce kratší než krátké [a] v samostatně stojící otevřené slabice, z čehož vyvozuje, že rozdíl mezi krátkými a dlouhými vokály nezávisí na jejich absolutním trvání, ale hodnotíme jej ve vztahu k samohláskám v nejbližším okolí. Chlumský (1928: 116) omezuje toto okolí pouze na krátký úsek, zejména uvnitř jednoho taktu.

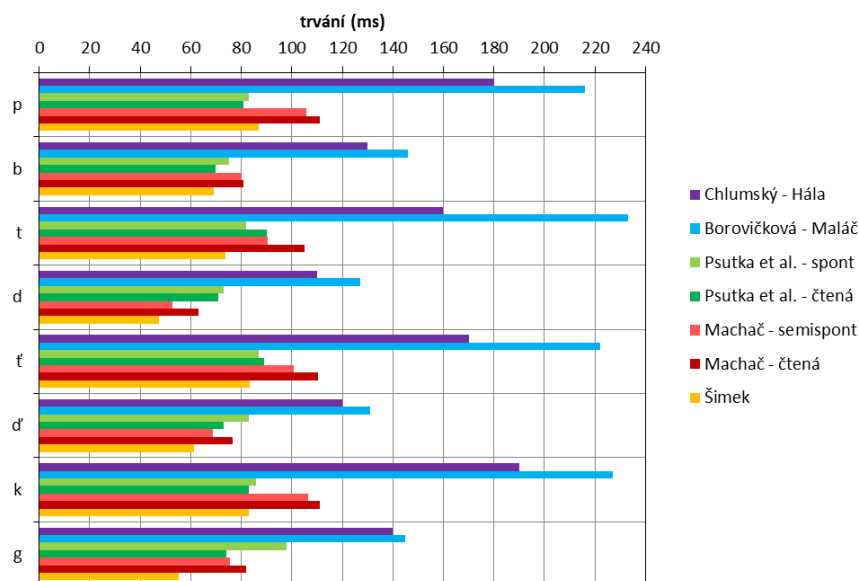
### 5.1.2 Konsonanty

Hála (1962) popisuje charakter konsonantů jako dynamický (oproti statickému charakteru vokálů), který se staví na odpor prodlužování hlásek v čase. A skutečně – u konsonantů můžeme pozorovat o něco nižší variabilitu, a to i v datech Psutky et al. (2006), kde jsou směrodatné odchylky opět několikrát vyšší než u jiných studií. Průměrná trvání českých konsonantů shrnuje tabulka 5.3 níže.

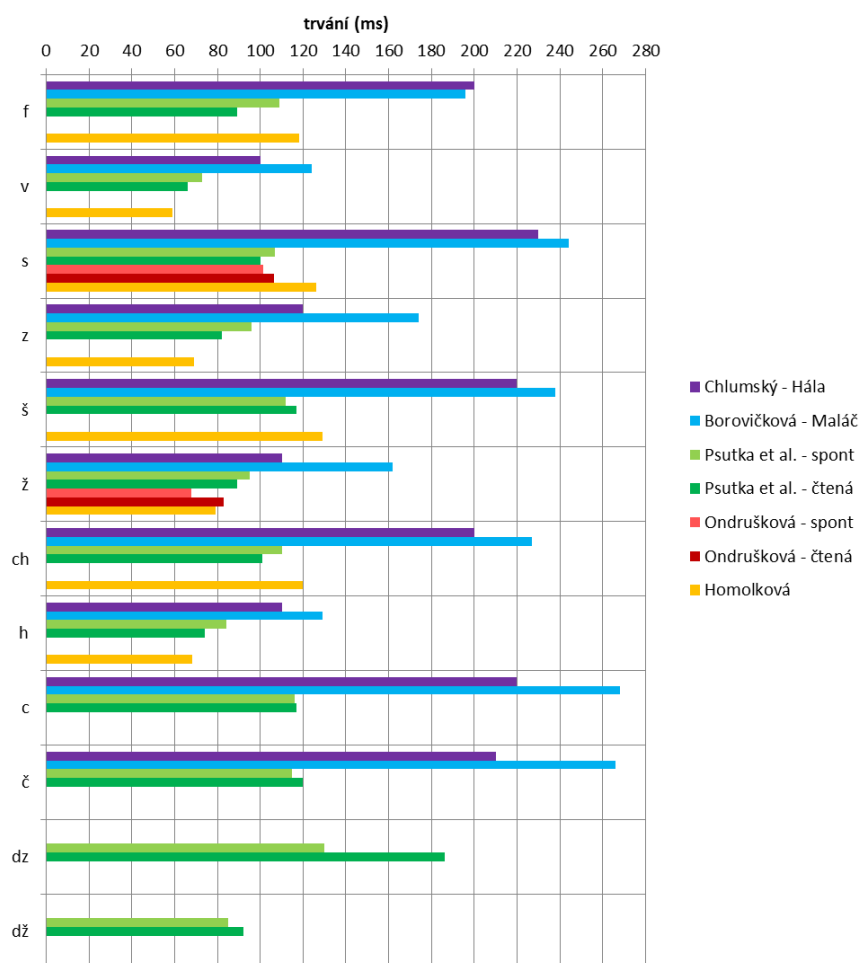
	Chlumský - Hála	Borov. - Maláč	Psutka et al.				Machač				Homolková		Šimek	
	Průměr	Prům.	Spontánní		Čtená řeč		Semispont.		Čtená řeč		Prům.	s	Prům.	s
			Prům.	s	Prům.	s	Prům.	s	Prům.	s			Prům.	s
p	180	216	83	64	81	63	105,6	18,5	111	15,9			86,9	22,2
b	130	146	75	40	70	26	80,3	14	80,8	12,6			69	20,6
t	160	233	82	42	90	71	90,4	21,7	104,9	10,7			73,9	21,6
d	110	127	73	39	71	61	52,7	15,2	62,9	11,9			47,3	17,1
ť	170	222	87	37	89	36	100,8	17,7	110,3	12,2			83,4	23,8
ď	120	131	83	37	73	31	68,8	11,2	76,7	12,1			61,3	19,8
k	190	227	86	33	83	35	106,6	16,5	111	13			83,1	23,1
g	140	145	98	78	74	53	75,4	13,2	81,9	10,5			55,2	21,4
f	200	196	109	77	89	51	<b>Ondrušková</b>				118	23		
v	100	124	73	40	66	28	Spontánní		Čtená		59	15		
s	230	244	107	52	100	52	101,5	27,8	106,6	15,8	126	24		
z	120	174	96	50	82	51					69	14		
š	220	238	112	59	117	94					129	22		
ž	110	162	95	48	89	37	67,6	25,1	82,8	19,5	79	17		
ch	200	227	110	56	101	55					120	24		
h	110	129	84	88	74	40					68	18		
c	220	268	116	52	117	85								
č	210	266	115	42	120	72								
dz			130	97	186	190								
dž			85	109	92	36								
j	90	146	82	57	73	41	26,4	30,4	39,3	15,8				
m	120	161	87	57	89	62								
n	100	158	82	48	74	44	<b>Vernerová</b>							
ň	110	165	78	41	75	39	Neslabičné		Slabičné					
r	30 (70)	138	84	36	79	32	45,1	8,7	74,9	23,8				
ř	120	195	72	34	66	31								
l	70	148	80	48	74	40	34,3	9,4	69,7	21,9				

Tabulka 5.3: Trvání českých konsonantů podle Chlumského (1928) a Hály (1962: 215), Borovičkové a Maláče (1967: 29), Psutky et al. (2006: 58), Vernerové (2006: 43), Machače (2006: 90), Homolkové (2009: 35), Šimka (2010: 40) a Ondruškové (2011: 51). Směrodatná odchylka (pokud byla k dispozici) je označena jako s. Hodnoty jsou uvedeny v milisekundách.

V případě vizualizace hodnot trvání konsonantů bude účelnější seskupit jednotlivé hlásky do skupin. Porovnání jednotlivých zdrojů zobrazují obrázky 5.2a-d, kde jsou konsonanty pro přehlednost rozděleny na explozivny, frikativy a afrikáty, likvidy a ostatní.

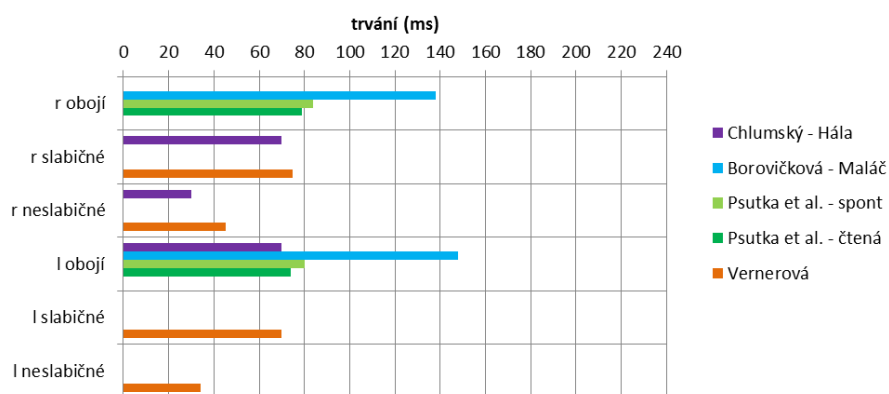


Obrázek 5.2a: Trvání českých exploziv podle Chlumského (1928) a Hály (1962: 215), Borovičkové a Maláče (1967: 29), Psutky et al. (2006: 58), Machače (2006: 90) a Šimka (2010: 40).



Obrázek 5.2b: Trvání českých frikativ a afrikát podle Chlumského (1928) a Hály (1962: 215), Borovičkové a Maláče (1967: 29), Psutky et al. (2006: 58), Homolkové (2009: 35) a Ondruškové (2011: 51).

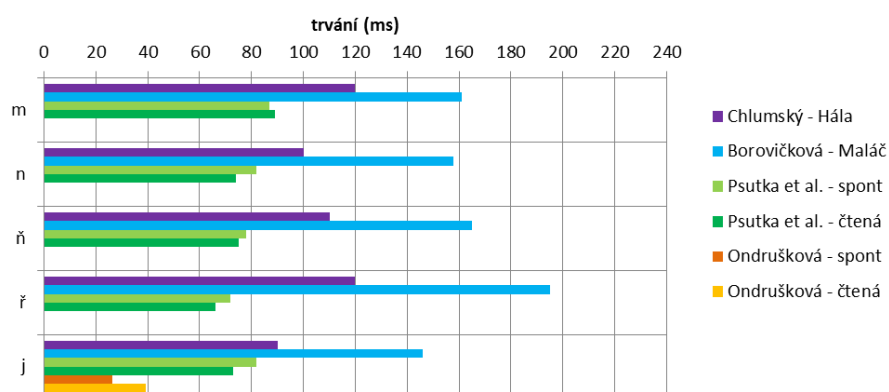
Opět zde vidíme nápadnou výjimečnost hodnot Borovičkové a Maláče (1967), které jsou oproti ostatním přibližně dvojnásobné. Tak je tomu i v případě dalších skupin konsonantů. Psutka et al. (2006) měřili také trvání [dz] a [dž] – z jejich textu se však nedá vyvodit, zda šlo o alofonické varianty [c] a [č] nebo o plnohodnotné fonémy ve slovech cizího původu. Navíc tyto hodnoty vykazují tak vysoké směrodatné odchylky, že je nelze označit za spolehlivé.



Obrázek 5.2c: Trvání českých likvid podle Chlumského (1928) a Hály (1962: 215), Borovičkové a Maláče (1967: 29), Psutky et al. (2006: 58) a Vernerové (2006: 43).

Borovičková a Maláč (1967) ani Psutka et al. (2006) nerozlišovali slabičnou a neslabičnou pozici českých likvid. Hála (1962) uvádí hodnoty pro jedno a vícekmtné [r]. Pro přehlednost jsou zařazeny do stejné skupiny se slabičným a neslabičným [r] Vernerové (2006), jejichž hodnotám relativně odpovídají. Důležité je též podotknout, že Vernerová (2006) uvádí hodnoty normalizované vzhledem k průměrnému artikulačnímu tempu mluvčího.

Na obr. 5.2d jsou porovnány zbylé české konsonanty, tj. nazály, frikativní vibranta [ř] a aproximanta [j].



Obrázek 5.2d: Trvání českých nazál, frikativní vibranty [ř] a aproximanty [j] podle Chlumského (1928) a Hály (1962: 215), Borovičkové a Maláče (1967: 29), Psutky et al. (2006: 58) a Ondruškové (2011: 51).

### 5.1.3 Obecné tendence

Shrňme na tomto místě tendence týkající se trvání českých hlásek, na kterých se citované studie shodují:

- Neznělé konsonanty jsou delší než jejich znělé protějšky (Hála, 1962: 215; Psutka et al., 2006: 58; Machač, 2006: 47 a Šimek, 2010: 40 pro explozivy; Homolková, 2009: 35 pro frikativy); Machač (2006: 48) k tomu ještě dodává, že na tomto trendu se podílí jak závěrová, tak i explozivní část – obě jsou u neznělých delší.
- Nejdější konsonanty jsou neznělé frikativy a afrikáty (Hála, 1962: 215; Psutka et al., 2006: 58).
- Nejkratší explozivy jsou alveolární (Hála, 1962: 215; Machač, 2006: 48).
- Neznělé explozivy mají větší podíl explozivní fáze než znělé (Hála, 1962: 215; Machač, 2006: 48)
- Alveolární a postalveolární frikativy trvají déle než ostatní frikativy<sup>1</sup> (Homolková, 2009: 40; této hypotéze odpovídají i výsledky Chlumského (1928) a Borovičkové a Maláče (1967); u Psutky et al. (2006) není tato tendence tak významná).
- Afrikáty jsou kratší než spojení dvou jim odpovídajících hlásek (Psutka et al., 2006: 59; výsledky Chlumského a Borovičkové a Maláče též odpovídají).

## 5.2 Faktory ovlivňující trvání hlásek

### 5.2.1 Artikulační tempo

Je nasnadě, že se změnou artikulačního tempa se změní také trvání hlásek. Zvyšování artikulačního tempa bude mít na svědomí zkracování segmentů a pravděpodobně též neutralizaci některých temporálních distinkcí (Klatt, 1976: 1210). Avšak tato změna není lineární, existují limity, za které již hlásku zkrátit či prodloužit nelze (Klatt 1976: 1216, Solé & Ohala, 2010). Také výsledky pro češtinu naznačují, že změna tempa postihuje některé hlásky více než jiné. Kaiserová a Janota (1964: 247) zkoumali trvání českých vokálů při čtení normálním a maximálně rychlým tempem a došli k závěru, že zrychlení zkrátilo více dlouhé než krátké vokály.

Machač (2006), Šimek (2010) i Homolková (2009) korelovali trvání exploziv, respektive frikativ se změřeným artikulačním tempem a došli shodně k závěru, že korelace je maximálně středně významná, a to hlavně u neznělých, znělé se tempu přizpůsobují jen málo. Machač (2006: 117) i Šimek (2010: 98) konstatují, že největší odolnost vůči tempu (tedy nejnížší

---

<sup>1</sup> Toto zjištění je obzvláště zajímavé ve srovnání s tendencí u exploziv, která je opačná.

korelaci) vykazuje trvání explozivy [g]. Podobně označuje Ondrušková (2011: 138) za stabilní vůči změnám tempa i frikativu [s].

Většina temporálních modelů (nebo modelů, které nějak zahrnují trvání segmentů) si změn artikulačního tempa nevšímá a předpokládá, že tempo je v rámci věty neměnné, což způsobuje např. menší přirozenost syntézy, či menší úspěšnost rozpoznávače řeči (Carlson, 1991: 244; O'Shaughnessy, 1995: 600; v češtině tento efekt ověřila Churaňová, 2013: 15).

Výsledky Dankovičové (2001) naznačují, že v češtině se variace tempa řeči projevují v rámci prozodické fráze.<sup>2</sup> Dankovičová zjistila výrazný efekt zpomalování v rámci prozodické fráze směrem k jejímu konci, a to ve dvou studiích u sedmi různých mluvčích. V rámci klauze a úseku mezi dvěma pauzami se tento efekt v jejích datech neprojevil. Zde je nutné podotknout, že klasifikace prozodické fráze není zcela neproblematickou záležitostí, jak by se mohlo zdát z definice Dankovičové (2001: 27), která je převzata od Palkové (1994: 162–163), a důkladnější popis akustických charakteristik prozodických předělů různé hloubky stále ještě čeká na budoucí výzkum.

### 5.2.2 Přízvuk

Již Chlumský (1928) se zabýval vlivem přízvuku a délky slova nebo taktu na trvání vokálů. Shrnuje, že v češtině přízvuk výrazný vliv nemá, pouze nepřízvučná pozice umožňuje zkracování fonologicky dlouhých vokálů (Chlumský, 1928: 92), naopak přízvučná spíše zachovává jejich délku.

Přestože se zdá, že český slovní přízvuk není standardně realizován změnou trvání hlásek (Janota & Palková, 1974), co se týče kontrastivního či emfatického přízvuku na větné úrovni, je v češtině realizován lokálním zpomalením (Volín, 2009). Také zahraniční výzkum podporuje hypotézu, že slova se pod kontrastivním přízvukem prodlužují. Niebuhr (2010) našel v němčině (nejen) temporální značení pozitivního a negativního emfatického přízvuku, kdy v prvním případě se prodlužuje vokál a ve druhém konsonant přízvučné slabiky – zároveň se v těchto případech projevila i kompenzace, kdy konsonant se prodlužoval na úkor délky vokálu a naopak (Niebuhr, 2010: 185).

### 5.2.3 Temporální kompenzace

Zjednodušeně řečeno, pojem temporální kompenzace zahrnuje jevy, kdy segmenty svým trváním ovlivňují segmenty sousedící. Delší hlásky vyvolávají zkrácení svého okolí a naopak, celkové trvání vyšší jednotky (např. slabiky) zůstává stejné.

---

<sup>2</sup> Intonační fráze v terminologii Dankovičové (2001), promluvový úsek v terminologii Palkové (1994).

Tento jev je dobře zdokumentován v mnoha jazycích (Port et al., 1980), v češtině si jej také někteří autoři všímají. Chlumský (1928: 103) uvádí, že skupina konsonantů následující po vokálu způsobí jeho zkrácení; nachází kompenzaci i v opačném směru – po dlouhém vokálu jsou konsonanty kratší než po krátkém (Hála, 1962: 215). Tuto tendenci však nepotvrzuje Machač (2006: 108) pro explozivy, na druhou stranu nachází souvislost mezi délkou explozivy a mírou otevřenosti vokálu – před nízkým vokálem je exploziva v průměru kratší než před vysokými. Borovičková a Maláč (1967: 61) zmiňují též vliv znělosti konsonantu – vokál následující pro znělém (tedy kratším) konsonantu má tendenci být delší než po neznělém. Tentýž výsledek pro frikativy uvádí i Homolková (2009: 95), a to i v opačném pořadí, kdy následuje frikativa po vokálu, a tedy není součástí stejné slabiky. Stejně jako Machač nenachází Homolková vliv opačný, tedy vliv fonologické délky vokálu na trvání frikativy.

Studenovský (2012: 118) pozoruje kompenzaci u trvání diftongů v závislosti na struktuře přetury – diftong následující po dvou konsonantech je kratší než pokud mu předchází jen jeden.

Kompenzací můžeme nazvat i přizpůsobování trvání segmentu délce vyšší jednotky. Dankovičová (2001: 51 nn.) nachází ve svém materiálu významný efekt délky slova vyjádřené počtem jeho slabik – čím delší slovo, tím kratší je trvání hlásek uvnitř něj a tedy vyšší artikulační tempo. Dále však spekuluje, že příčinou by kromě kompenzace mohla být též tendence češtiny k jednodušší slabičné struktuře u delších slov. Průměrné hodnoty artikulačního tempa různě dlouhých slov nalezneme v tabulce 5.4.

Počet slabik ve slově	Průměrné artikulační tempo (sl/s)
1	3,83
2	6,14
3	6,65
4	7,07
5	7,51 ve 3slovných frázích 8,13 ve 4slovných frázích

Tabulka 5.4: Průměrné artikulační tempo ve slabikách za sekundu jednoslabičných a víceslabičných českých slov podle Dankovičové (2001: 96).

Lze tedy říci, že tendence k temporální kompenzaci se v češtině zcela jistě uplatňují, nicméně výsledky jsou zatím kusé a nesystematické. Bylo by jistě zajímavé tento problém pojmout uceleně, též s ohledem na percepční stránku věci.

#### 5.2.4 Vliv pozice v rámci vyšší jednotky

Většina zmiňovaných českých autorů též zkoumala vliv pozice hlásky v rámci mluvního taktu na její trvání. Borovičková a Maláč (1967: 61) konstatují, že hlásky jsou nejdelší ve finální pozici v taktu. Tuto hypotézu ověřovali dále ještě Šimek (2010: 81) pro explozivy (vzhledem k nedostatku materiálu ale tuto tendenci pozoroval pouze v případě [t] a [d]) a Vernerová (2006: 47) pro slabičné likvidy (u neslabičných se rozdíl mezi pozicemi neprojevil). Machač (2006: 87) i Šimek (2010: 80) pozorují u exploziv tendenci k postupnému krácení směrem od začátku mluvního taktu – tj. iniciální exploziva je delší než každá další (s výjimkou finálních). Tuto tendenci našla u některých frikativ i Homolková (2009: 50–51), nicméně hodnotí ji jen jako slabou. Opět konstatujeme, že tato tematika je v češtině prozatím zpracována neuspokojivě.

Co se týče pozice hlásky ve vyšší prozodické jednotce, než je takt, mnoho literatury je věnováno závěrovému zpomalování (např. Klatt, 1976: 1211 n., 1219). Z výsledků Dankovičové (2001) lze vyvodit, že v češtině je doménou závěrového zpomalování prozodická fráze. Nicméně míra zpomalování by v češtině mohla být závislá na mluvčím. Dankovičová (2001: 80 a 93) nachází variabilitu ve stupni zpomalování mezi čtyřmi neprofesionálními mluvčími – porovnávala artikulační tempa slov v různých pozicích v prozodické frázi. Stejný efekt byl nalezen i ve studii Weingartové (2013) u tří profesionálních mluvčích ve čteném projevu, kde byla ovšem využita Volínova (2009) metoda měření lokálního artikulačního tempa podle vzdálenosti dvou po sobě následujících vokálních jader.

#### 5.2.5 Vliv mluvního stylu

Další zajímavý vhléd do problematiky trvání hlásek poskytuje srovnání různých mluvních stylů. Ač je třeba dbát opatrnosti při porovnávání studií od různých autorů prováděných na různém materiálu, je možné zde vyčíst obecnou tendenci k větší variabilitě hodnot trvání při spontánnějších projevech (Psutka et al., 2006: 59; Machač, 2006: 96; Homolková, 2009: 65; Ondrušková, 2011: 74).

Co se týče průměrně delšího či kratšího trvání hlásek v některém z typů projevů, nejsou výsledky jednotné – Psutka et al. (2006: 59) uvádějí průměrně delší hodnoty ve spontánních projevech, Machač (2006: 91) a Šimek (2010: 55) nalézají u exploziv opačnou tendenci, která pravděpodobně odráží rozdíly v artikulačním tempu. Rovněž výsledky zahraničních studií nejsou zcela konzistentní, pro diskusi viz Weingartová & Volín, 2014.



### 5.2.6 Vliv kategorie slova

Vztah mezi kategorií slova a jeho artikulačním tempem, případně trváním jeho segmentů je také předmětem zkoumání. Obvykle se slova dělí na autosémantická (*content words*) a synsémantická (*function words*), přičemž autosémantická bývají vyslovována pomaleji (O'Shaughnessy, 1995: 602; Dankovičová, 2001: 97). Určité předběžné výsledky na toto téma můžeme nalézt i ve studii Volína a Weingartové (2012), ze kterých vyplývá, že synsémantická slova se zdají být méně stabilní a podléhají více individuálnímu chování mluvčího než autosémantická. Souvislost s touto problematikou lze také hledat v oblasti textových frekvencí slova – synsématicnost je často spojena s vyšším výskytem, zatímco autosémantických slov je méně, a mnohá z nich nejsou tím pádem častá. Dá se očekávat, že slovo méně obvyklé bude vysloveno pomaleji, zatímco slovo vysoce frekventované bude vysloveno rychleji.

Další vlivy na trvání hlásek nebyly pro češtinu dosud systematicky zkoumány. Patří mezi ně zejména faktory týkající se mluvčího (nálada či afektivní stav, fyziologický stav, řečové vady) a posluchače (tedy percepce rytmičnosti promluvy).

## 6 Výzkumné otázky a cíle práce

Prvním z cílů práce je podrobně popsat temporální charakteristiky českých hlásek a vliv různých faktorů na tyto charakteristiky. Deskripce trvání hlásek a jejich změn pod vlivem segmentálních a prozodických faktorů pak bude základem pro vytvoření temporálního modelu češtiny. Druhým z cílů je prozkoumat temporální charakteristiky, rozsah jejich individuality a možnosti využití pro identifikaci mluvčího.

Využitím některých temporálních charakteristik pro rozpoznávání mluvčího se již zabýval článek *Rhythm metrics for speaker identification in Czech* (Weingartová, 2013), kde bylo prozkoumáno devět globálních temporálních ukazatelů a jeden lokální. Ukazatele globální se neukázaly jako příliš vhodné pro postižení individuálních rozdílů, z důvodů zmíněných v předchozích kapitolách. Naopak lokální ukazatel LAR (Volín, 2009), který zachycuje vzdálenosti slabičných jader, byl výrazně úspěšnější. Při aplikaci na koncové slabiky úseků dokázal rozpoznat různé způsoby závěrového zpomalování u tří mluvčích. Jedna z našich hypotéz tedy předpokládá, že kontury lokálního artikulačního tempa dokážou systematicky postihnout individualitu mluvčího v temporální struktuře jeho řeči.

Možnosti tohoto ukazatele – a lokálního pojetí rytmu vůbec – jsou však mnohem širší. LAR lze využít nejen k celkovému popisu průběhu temporálních modulací řeči u jednotlivých mluvčích, ale také k vytvoření průměrných vzorců pro temporální průběhy vět v češtině (podobně jako např. u Volína a Skarnitzla, 2007). Odchyly od těchto průměrů by pak mohly sloužit jako charakteristiky jednotlivých mluvčích (či jejich skupin).

Prvním krokem v této práci bude extrakce průměrných hodnot trvání jednotlivých hlásek z dostupného materiálu (viz níže) a jejich porovnání s hodnotami uvedenými ve starší literatuře. Tyto hodnoty mohou být využity jednak jako referenční, jednak například pro vylepšování řečové syntézy či rozpoznávání řeči. A právě tyto průměrné hodnoty pak zde budou sloužit jako vstupní data do modelu temporálních charakteristik.

Dále je naším cílem modelovat kontextuální vlivy na trvání hlásek – umístění v prozodické struktuře, délka nadřazené jednotky, struktura slabiky, hláskové okolí, atd. Všechny tyto faktory by podle výsledků předchozích českých i zahraničních studií měly trvání hlásek nenáhodně ovlivňovat. Po vložení všech těchto pravidel do modelu budou výsledkem průměrné kontury zacházení s temporálními charakteristikami (závislé na textu, nikoliv na mluvčích). Po aplikaci modelu na řečový materiál vzniknou rezidua (tedy temporální odchylky, které model nevysvětluje). Podle našich předpokladů by měla poukázat na individuální temporální vzorce.

Využitým materiálem je korpus Minidialogy-H – krátké hrané dialogy, které přestože byly nahrávány ve studiu Fonetického ústavu, vykazují vysokou míru spontaneity. Korpus obsahuje 24 ručně segmentovaných dialogů od 34 mluvčích.

Důraz je především kladen na spolehlivou manuální segmentaci, která by měla zajistit co nejpřesnější extrakci hodnot trvání. Automatické nucené zarovnávání, jehož důsledkem je výrazně vyšší chybovost, tedy využíváme pouze v prvních fázích úpravy dat, dále hranice mezi segmenty upravujeme ručně.

## 7 Metoda a materiál

### 7.1 Korpus Minidialogy-H a jeho anotace

Zvukový materiál využitý v této práci pochází z korpusu *Minidialogy*, subkorpusu *H*, který je v současné době stále rozvíjen ve Fonetickém ústavu FF UK v Praze. Celý korpus obsahuje nahrávky 50 mluvčích, kteří ve dvojicích četli krátké dialogy sestávající obvykle z 5 replik. Mluvčí byli s textem dialogů předem seznámeni a měli je pronést co nejpřirozeněji. Pokud došlo k dysfluencím či nepřirozeným realizacím, byli mluvčí požádáni experimentátory o další opakování dialogu.<sup>3</sup> Každý dialog byl později (jiný den) přečten ještě jednou s mluvčími v opačných rolích; každá replika je tak pronesena všemi mluvčími. Texty obsažené v korpusu *Minidialogy-H* tvoří přílohu č. 1.

Nahrávky byly pořízeny ve studiu Fonetického ústavu, s využitím elektretového mikrofonu IMG ECM2000, zvukové karty SB Audigy 2ZS, při vzorkovací frekvenci 32 kHz a kvantizaci 16 bitů. Celý dialog byl posléze rozstříhán na jednotlivé repliky.

V této práci je využita část korpusu *Minidialogy-H*, která je v tuto chvíli zpracována nejdůkladněji, a to nahrávky od 34 mluvčích, z toho 26 žen a 8 mužů ve věku 20–25 let. Další údaje o mluvčích nejsou bohužel k dispozici, neboť v době, kdy byl tento korpus pořizován, bylo nahrávání striktně anonymní. Jeden z mluvčích vykazuje řečovou vadu, k čemuž jsme při analýzách přihlédlí.

Přečtených dialogů je celkem 24, přičemž každý kromě jednoho obsahuje pět replik (zbývající dialog má pouze čtyři). Dialogy jsou seskupeny po čtyřech a každý ze čtveřice obsahuje jako třetí repliku cílovou větu, která se liší pouze tvarem přísudku (v jednotném nebo množném čísle) a modalitou věty (oznamovací nebo tázací).

Cílové věty na příkladu čtveřice dialogů H1 tedy vypadají takto:

H1a: *Řekneš jim, co si myslíš.*

H1b: *Řekneš jim, co si myslíš?*

H1c: *Řeknete jim, co si myslíte.*

H1d: *Řeknete jim, co si myslíte?*

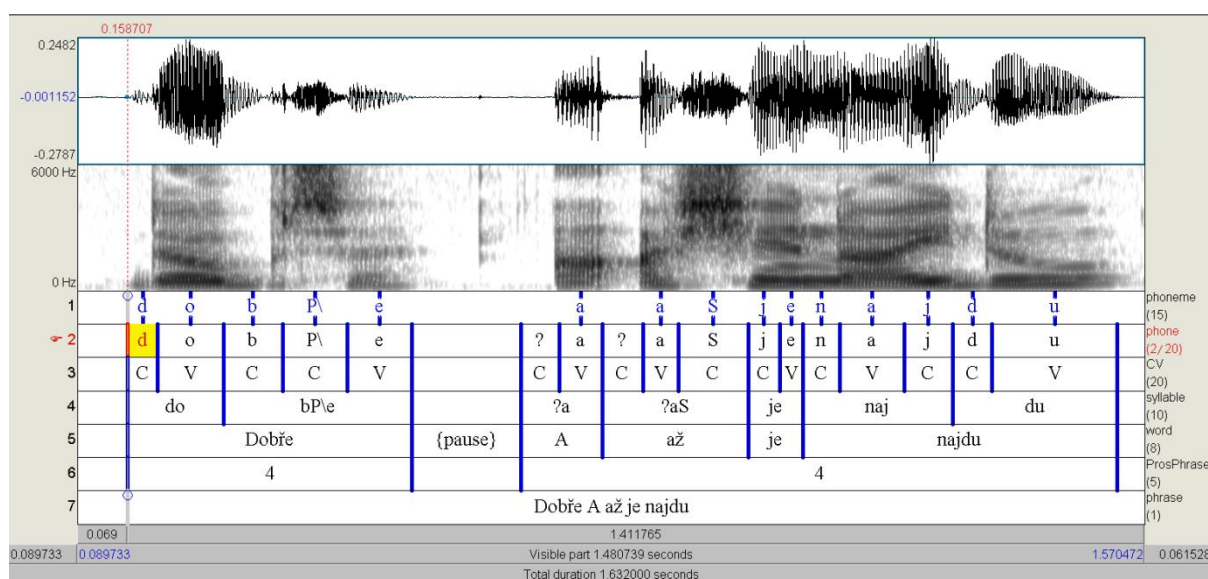
Kódování replik vypadá následovně: První dva znaky určují čtveřici dialogů (H1–H6), třetí znak zastupuje jeden dialog ze čtveřice, tj. typ cílové věty (a–d), pátý znak je pořadové číslo repliky v dialogu a sedmý až desátý znak je kód mluvčího. Tedy například kód H6c\_3\_FEJA označuje třetí repliku z dialogu H6c (tj. oznamovací větu s přísudkem v množném čísle)

<sup>3</sup> I přes to obsahuje korpus určité množství dysfluencí, heziticí a podobných jevů, viz oddíl 7.1.5 Vrstva slov.

vyslovenou mluví FEJA. Kódy mluvích jsou vždy čtyřpísmenné, ženy končí na A, muži na souhlásku.

Text korpusu obsahuje celkem 119 unikátních replik. Vynásobíme-li to počtem mluvích, dostaneme celkový počet replik analyzovaných v této práci, což je 4046; analyzovaných slov je přes 27 000. Řečový materiál trvá celkem déle než dvě hodiny.

Nahrávky jsme ručně označovali v programu *Praat* (Boersma & Weenink, 2014) na úrovni prozodických frází, slov, slabik, typů hlásek, hlásek a fonémů.<sup>4</sup> Příklad anotovaného TextGridu ukazuje obrázek 6.1. Jednotlivým vrstvám budou věnovány následující oddíly.



Obrázek 7.1: Příklad anotované repliky H1a\_2\_PCNA z korpusu Minidialogy-H v editačním okně programu *Praat*. Zvuk je označován na vrstvě fonémů, hlásek, typů hlásek, slabik, slov a prozodických frází (shora).

### 7.1.1 Vrstva hlásek

Vrstva slov a hlásek byla vytvořena automaticky programem *Prague Labeller* (Pollák et al., 2007), který využívá algoritmus nuceného zarovnávání založený na skrytých Markovových modelech (HMM). Hranice však musely být manuálně opraveny (s využitím principů a kritérií uvedených v knize Machače a Skarnitzla, 2009). Vodítkem pro umístění hranic byl zejména největší nárůst či úbytek energie v amplitudě oscilogramu, v případě vokálů začátek či konec formantové struktury ve spektrogramu. U obtížně určitelných míst byla vzhledem k omezením daným lidským sluchem a požadavku co největší konzistentnosti (jak v rámci

<sup>4</sup> Ruční značkování prováděly kromě nás částečně též jiné osoby, které však byly námi proškoleny a jejich práce následně kontrolována. Seznam těchto osob viz poděkování. Za automatické vytvoření vrstvy slabik a fonémů děkujeme Tomáši Bořilovi.

jednoho anotátora, tak také mezi anotátory) dávána přednost vizuálním vodítkům před sluchovými. Pokud chyběla i ta, byla hranice umístěna do poloviny přechodové oblasti.

Rovněž byla manuálně opravována transkripce hlásek (příčemž vrstva slov zůstala v ortografické podobě, jak je vidět z obrázku 7.1). Z technických důvodů jsme zachovali transkripci SAMPA, která je výstupem Prague Labelleru a obsahuje pouze ASCII znaky. Do budoucna ale předpokládáme konverzi transkripce celého korpusu do znaků české IPA.

Při transkripci vrstvy hlásek byla vzata v úvahu asimilace znělosti i místa tvoření (např. ve slovech *pořádně*, *hodně* či *trumf*), desonorizace (častá zejména u iniciálních slov začínajících na [z]) a další změny hlásek (např. změna [h] na [d] ve slově *tohle*). Desonorizace nebyla značena pouze u [h] ani [ř], jelikož tento kontrast není v češtině funkční, neznělé protějšky obou hlásek nemají distinktivní platnost.<sup>5</sup> Všechny tyto jevy byly tedy transkribovány podle percepčního hodnocení jednotlivých realizací. Pokud byl výrazně změněn některý z rysů definujících hlásku, tj. znělost, místo nebo způsob artikulace, a tato změna vedla k percepčnímu dojmu jiné hlásky, byla transkripce podle toho upravena.

Při značkování slov s kolísající vokalicou kvantitou (např. *jo*, *ne*, *tím*, *nevím*, *není*), jsme se opět řídili percepčním dojmem, přičemž u nejasných případů byla přednost dána ortoepické formě.

Při transkripci jsme nicméně přihlíželi k fonologickému systému češtiny a hlásky, které nejsou v tomto systému obsaženy, jsme transkribovali jejich nejbližším českým ekvivalentem (to se týkalo zejména posunů v kvalitě vokálů).

Oproti zvyklostem u jiných korpusů budovaných na Fonetickém ústavu jsme se u tohoto korpusu rozhodli nezarovnávat hranice na průchod zvukové vlny v oscilogramu nulou, protože by tato změna mohla deformovat trvání hlásek, obzvláště v případě míst s velmi nízkou základní frekvencí, kdy místa protnutí vlny a nulové amplitudy jsou od sebe dost vzdálena (tento postup je též doporučen v manuálu Machače a Skarnitzla, 2009: 24).

### 7.1.2 Vrstva fonémů

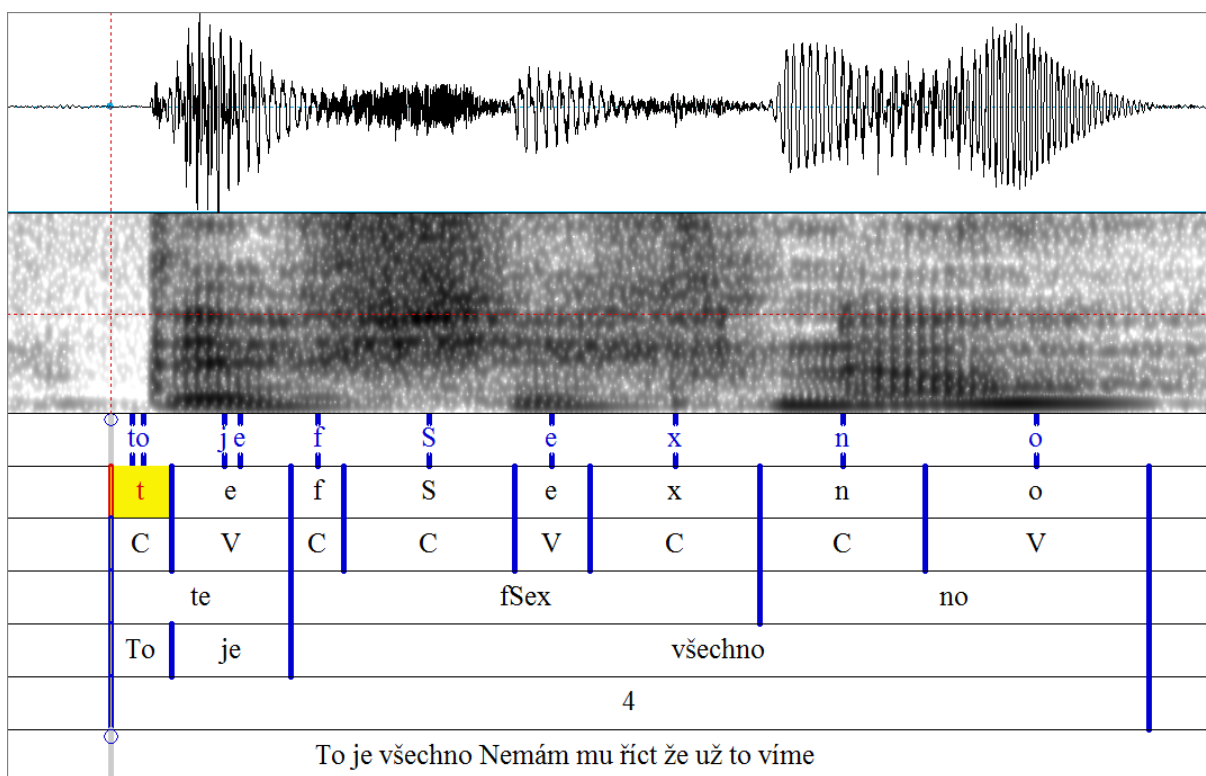
K zachycení asimilací, desonorizací či elizí hlásek byla potřeba vrstva fonémů, pro všechny mluvčí totožná – tedy nezávislá na konkrétní realizaci. Do TextGridů byla přidána automaticky na základě zdrojového souboru, který tvoří přílohu č. 2. Při vytváření vrstvy fonémů jsme učinili několik koncepčních rozhodnutí – u slov *no*, *jo*, nebo *ne* byl jako foném ponechán krátký vokál na základě ortografie; u tvarů slovesa *být* s *j* na začátku bylo /j/

---

<sup>5</sup> Toto řešení bylo zvoleno jako nejeftektivnější vzhledem k tématice práce, do níž tyto jevy přímo nespádají. Desonorizaci českých hlásek se věnovali např. Machač (2008) a Skarnitzl (2011: 202ff.).

ponecháno jako foném. Co se týče asimilací znělosti, bylo postupováno podle ortoepických pravidel (Palková, 1994). U obstruentů před sonorami, kde by mohlo dojít k tzv. moravské asimilaci (například *snad nebudu*), byla jako foném ponechána neznělá varianta. V případě asimilací přes hranici slova, kde bylo lze předpokládat prozodický předěl (tedy například *myslíš, že*), jsme asimilaci na vrstvě fonémů neznačili. Ve slovech *priorita* a *idiot* bylo /j/ ponecháno jako foném. Zvláštní kategorii pak tvoří částice *hm*, které jsme na této vrstvě přiřadili vlastní značku.

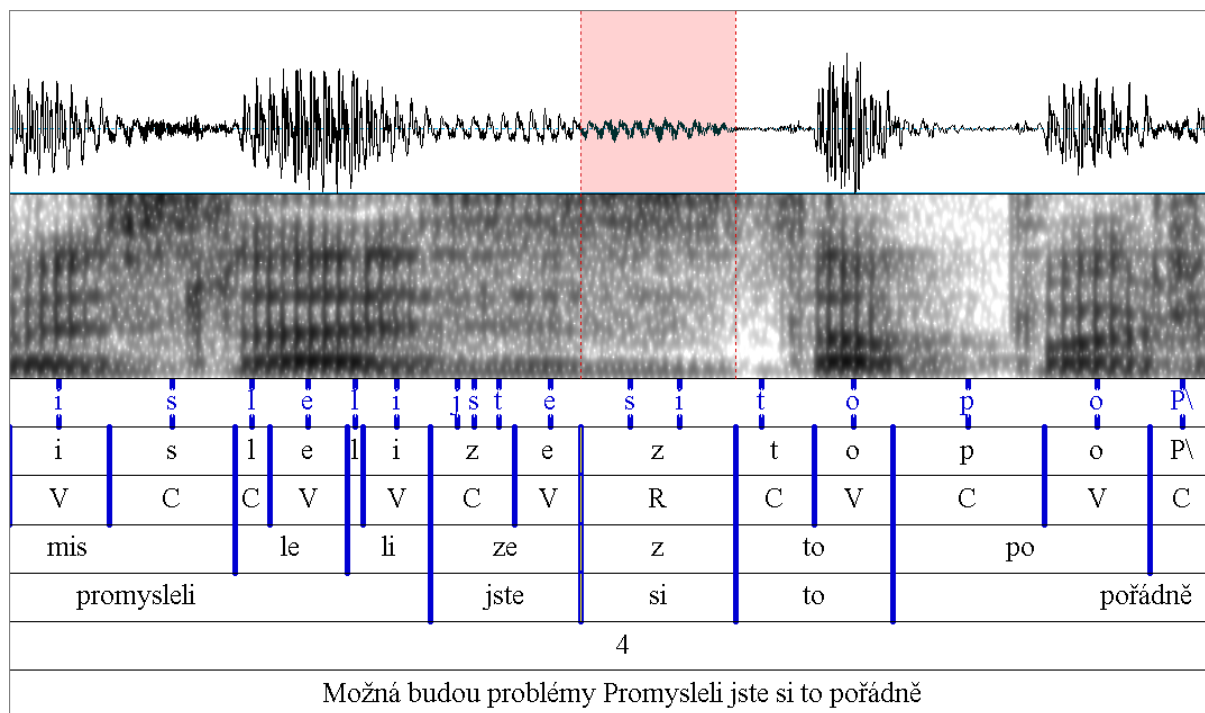
Značkování elizí probíhalo tak, že elidovaný segment byl z vrstvy hlásek zcela smazán, nicméně jemu příslušející značka fonému byla přesunuta nad sousední hlásku patřící do stejné slabiky. Příklad je zobrazen níže (obrázek 7.2).



Obrázek 7.2: Příklad elize hlásek [o] a [j] v replice H2a\_4\_DANA z korpusu Minidialogy-H. Fonémy příslušející elidovaným hláskám v TextGridu zůstaly a byly přesunuty nad sousední hlásky stejné slabiky. Na vrstvě slabik byla pak následně provedena resylabifikace řetězce *to-je* na *te*.

### 7.1.3 Vrstva typů hlásek

Skriptem byla v programu *Praat* vytvořena vrstva, ve které jsou označeny konsonanty jako C, vokály jako V a slabičné konsonanty jako R. Vzhledem k občasným elizím se v korpusu vyskytují kromě standardních slabičných konsonantů /r/, /l/ a /m/ též jiné – příklad ukazuje obrázek 7.3, kde v důsledku elize vokalického jádra slabiky vzniklo slabičné [z].



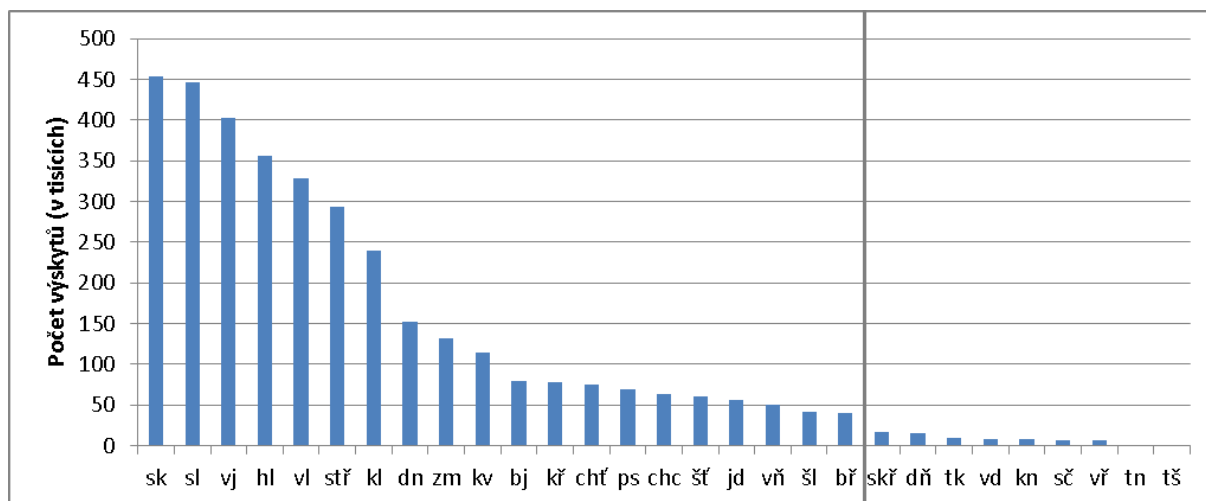
Obrázek 7.3: Příklad slabičného [z] v replice H1d\_1\_MORC z korpusu Minidialogy-H vzniklého elizí jádra slabiky.

#### 7.1.4 Vrstva slabik

Vrstva slabik byla podobně jako vrstva fonémů vytvořena automaticky na základě vrstvy hlásek a dané sylabifikace. Zdrojový text s vyznačenou sylabifikací lze najít v příloze č. 3. Rozčlenění slov na slabiky ale není vždy triviální – problémy tvoří konsonantické shluky, které mohou být buďto celé součástí prétury, nebo se rozdělí mezi préturu a kodu předchozí slabiky. V nejasných případech bylo postupováno podle Kučery a Monroea (1968: 48–49) a přihlédnuto k frekvenci shluku na začátku slova podle korpusu současné češtiny SYN2005 (Bartoň et al., 2009).

Obrázek 7.3 ukazuje počty výskytů zkoumaných shluků na začátcích slov, přičemž hranice byla stanovena mezi shluky *skř* a *bř*, kde dochází ke skokovému více než zdvojnásobení počtu výskytů (z 16 674 na 40 166 výskytů). Shluky nad touto hranicí byly tedy brány jako prétury, shluky pod touto hranicí byly rozděleny mezi préturu a kodu předcházející slabiky.





Obrázek 7.4: Počet výskytů vybraných konsonantických shlučků v korpusu Minidialogy-H na začátku slova v korpusu současné češtiny SYN2005 z celkového počtu 122 419 948 pozic (Bartoň et al., 2009: 16). Ostatní shlučky mají v korpusu SYN2005 výrazně vyšší počet výskytů a nejsou zobrazeny.

Text v příloze také počítá s ortoepickou výslovností hlásek bez elizí a jiných změn v promluvách. Místa, která neodpovídala slabikám určeným ve zdrojovém textu, jsme prošli jedno po druhém a opravili podle skutečnosti (a percepčního dojmu slabik). Do resylabifikace jsme se však nepouštěli v případech, kdy by vznikl nestandardní hláskový shluč. Tedy například zmizelo-li [d] ve slově *bude*, byly ponechány dvě slabiky *bu-e*, protože diftong /ue/ není v českém inventáři. K resylabifikaci přes hranici slova jsme se uchýlili opět pouze v případě, kdy by elize způsobila nestandardní slabiku (viz např. obrázek 7.2 na str. 39, kdy by bez resylabifikace zbylo slabičné [t]).

### 7.1.5 Vrstva slov

Na vrstvě slov jsou kromě ortografického přepisu slov zachyceny také informace o pauzách, dysfluencích, hezitacích a příbuzných jevech. Využité značky jsou tyto:

{pause}: tichá pauza

{breath}: nádechová pauza

{dsfl}: dysfluence (označuje buďto dysfluentně realizované slovo, nebo dysfluentní část výpovědi, která není jasnou součástí některého slova)

{hesit}: hezitace (označuje slovo, které obsahuje prodloužení některé hlásky jako projevu váhání; samostatná hezitace se v korpusu nevyskytuje)

{wdch}: změna textu (mluvčí přečetl jiné slovo, než měl předepsáno, případně změnil pořadí slov)

Oproti obecně přijímaným zvyklostem v literatuře a též oproti pokynům platným pro jiné korpusy budované na Fonetickém ústavu (viz např. Weingartová et al., 2014) byly pauzy na prozodických předělech označovány i v případech, kdy jejich trvání nedosahovalo 120 ms. Rozhodnutí značit i kratší pauzy bylo učiněno vzhledem k tomu, že cílem práce je získat co nejpřesnější údaje o trvání hlásek, a i velmi krátká pauza rozdělená mezi okolní hlásky by kontaminovala jejich trvání. Také je třeba podotknout, že celý korpus byl segmentován ručně, a tedy nemohlo dojít k označení např. dlouhého závěru neznělé explozivы jako pauzy, což je jeden z typických problémů automaticky segmentovaných korpusů.

### 7.1.6 Vrstva prozodických frází

Na této vrstvě jsou vyznačeny začátky a konce prozodických frází a síla předělů. Využita je indexace předělů podle systému ToBI (viz např. Beckman & Elam, 1997), a to index 3 a 4. Předěl síly 4 označuje plnohodnotný konec prozodické fráze, signalizovaný temporálními i intonačními prostředky (obvykle zpomalením či pauzou a koncem melodému). Pokud jeden z těchto signálů chybí, je využit index 3. Malé *p* připojené k indexu indikuje prozodickou neplynulost, zadržnutí.

## 7.2 Extrakce dat

Informace zjišťované u každé položky jsou následující:

- Replika (kód)
- Mluvčí (kód)
- Hláška
- Typ hlásky (C pro konsonanty, V pro vokály a R pro slabičné konsonanty)
- Trvání hlásky (v ms)
- Trvání hlásky vzhledem k trvání slabiky (v %)
- Trvání hlásky vzhledem k trvání slova (v %)
- Předcházející hláška
- Následující hláška
- Foném(y) přináležející hlásce
- Slabika, v níž se segment nalézá
- Struktura slabiky (např. CVC)
- Pozice hlásky ve slabice (prétura/nukleus/koda)
- Trvání slabiky (v ms)
- Pozice slabiky ve slově (iniciální/mediální/finální/individuální)

- Pozice hlásky ve slově (iniciální/mediální/finální/individuální)
- Slovo
- Trvání slova (v ms)
- Délka slova ve slabikách (kanonická, elize nebereme v úvahu)
- Délka slova ve fonémech (kanonická, elize nebereme v úvahu)
- Pozice slova v prozodické frázi (iniciální/mediální/finální/individuální)
- Délka prozodické fráze v ms
- Délka prozodické fráze ve slabikách
- Délka prozodické fráze ve slovech (grafických, nikoliv fonologických)
- Hloubka následujícího předělu (3 nebo 4, popř. s indexem *p*)

### 7.3 Čištění dat

Celkový počet hlásek v našem materiálu je 113 216. Posléze byly ovšem vyřazeny položky z dysfluentních slov a ze slov s hezitací (příznaky {dsfl} a {hesit}), což činilo dohromady 296 hlásek (přibližně 0,3 % celého materiálu). Celkově tedy analyzujeme 112 920 hlásek.

Dále byly pro většinu analýz zablokovány hlásky ze slov „hm“, „jo“ a „no“ (pokud nebyly, je to v textu explicitně zmíněno). Předpokládáme totiž, že tyto částice budou mít specifické temporální chování související s jejich diskurzí funkcí. Do této skupiny byly zařazeny též dvě realizace slova „aha“, které vykazovaly podobně nestandardní temporální chování.

### 7.4 Statistická analýza dat

V celém textu pracujeme s hladinou významnosti  $\alpha = 0,05$ , tedy 5 %.

Protože analyzovaných případů je velké množství, je nutné pro statistické testy data shlukovat do skupin. V opačném případě by v t-testech a analýzách rozptylu počet stupňů volnosti uvnitř skupin dosahoval až desetitisíců a významnost výsledků by se tak uměle zvyšovala, čímž bychom podstupovali značné riziko chyby 1. druhu (Volín, 2007: 33–34). Počet stupňů volnosti jsme snížili shlukováním dat do skupin, a dále jsme analyzovali tyto skupinové průměry. Při deskriptivních analýzách seskupujeme data vždy podle mluvčího, čímž se snažíme odfiltrovat individuální variabilitu a získat generalizovatelnější reprezentativní hodnoty. Od toho ustupujeme při analýzách souvisejících s identifikací, tam je naopak variabilita mezi mluvčími a v rámci mluvčích zásadní.

Například seskupení podle mluvčího (34 úrovní, podle 34 mluvčích) a pozice hlásky ve slově (4 úrovně: iniciální, mediální, finální a individuální) bude znamenat, že spočítáme průměr pro všechny iniciální hlásky každého mluvčího, mediální hlásky každého mluvčího, atd., a tyto

průměry pak použijeme jako vstup do statistických testů. Počet případů se tak sníží na 136 (34\*4).

Co se týče konkrétních statistických testů, v deskriptivní části práce byly použity dvouvýběrové t-testy pro nezávislá měření (v případě, že byly porovnávány pouze dvě skupiny), dále analýzy rozptylu (zkr. ANOVA), a to jedno- i vícefaktorové, a Tukeyho HSD post-hoc testy. K některým dílčím analýzám využíváme také Pearsonův korelační výpočet.

## 7.5 Artikulační tempo a globální temporální ukazatele

Pro potřeby kvantifikace tempa řeči používáme artikulační tempo, tedy trvání lingvistické jednotky za jednotku času s vyřazením pauz. Průměrné artikulační tempo jednoho mluvčího je spočítáno jako průměr artikulačních temp ve všech jím pronesených replikách. Repliky obsahující dysfluence a hezitace byly vyřazeny. Artikulační tempo bylo vyjádřeno ve slabikách a hláskách za sekundu, přičemž šlo o reálně vyslovené slabiky i hlásky, k elidovaným jednotkám nebylo přihlíženo. Mluvní tempo (tedy se zahrnutím pauz) počítáno nebylo ze dvou důvodů: jednak z literatury vyplývá, že artikulační tempo je stabilnější v rámci jednoho mluvčího (např. Goldman-Eisler, 1968; Künzel, 1997 nebo Jessen, 2007), jednak náš materiál obsahuje vzhledem k metodice sběru tak málo pauz uvnitř repliky jednoho mluvčího, že by hodnoty nebyly nijak odlišné od tempa artikulačního.

Co se týče globálních temporálních ukazatelů, z důvodů zmíněných v oddílu 2.1 zde opouštíme zavádějící název „rytmické ukazatele“. Termín globální také navozuje žádoucí kontrast k ukazateli lokálního artikulačního tempa (LAR, viz níže). Přestože v některých zdrojích můžeme nalézt tvrzení, že index párové variability (PVI) je lokálním ukazatelem (např. Loukina et al. 2011), není tomu tak. PVI sice bere v úvahu lokální rozdíl mezi po sobě následujícími intervaly, tento rozdíl však průměruje přes celou promluvu.

V této práci budou ukazatele průměrovány vždy v jedné replice, bez zahrnutí pauz. Také zde byly vyřazeny repliky s dysfluencemi a hezitacemi.

Globální temporální ukazatele byly spočteny následovně:

- %V: procento vokalických intervalů v řeči, tedy  $\%V = (d_{vt}/d_t) \times 100$ , kde  $d_{vt}$  je celkové trvání vokalických intervalů a  $d_t$  celkové trvání.
- $\Delta V/\Delta C$ : směrodatná odchylka trvání vokalických, respektive konsonantických intervalů v řeči, tedy:

$$\Delta V = \sqrt{\frac{\sum (d_v - d_{vavg})^2}{n-1}},$$

kde  $d_v$  je trvání jednoho vokalického intervalu a  $d_{vavg}$  průměrné trvání všech vokalických intervalů; analogicky pro  $\Delta C$ .

- VarcoV/VarcoC: Variační koeficient  $\Delta V$ , respektive  $\Delta C$ , tzn. dělené průměrným trváním vokalických, resp. konsonantických intervalů, tedy:  $\text{VarcoV} = \Delta V / d_{vavg}$ . Analogicky pro VarcoC.
- rPVI-V/rPVI-C: nenormalizované (angl. *raw*) indexy párové variability pro vokalické, resp. konsonantické intervaly. Spočítají se podle vzorce:

$$rPVI = [\sum_{k=1}^{n-1} |d_k - d_{k+1}| / (n-1)],$$

kde  $d_k$  je trvání k-tého intervalu a  $n$  počet intervalů.

- nPVI-V/nPVI-C: normalizované indexy párové variability pro vokalické, resp. konsonantické intervaly. Vzorec výpočtu je následující:

$$nPVI = 100 \times \left[ \sum_{k=1}^{n-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (n-1) \right]$$

## 7.6 LAR

Ukazatel lokálního artikulačního tempa (LAR) je převzat od Volína (2009), který jej počítá jako převrácenou hodnotu vzdálenosti dvou po sobě následujících slabičných jader:

$$LAR_{pk-pk} = \frac{1}{Dur_{pk-pk}},$$

kde  $Dur_{pk-pk}$  je vzdálenost středů dvou po sobě následujících vokálů. Ve světle nejnovějšího výzkumu v češtině (Volín et al., 2014), který naznačuje, že percepční centrum slabiky se pohybuje kolem onsetu slabičného jádra, případně dokonce v průtoku slabičky, jsme se rozhodli měřit LAR jako vzdálenost onsetů dvou následujících slabičných jader (tj. vokálů, případně slabičných konsonantů). Na rozdíl od Volína (2009) a Weingartové (2013) nevyhlazujeme křivky LAR klouzavým průměrem. V jejich studiích to bylo nutné pro vyhlazení výkyvů křivky v důsledku hláskové struktury textu (například kontura LAR by tak ukazovala na „zpomalení“ v každém souhláskovém shluku) – zde to potřeba nebude, protože kontury budou normalizovány vůči modelovým konturám, které již hláskovou strukturu berou v úvahu (viz níže, kapitola 10).

## 8 Deskripce trvání hlásek

V tomto oddílu bude prozkoumáno trvání hlásek ve výše popsaném materiálu s přihlédnutím k pozici ve slově, ve slabice, případně v prozodické frázi, hláskovému okolí a dalším parametrům. Pokusíme se též poskytnout průměrné hodnoty, které by mohly být považovány za reprezentativní (tj. průměrné trvání z pečlivě vybraného vzorku, kde by se mělo projevat co nejméně faktorů), a to v oddílech 8.2 a 8.4.

Naměřené hodnoty jsou rovněž shrnuty v tabulkách v příloze č. 4.

### 8.1 Vokály

Celkový počet analyzovaných vokálů je 45 645. Vyřazení slov *jo* a *no* (viz výše) způsobilo ztrátu hlásky [o:]. Tato hláska bude níže analyzována samostatně. Počty výskytů jednotlivých vokálů zobrazuje tabulka 8.1a. Jak vidíme, v korpusu se nevyskytují diftongy [au] ani [eu].

Krátký vokál	Poč. výskytů	Dlouhý vokál	Poč. výskytů
a	7777	a:	2693
e	13874	e:	227
i	5478	i:	3273
o	8048	o:	0
u	3017	u:	235
Diftong		Poč. výskytů	
ou	1019	Celkem	45641

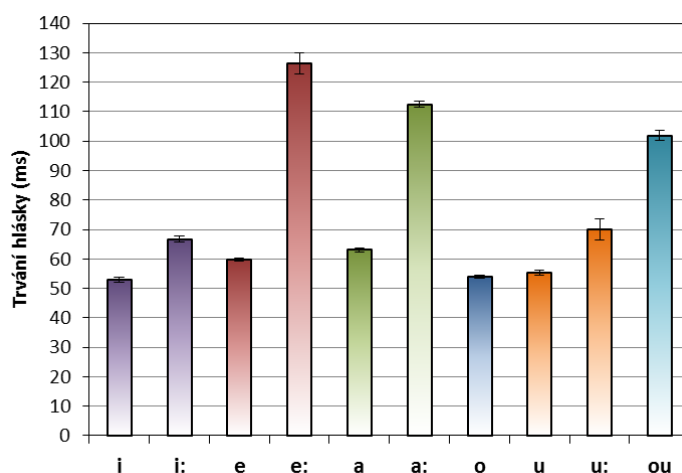
Tabulka 8.1a: Počty výskytů jednotlivých vokálů v korpusu Minidialogy-H.

Vokál	Iniciální	Mediální	Finální	Individuální
a	287	5699	1686	105
a:	1	2284	408	
e	35	6880	6863	96
e:		102	125	
i	61	3164	2250	3
i:		2214	1059	
o	100	4374	3530	44
u	167	1902	948	
u:		201	34	
ou		407	613	
Celkem	651	27227	17515	248

Tabulka 8.1b: Počty výskytů jednotlivých vokálů v korpusu Minidialogy-H rozdělené podle pozice ve slově.

Tabulka 8.1b rozděluje počty výskytů jednotlivých vokálů podle pozice ve slově. Individuální hláska znamená, že sama tvoří celé slovo. 96 individuálních [e] vzniklo elizí [j] ve slově *je* (viz obr. 7.2). Zde je nutno podotknout, že předcházel-li vokalické předložce nebo spojce ráz, nebyl vokál klasifikován jako individuální, ale jako finální.

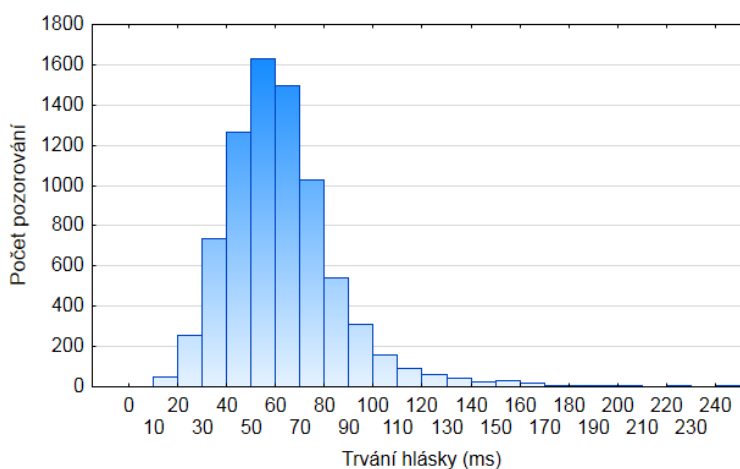
Průměrná trvání jednotlivých vokálů z neseskupených dat ukazuje obrázek 8.1. Konkrétním vokálům budou věnovány následující podkapitoly.



Obrázek 8.1: Průměrné trvání jednotlivých vokálů z neseskupených dat. Svorky označují 95% interval spolehlivosti.

### 8.1.1 Vokály [a] a [a:]

Histogram níže zobrazuje trvání vokálu [a] ve všech pozicích a všech hláskových okolích. Celkový počet analyzovaných [a] je 7777, pro počty na jednotlivých pozicích viz tabulka 8.1b. Průměrné trvání je 63 ms se směrodatnou odchylkou 24,7 ms; medián je 59,8 ms. Minimální trvání [a] je 9,9 ms, maximální 295,9 ms, 10. percentil je přitom 37,4 ms a 90. percentil 90,1 ms.

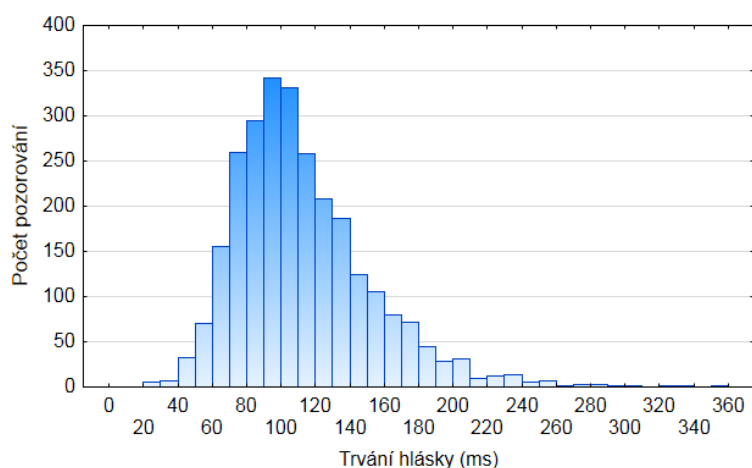


Obrázek 8.2: Histogram trvání vokálu [a] ve všech pozicích (iniciální, mediální, finální, i samostatně) a všech hláskových okolích.

Iniciální [a] (tedy první ve slově, bez rázu) mají průměrné trvání 62,6 ms, medián 61, 1 ms a směrodatnou odchylku 16,7 ms. Mediální [a] trvají průměrně 62,4 ms se směrodatnou odchylkou 22,9 ms a mediánem 59,5 ms. Finální jsou nejdelší, s průměrem 65,2 ms, mediánem 60,4 ms a směrodatnou odchylkou 31,8 ms. Individuálních [a] je nejméně a jejich průměr činí 61,3 ms, směrodatná odchylka 15,8 ms a medián 60,3 ms.

V datech seskupených podle mluvčího a pozice hlásky ve slově byl nalezen rozdíl mezi průměrným trváním mediálních a finálních [a], ale jen okrajově statisticky významný:  $t(66) = 1,9$ ;  $p = 0,06$ . Při seskupení do menších skupin podle mluvčího, pozice hlásky ve slově a navíc pozice slova v prozodické frázi pak rozdíl mezi mediálními a finálními dosáhne dokonce vysoké významnosti:  $t(270) = 6,5$ ;  $p < 0,001$ . Iniciální a individuální [a] není možné takto porovnávat vzhledem k jejich nízkému výskytu.

Histogram na obr. 8.3 ukazuje rozložení trvání dlouhého vokálu [a:], opět ve všech pozicích a všech hláskových okolích. Analyzovaných položek je 2693 a kromě jedné iniciální (která vznikla elizí [n] ve slově *nákresy*) jsou všechny v mediální či finální pozici. Celkové průměrné trvání [a:] činí 112,5 ms se směrodatnou odchylkou 39,6 ms a mediánem 105,2 ms. Minimální trvání je 20,6 ms, maximální 360 ms, 10. percentil 69,9 ms a 90. percentil 165,9 ms.



Obrázek 8.3: Histogram trvání vokálu [a:] ve všech pozicích (iniciální, mediální a finální) a všech hláskových okolích.

Mediální [a:], která tvoří většinu vzorku, mají průměrné trvání 113 ms, směrodatnou odchylku 39,7 ms a medián 104,9 ms. Finální [a:], jichž je přibližně šestkrát méně, mají průměrné trvání 109,6 ms se směrodatnou odchylkou 39,1 ms a mediánem 107 ms.

Tento zvláštní výsledek (průměrně delší trvání mediálních oproti finálním) je pravděpodobně způsoben dvěma faktory: U mediálních je o něco větší podíl položek ze slov ve finální pozici

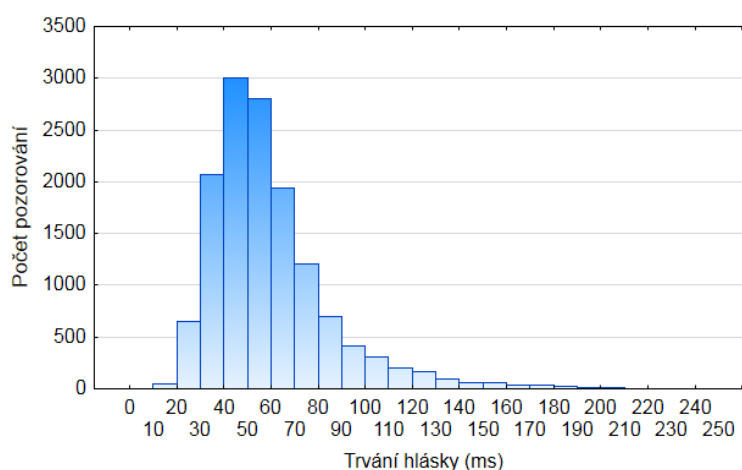


v prozodické frázi a naopak u finálních je poměrně velký podíl vokálů ze slova *já* (jež se vyskytuje pouze v iniciální a mediální pozici v prozodické frázi), které mají podprůměrně krátká trvání (kolem 70 ms). Lze navíc také předpokládat, že mediální vokály v posledních slabikách, po nichž ještě následuje finální konsonant, také podléhají frázovému zpomalování. Podrobněji se tomuto jevu budeme věnovat v oddílu 8.1.8.

Rozdíly mezi trváním mediálních a finálních [a:] nejsou ovšem statisticky významné, a to ani na neseskupených datech.

### 8.1.2 Vokály [e] a [e:]

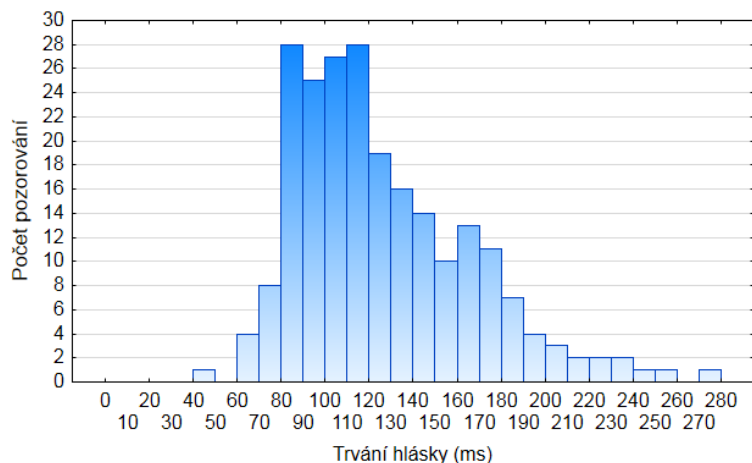
V histogramu na obr. 8.4 jsou zobrazena trvání všech hlásek [e] bez ohledu na pozici ve slově či hláskové okolí. Tato [e] mají průměrné trvání 59,7 ms se směrodatnou odchylkou 26,5 ms a mediánem 53,8 ms. Minimální nalezené trvání [e] je 7,5 ms, maximální 250,1 ms, 10. percentil 34,2 ms a 90. percentil 91,2 ms.



Obrázek 8.4: Histogram trvání vokálu [e] ve všech pozicích (iniciální, mediální, finální, i samostatné) a všech hláskových okolicích.

Iniciální [e], jichž je ale pouze 35, mají průměrné trvání 46 ms se směrodatnou odchylkou 12,2 ms a mediánem 46,6 ms. Mediální mají průměr 55,5 ms, směrodatnou odchylku 19 ms a medián 53 ms. Finální [e] trvají průměrně 63,9 ms se směrodatnou odchylkou 31,8 ms; medián je 54,7 ms. Individuální [e] (vzniklá elizí [j] ve slově *je*) mají průměrné trvání 61,1 ms, směrodatnou odchylku 21,5 ms a medián 56,9 ms. Rozdíl mezi trváním mediálních a finálních [e] je statisticky vysoce významný, a to dokonce při seskupení jen podle mluvčího a pozice hlásky ve slově:  $t(66) = 6,2$ ;  $p < 0,001$ . Iniciální a individuální opět nemá smysl statisticky analyzovat kvůli malému počtu výskytů.

Dlouhých [e:] je v korpusu nejméně, což koresponduje i s frekvencí této hlásky v češtině. Vyskytuje se zejména ve slovech *problémy*, *šéf* a *malér*, dále pak v některých realizacích slova *ne* a v několika případech, kdy mluvčí přečetli slovo *zbytečně* jako *zbytečné*. Je tedy nutné mít na paměti, že tento vzorek není příliš reprezentativní.



Obrázek 8.5: Histogram trvání vokálu [e:] ve všech pozicích (mediální a finální) a všech hláskových okolích.

V histogramu 8.5 se zřetelně projevuje malý počet výskytů – v grafu jsou vidět lokální maxima, která budou pravděpodobně odpovídat jednotlivým slovům. Nicméně celkový průměr trvání [e:] je 126,3 ms se směrodatnou odchylkou 39,7 ms, medián je 115,7 ms.

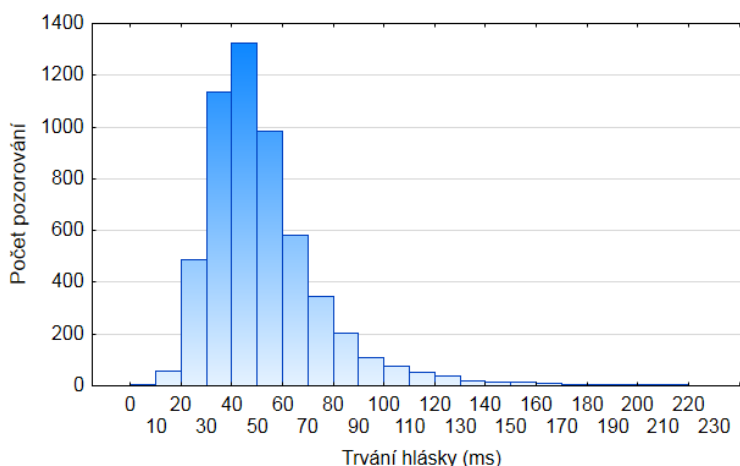
Vezmeme-li všechna mediální [e:], jichž je přibližně polovina, zjistíme, že jejich průměrné trvání je 129,5 ms, směrodatná odchylka 32,6 ms a medián 126 ms. U finálních nastává podobná situace jako u [a:], jejich trvání je průměrně kratší (123,7 ms, sm. odch. 44,6 ms a medián 107 ms) než u mediálních. Tento efekt je přitom statisticky významný při seskupení dat podle mluvčího a pozice hlásky ve slově:  $t(66) = 2,08$ ;  $p = 0,04$ . Je nicméně způsobený nevyrovnaností vzorku dat – finální [e:] jsou téměř všechna z jednoslabičného slova *ne*, většinou v mediální pozici v prozodické frázi (70 %). Všechna mediální [e:] jsou ale z finálních slov v prozodické frázi, a tedy zde bude hrát roli spíše závěrové zpomalování (viz také podkapitola 8.1.8).

### 8.1.3 Vokály [ɪ] a [i:]

Obrázek 8.6 ukazuje rozložení hodnot trvání [ɪ] ve všech pozicích a všech hláskových okolích. Průměrné trvání [ɪ] je 52,9 ms se směr. odchylkou 24,2 ms a mediánem 48 ms. Minimální trvání je 7 ms, maximální 230 ms, 10. percentil je 30 ms a 90. percentil 80,6 ms.

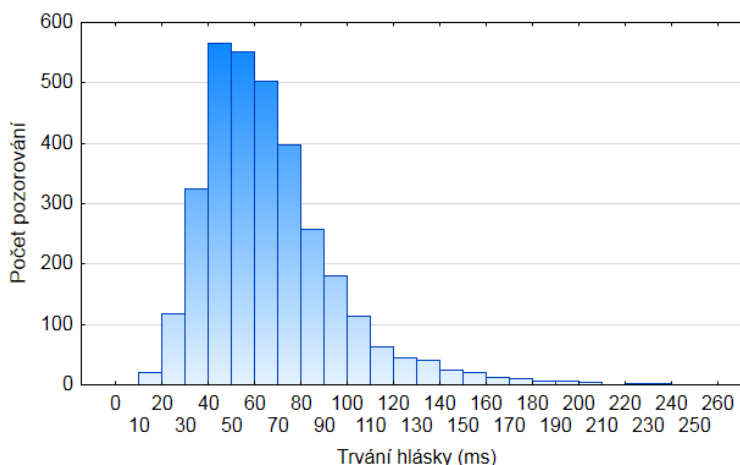
Iniciálních [ɪ] je pouze 61 a mají průměrné trvání 45,5 ms, směr. odchylku 10,2 ms a medián 45,4 ms. Mediální mají průměr 51,1 ms, směrodatnou odchylku 22,8 ms a medián 46,4 ms. Finální [ɪ] jsou pak nejdelší s průměrným trváním 55,7 ms, směrodatnou odchylkou 26 ms a

mediánem 49,9 ms. Rozdíl mezi průměrem mediálních a finálních vokálů je statisticky významný při seskupení dat pouze podle mluvčího a pozice hlásky ve slově,  $t(66) = 3,5$ ;  $p < 0,001$ . Individuální [i] jsou pouze tři (vznikly vložením spojky *i* nebo elizí [b] ve slově *by*).



Obrázek 8.6: Histogram trvání vokálu [i] ve všech pozicích (iniciální, mediální, finální i samostatně) a všech hláskových okolích.

Dlouhá [i:] se vyskytují v mediální a finální pozici, histogram 8.7 zobrazuje celkové rozložení jejich trvání. Průměrně trvají 66,7 ms, směrodatná odchylka 29,7 ms, minimum 7,7 ms, maximum 336 ms, 10. percentil 36,6 ms a 90. percentil 102 ms.

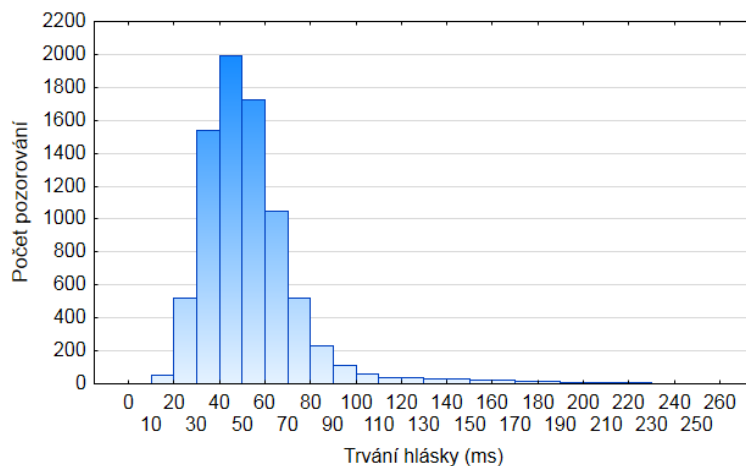


Obrázek 8.7: Histogram trvání vokálu [i:] ve všech pozicích (mediální a finální) a všech hláskových okolích.

Průměrné trvání mediálních [i:] je 64,9 ms se směrodatnou odchylkou 25,8 ms, medián je 61,3 ms. Finální [i:] trvají průměrně 70,4 ms se směrodatnou odchylkou 36,4 ms a mediánem 61,1 ms. Průměry mediálních a finálních jsou opět statisticky významně odlišné, a to při seskupení podle mluvčího a pozice hlásky ve slově:  $t(66) = 2,9$ ;  $p = 0,004$ .

### 8.1.4 Vokál [o]

Hlávka [o] má v našem materiálu průměrné trvání 55,4 ms, směrodatnou odchylku 26 ms a medián 50,3 ms. Minimální nalezené trvání je 12,9 ms, maximální 303,8 ms, 10. percentil 32 ms a 90. percentil 80 ms. Rozložení hodnot zobrazuje graf na obr. 8.8.



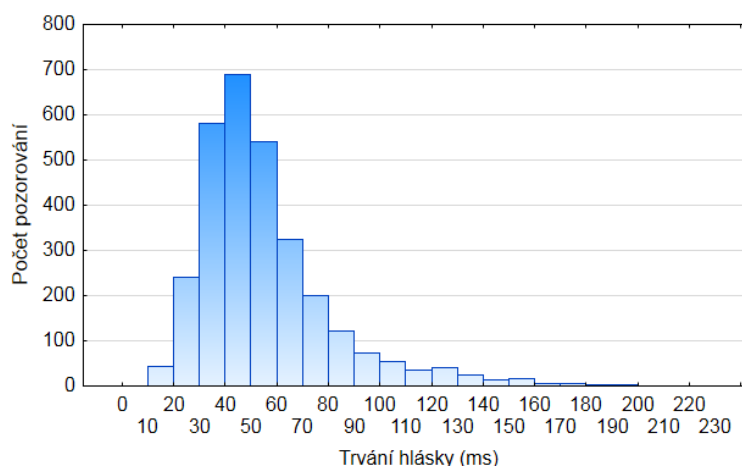
Obrázek 8.8: Histogram trvání vokálu [o] ve všech pozicích (iniciální, mediální, finální a samostatné) a všech hláskových okolích.

Iniciální [o] se vyskytuje pouze ve 100 případech, nicméně jeho průměrné trvání je 58 ms, směrodatná odchylka 16,4 ms a medián 57,2 ms. Mediální [o] trvá průměrně 53,6 ms, směrodatnou odchylku má 19,3 ms a medián 51,3 ms. Finální pak průměrně trvá 57,4 ms, směrodatná odchylka 32 ms a medián 48,8 ms. Samostatných [o] je pouze 44 a trvají průměrně 56,5 ms se směrodatnou odchylkou 22,7 ms a mediánem 50,7 ms. Finální [o] jsou opět významně delší než mediální při seskupení podle mluvčího, pozice hlávky ve slově a pozice slova v prozodické frázi:  $t(270) = 7,06$ ;  $p < 0,001$  (seskupení jen podle mluvčího a pozice hlávky ve slově v tomto případě k významnosti rozdílu nestačí).

Dlouhý protějšek, [o:], se v protříděných datech nevyskytuje a bude analyzován dále samostatně.

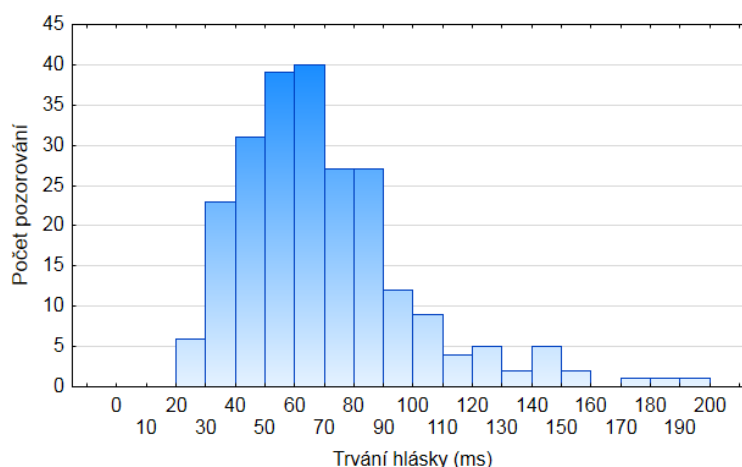
### 8.1.5 Vokály [u] a [u:]

Krátké [u] se v našem materiálu nalézá v iniciální, mediální a finální pozici, nikoliv samostatné (předložka *u* se v textu nevyskytuje). Všechna [u] dohromady mají průměrné trvání 55,3 ms, směrodatnou odchylku 26,3 a medián 49,3 ms. Minimální trvání je 8,8 ms a maximální 229 ms, 10. percentil 30,4 ms a 90. percentil 86,2 ms.



Obrázek 8.9: Histogram trvání vokálu [u] ve všech pozicích (iniciální, mediální a finální) a všech hláskových okolích.

Iniciální [u] trvá průměrně 54,8 ms, se směrodatnou odchylkou 16,2 ms a mediánem 54,5 ms (počet výskytů je ale opět nízký, 167). Mediální pak má průměrné trvání 48,2 ms, směrodatnou odchylku 17,8 ms a medián 45,9 ms. Finální [u] trvají průměrně 69,6 ms se směrodatnou odchylkou 34,9 ms a mediánem 58,9 ms. Rozdílnost průměrů trvání mediálních a finálních [u] je statisticky vysoce významná,  $t(66) = 11,2$ ;  $p < 0,001$  (při seskupení podle mluvčího a pozice hlásky ve slově), finální jsou výrazně delší.

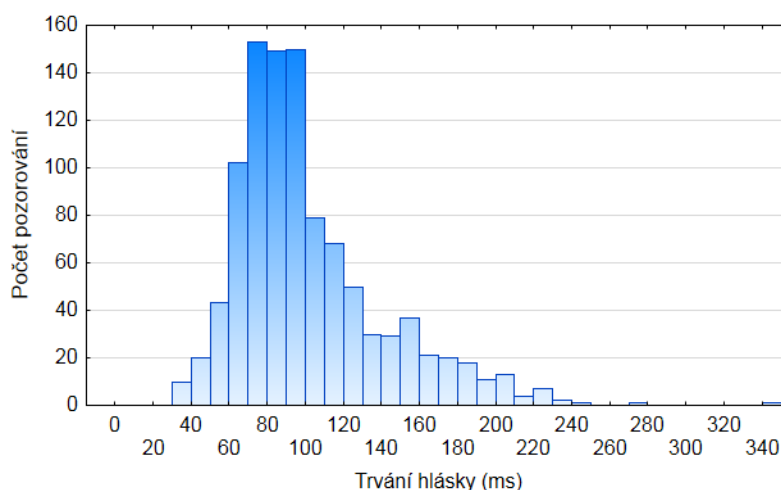


Obrázek 8.10: Histogram trvání vokálu [u:] ve všech pozicích (mediální a finální) a všech hláskových okolích.

Dlouhé [u:] není příliš častá hláška, v našem materiálu se objevuje jen 235krát, z toho většinou v pozici mediální. Bez ohledu na pozici trvá [u:] průměrně 70 ms, směrodatná odchylka je 29,7 ms, medián 65,5 ms, minimum 22,2 ms, maximum 197,4 ms, 10. percentil 38,5 ms a 90. percentil 107,5 ms. Mediální [u:] trvají průměrně 62,4 ms, směrodatná odchylka 20,8 ms a medián 61 ms. Finální [u:] jsou opět delší, s průměrem 115,4 ms, směrodatnou odchylkou 34,5 ms a mediánem 112,4 ms. Tento rozdíl je vysoce statisticky významný při seskupení položek podle mluvčího a pozice hlásky ve slově:  $t(66) = 8,7$ ;  $p < 0,001$ .

### 8.1.6 Diftong [ou̯]

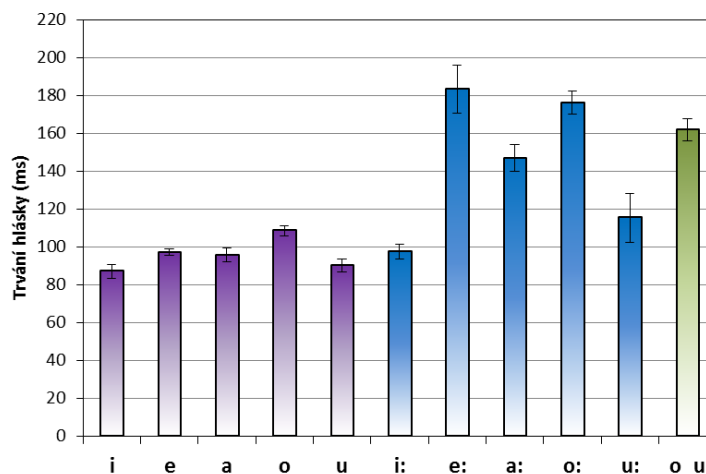
Dvojhlasčka [ou̯] se v našem korpusu vyskytuje celkem 1019krát. Její průměrné trvání je 101,9 ms, směrodatná odchylka 39,2 ms, medián 91,7, minimum 32,5 ms, maximum 342,2 ms, 10. percentil 63,5 ms a 90. percentil 159,1 ms. Menší část výskytů tvoří mediální hlásky, které mají průměrné trvání 98,7 ms, směrodatnou odchylku 30,6 ms a medián 92,5 ms. Finální [ou̯] jsou pak v průměru dlouhé 104 ms, se směrodatnou odchylkou 43,8 ms a mediánem 91,5 ms. Tyto dvě skupiny se od sebe opět významně liší,  $t(66) = 8,8$ ;  $p < 0,001$  (při seskupení podle mluvčího a pozice hlásky ve slově), finální vykazují delší trvání.



Obrázek 8.11: Histogram trvání diftongu [ou̯] ve všech pozicích (mediální a finální) a všech hláskových okolích.

### 8.1.7 Vokál [o:]

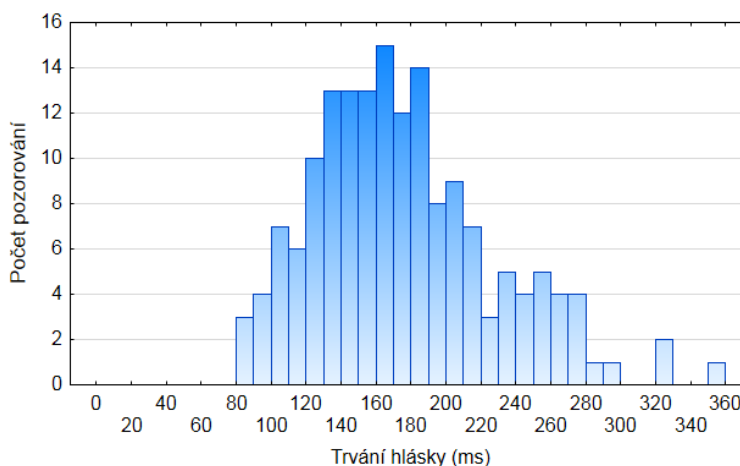
Dlouhý vokál [o:] se v korpusu Minidialogy vyskytuje pouze v některých realizacích částic *jo* a *no* (kvantita vokálu byla hodnocena percepčně), čili jde ve všech případech o finální pozici ve slově. Celkově je slov s výskyty [o:] 175 a z toho 114 (zhruba 65 %) tvoří samostatnou prozodickou frázi. Dalších 50 je ve finální pozici prozodické fráze (tvořené slovy *no jo*) a pouze 11 jich je v iniciální či mediální pozici. Porovnávat tedy celkové průměrné trvání [o:] s ostatními vokály by bylo značně zkreslující – můžeme jej pouze porovnávat s trváním všech finálních vokálů ve slovech, které stojí na konci prozodické fráze.



Obrázek 8.12: Průměrné neseskupené trvání finálních vokálů ve finálních a individuálních slovech v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Obr. 8.12 ukazuje průměrné trvání finálních vokálů ve finálních (a také individuálních) slovech. Vidíme, že trvání [o:] se nejvíce blíží trvání [e:]. To ovšem může být dáno také podobností slov, ve kterých se tyto dvě hlásky v této pozici vyskytují (*jo, no* vs. *ne*) – nezapomínejme také, že tyto dva vokály mají v korpusu zdaleka nejméně výskytů.

Finální [o:] (z finálních a individuálních slov v prozodické frázi) trvá průměrně 176,2 ms se směrodatnou odchylkou 52,7 ms a mediánem 167,4 ms. Minimální trvání je 83,4 ms, maximální 355,9 ms, 10. percentil 115,3 ms a 90. percentil 254,7, viz histogram 8.13. Oproti finálnímu krátkému [o] je finální [o:] o 68 ms a 1,6krát delší.



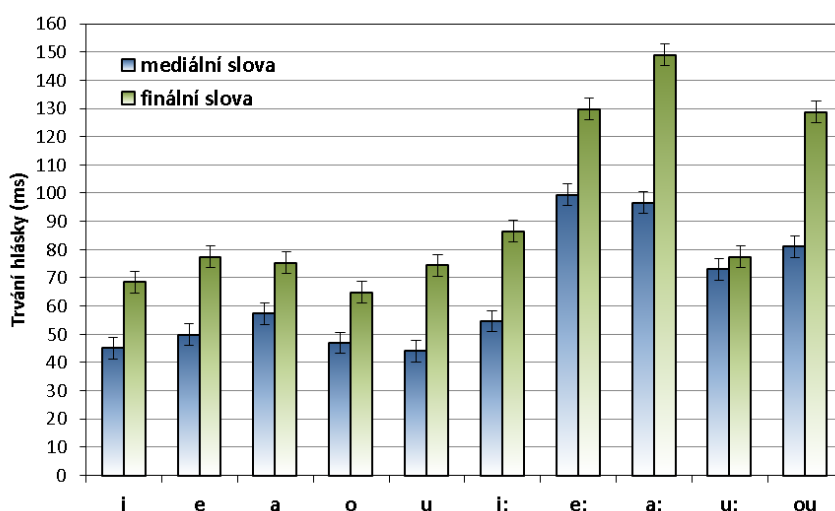
Obrázek 8.13: Histogram trvání vokálu [o:] ve finální pozici v individuálních a finálních slovech.

### 8.1.8 Vliv frázového zpomalování na trvání vokálů

Graf 8.14 zobrazuje průměrné trvání vokálů v mediálních a finálních slovech prozodické fráze ukončené předělem hloubky 4 (tj. plným prozodickým předělem). Rozdíly oproti průměrnému trvání finálních vokálů z obrázku 8.11 jsou dány tím, že zde jsme zobrazili seskupená data

(viz níže) a neomezili se pouze na skutečně finální vokály – tj. bereme v úvahu vokály ze všech slabik finálních slov.

Vidíme, že ve všech případech jsou vokály z finálních slov v prozodické frázi delší než ze slov mediálních. K ověření statistické významnosti těchto rozdílů byla využita dvoufaktorová ANOVA s faktory HLÁSKA (a, a:, e, e:, i, i:, o, u, u:, ou) a POZICE (v prozodické frázi: mediální vs. finální). Případy byly před tím seskupeny podle mluvčího, hlásky a pozice slova v prozodické frázi. Efekt pozice je statisticky významný:  $F(1, 660) = 1056,9; p < 0,001$ . Významná je též interakce HLÁSKA\*POZICE:  $F(9, 660) = 26,1; p < 0,001$ . Tukeyho post-hoc test prozrazuje, že k rozdílu přispívají téměř všechny vokály – kromě [u:] (který není významný vůbec) jsou rozdíly u jednotlivých vokálů významné na hladině  $p < 0,001$ .



Obrázek 8.14: Průměrné trvání jednotlivých vokálů ze slov v mediální a finální pozici v prozodické frázi. Svorčky označují 95% interval spolehlivosti.

Tabulka 8.2 pak zobrazuje koeficienty prodlužování ve finálních slovech, v prvním sloupci multiplikativní (tj. finální trvání vydělené trváním mediálním), ve druhém sloupci pak aditivní (tj. finální minus mediální). Vidíme, že multiplikativní rozdíl se pohybuje mezi 1,06 a 1,69 s průměrem těsně kolem 1,5 – čili vokály ve finálních slovech jsou průměrně 1,5krát delší než v mediálních slovech. Nejmenší rozdíl je u [u:], největší pak u krátkého [u]. Nelze říci, že by se dlouhé vokály chovaly jinak než krátké, například byly prodlužovány systematicky více či méně. Také diftong [ou] se v tomto ohledu chová velmi podobně.

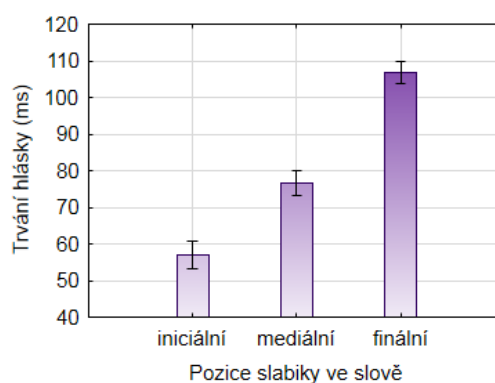
Aditivní rozdíly se pohybují až na jednu výjimku v řádu desítek milisekund, průměrně je vokál ve finálním slově delší o cca 28 ms. Nejméně se prodlužuje opět [u:] (o 4 ms), nejvíce [a:] (o 52 ms). Odhlédneme-li od [u:] (jehož počet výskytů je zdaleka nejnižší), lze pozorovat tendenci k většímu aditivnímu prodlužování dlouhých vokálů.



Vokál	Multipl. koef.	Adit. koef. (ms)
i	1,52	23,5
e	1,55	27,6
a	1,31	18
o	1,38	17,9
u	1,69	30,4
i:	1,58	31,8
e:	1,31	30,3
a:	1,54	52,4
u:	1,06	4,3
ou	1,59	47,6
<b>Průměr</b>	<b>1,45</b>	<b>28,4</b>

Tabulka 8.2: Koefficienty průměrných rozdílů mezi vokály z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

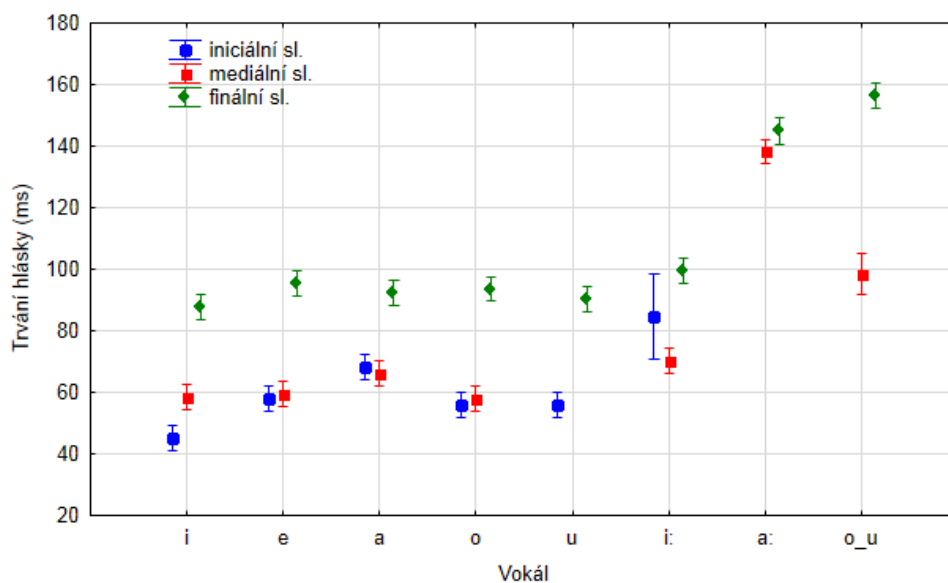
Nabízí se otázka, zda se tendence prodlužovat vokály ve finálních slovech vztahuje na celá slova, či zda je rozdíl, ve které ze slabik finálního slova se vokál nachází. Vokály z finálních slov byly tedy seskupeny podle mluvčího, hlásky a pozice slabiky ve slově a po seskupení byly vyřazeny položky tvořené pouze 1–5 případy, kterých bylo mnoho a zanášely by do výsledků nežádoucí artefakty. Kvůli tomu v analýze nejsou zastoupeny vokály [e:] a [u:]. Jednofaktorová ANOVA provedená na těchto seskupených datech ukázala vysoce signifikantní efekt pozice slabiky ve finálním slově:  $F(2, 654) = 203; p < 0,001$ , přičemž tento rozdíl je signifikantní pro všechny dvojice pozic (Tukeyho post-hoc test:  $p < 0,001$ ).



Obrázek 8.15: Průměrné trvání vokálů z finálních slov rozdělených podle pozice slabiky ve slově ze seskupených dat. Svorcky označují 95% interval spolehlivosti.

Obr. 8.15 ukazuje průměrné trvání všech vokálů ve finálních slovech (za kterými následuje plný prozodický předěl) podle pozice slabiky, jejíž nukleus tvoří. Je zřetelné, že vokály se výrazně prodlužují směrem ke konci slova.

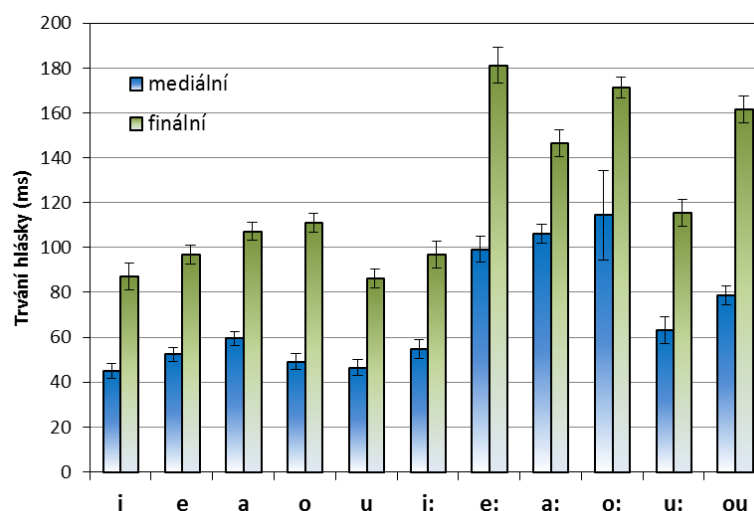
Dále byla provedena dvoufaktorová ANOVA, kde k faktoru POZICE SLABIKY byl přidán ještě faktor HLÁSKA, aby bylo možné zjistit, zda se všechny vokály chovají v tomto ohledu stejně.



Obrázek 8.16: Průměrné trvání jednotlivých vokálů z finálních slov podle pozice slabiky ve slově (barevně odlišené) ze seskupených dat. Svorcky označují 95% interval spolehlivosti.

Obrázek 8.16 ukazuje průměrná trvání vokálů ve všech skupinách. Statistické testy byly provedeny pro dvojice pozic zvláště kvůli nesterenosti zastoupení jednotlivých vokálů. Efekt pozice slabiky ve slově je významný pro dvojice iniciální vs. finální ( $F(1, 365) = 381,6; p < 0,001$ ) a mediální vs. finální ( $F(1, 436) = 572,1; p < 0,001$ ). Vokály v iniciálních slabikách se od mediálních významně neliší. S pomocí Tukeyho post-hoc testů můžeme doplnit, že iniciální od mediálních se liší pouze v případě [ɪ] ( $p < 0,001$ ) a [i:] ( $p = 0,01$ ). Mediální se od finálních významně liší u všech vokálů ( $p < 0,001$ ) kromě [a:] a iniciální se od finálních významně liší u všech vokálů ( $p < 0,001$ ) kromě [i:]. Iniciálních [i:] je ovšem řádově méně. Lze tedy uzavřít, že kromě [a:] (a [e:] a [u:]), jež nebyly do této analýzy zahrnuty) pozorujeme u všech vokálů na konci prozodické fráze trend k prodlužování jádra finální slabiky.

Paralelně k analýze zobrazené na obr. 8.14 můžeme také porovnat mediální vokály se zcela finálními, tj. pouze těmi, které stojí na konci prozodických frází (tedy těch nejvíce prodlužovaných). Obr. 8.17 pak ukazuje výsledky tohoto porovnání. Do této analýzy byla zahrnuta slova *hm*, *jo* a *no*, abychom mohli porovnat také vokál [o:]. Zahrnuli jsme i předěly hloubky 3, abychom zvýšili počet případů pro málo zastoupené vokály. Kvůli příliš vysokému počtu stupňů volnosti byly dvoufaktorové testy ANOVA (s faktory HLÁSKA a POZICE) provedeny pro krátké a dlouhé vokály zvláště.



Obrázek 8.17: Průměrné trvání vokálů z mediálních slov (modrá) a finálních vokálů z finálních slov (zelená) ze seskupených dat. Svorky označují 95% interval spolehlivosti.

Efekt pozice je pro krátké vokály vysoce signifikantní:  $F(1, 840) = 1693,2$ ;  $p < 0,001$ , stejně tak interakce HLÁSKA\*POZICE:  $F(4, 840) = 12,2$ ;  $p < 0,001$ . Podle Tukeyho post-hoc testu jsou významné rozdíly u každého z vokálů na hladině  $p < 0,001$ .

Také pro dlouhé vokály (a diftong [ou]) je efekt pozice vysoce významný:  $F(1, 468) = 481,3$ ;  $p < 0,001$ . Interakce HLÁSKA\*POZICE je také významná:  $F(5, 468) = 16,3$ ;  $p < 0,001$ , stejně jako rozdíly u každého z vokálů (Tukeyho post-hoc test:  $p < 0,001$ ).

I zde můžeme spočítat aditivní a multiplikativní koeficienty prodlužování finálních vokálů.

Vokál	Multipl. koef.	Adit. koef. (ms)
a	1,8	47,7
e	1,85	44,5
i	1,93	42,1
o	2,26	62
u	1,85	39,6
a:	1,38	40,4
e:	1,82	81,9
i:	1,77	42,2
o:	1,5	56,8
u:	1,83	52,3
ou	2,05	83
<b>Průměr</b>	<b>1,82</b>	<b>53,9</b>

Tabulka 8.3: Koeficienty průměrných rozdílů mezi vokály z mediálních slov a finálními vokály z finálních slov v prozodické frázi. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

Jak vidíme z tabulky 8.3, vezmeme-li v úvahu pouze finální vokály a nikoliv všechny vokály z finálních slov, jsou aditivní i multiplikativní koeficienty závěrového prodlužování vyšší než

v tabulce 8.2. Co se týče poměru finálních a mediálních hlásek, nejvíce je prodlouženo [o], které je ve finální pozici více než dvojnásobně dlouhé. Dvojnásobně je také prodloužen diftong [ou̯], většina vokálů se pohybuje kolem 1,8násobného prodloužení. Výrazně nejméně je prodloužen vokál [a:], pouze přibližně 1,4krát. V případě aditivních koeficientů se většina vokálů prodlužuje o 40–60 ms s výjimkou [e:] a [ou̯], které se prodlužují o více než 80 ms.

### 8.1.9 Trvání vokálů po rázu

Ráz (nebo obecněji glotalizace) je v našich datech značen jako hláska bez fonémického statusu, třebaže v češtině jej považujeme pouze za hraniční signál (Volín, 2012: 51). Jeho přítomnost či nepřítomnost a segmentace jeho hranic v korpusu je poměrně přímočará, přestože může být realizován různými způsoby (kanonicky jako exploziva, jako třepená fonace, případně jako fonace dyšná). Pouze v případě třepené fonace mohlo dojít k nejasnostem při značení hranic, pokud byl glotalizován celý vokál. V takových případech byla jako ráz označena první polovina vokálu.

Pár slov k míře glotalizace v našem materiálu: Pokud spočítáme slova začínající (kanonicky) na samohlásku a nezahrneme slova po pauze (kde je výskyt rázu obligatorní a souvisí s nastavením artikulátorů při přechodu z ticha k fonaci), dostaneme celkem 790 slov s potenciálem ke glotalizaci. Z toho u 486 případů se skutečně vyskytl ráz, což činí 61,5 %. Porovnáme-li toto číslo s Volínem (2012), zjistíme, že se nacházíme mezi jeho výsledky pro čtené projevy (92,6 %) a nepřipravené dialogy (53 %), což odpovídá mluvnímu stylu v našem experimentu, který také leží někde mezi oběma zmíněnými.

Do následujících analýz jsou nicméně zahrnuta i slova po pauze (tedy iniciální v prozodické frázi), protože nás zajímá vliv přítomnosti rázu na trvání následujícího vokálu. Nejsou ovšem zahrnuta slova, která začínají kanonicky na souhlásku, ale díky její elizi se pak iniciálním stal následující vokál.

Vokál	Počet	Průměr	Sm. odch.	Počet	Průměr	Sm. odch.	Poměr	Rozdíl
Vokály po rázu				Vokály bez rázu				
a	1376	<b>45,5</b>	20,1	169	<b>65,9</b>	17,4	<b>0,69</b>	<b>20,4</b>
i	90	<b>37,9</b>	12,7	47	<b>47,3</b>	9	<b>0,8</b>	<b>9,4</b>
o	377	<b>42,6</b>	13,1	102	<b>57,9</b>	16,3	<b>0,74</b>	<b>15,3</b>
u	309	<b>40,8</b>	12,8	35	<b>70,2</b>	16,5	<b>0,58</b>	<b>29,4</b>

Tabulka 8.4: Průměrné trvání iniciálních vokálů ve slovech začínajících na samohlásku s rázem (vlevo) a bez rázu (vpravo). Hodnoty trvání jsou uvedeny v milisekundách.

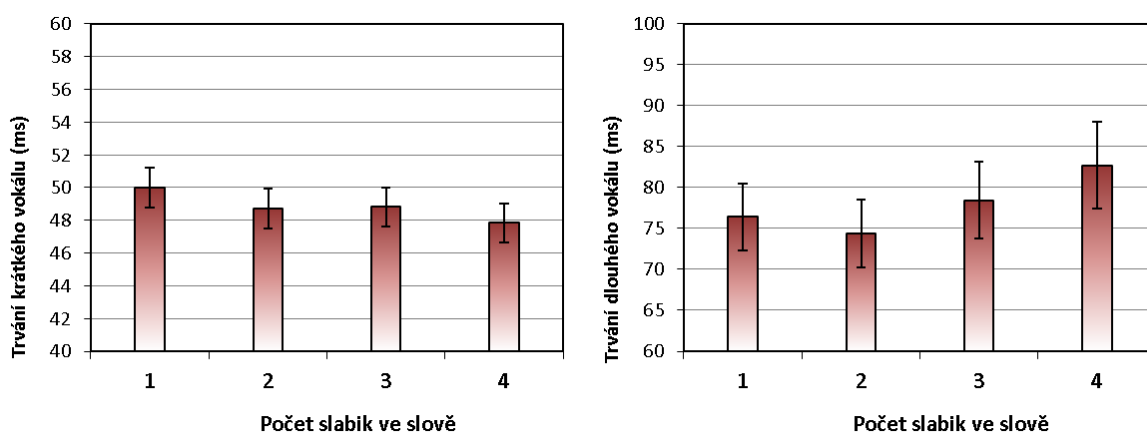
V tabulce 8.4 jsou zobrazeny hodnoty trvání iniciálních vokálů po rázu a bez rázu. V posledních dvou sloupcích je uveden jejich poměr (vokál po rázu vůči vokálu bez rázu) a

rozdíl. V korpusu jsou zastoupena pouze slova začínající na krátké samohlásky [a], [ɪ], [o] a [u]. Je zřetelné, že vokály po rázu jsou v průměru kratší než jejich protějšky, u kterých se ráz nevyskytuje. Přítomnost rázu zkracuje každý vokál trochu jinak, nejvíce [u], které je zkráceno na 0,58 svého trvání a nejméně [ɪ], kterého zůstává 80 %.

K ověření statistické významnosti těchto rozdílů byla opět provedena dvoufaktorová ANOVA na datech seskupených podle mluvčího, hlásky a přítomnosti glotalizace. Efekt přítomnosti glotalizace je vysoce významný:  $F(1, 234) = 307; p < 0,001$ . Také interakce HLÁSKA\*GLOTALIZACE je významná:  $F(3, 234) = 17,2; p < 0,001$ , přičemž významné jsou také rozdíly u každého vokálu zvlášť (Tukeyho post-hoc test:  $p < 0,001$ ). Podíváme-li se na průměrné hodnoty v tabulce 8.5, vidíme, že ačkoliv se trvání vokálů bez rázu poměrně hodně liší, vokály následující po rázu mají přibližně stejné trvání kolem 40 ms. To potvrzuje i post-hoc test – ze všech dvojic vokálů po rázu se od sebe signifikantně liší pouze [a] od [ɪ], ostatní nikoliv. Je možné, že jde o manifestaci podlahového efektu – ráz „ukrajuje“ z trvání vokálu, ale jen do určité míry, za niž už se mluvčí vokál zkrátit zdráhají.

### 8.1.10 Vliv délky vyšší jednotky na trvání vokálů

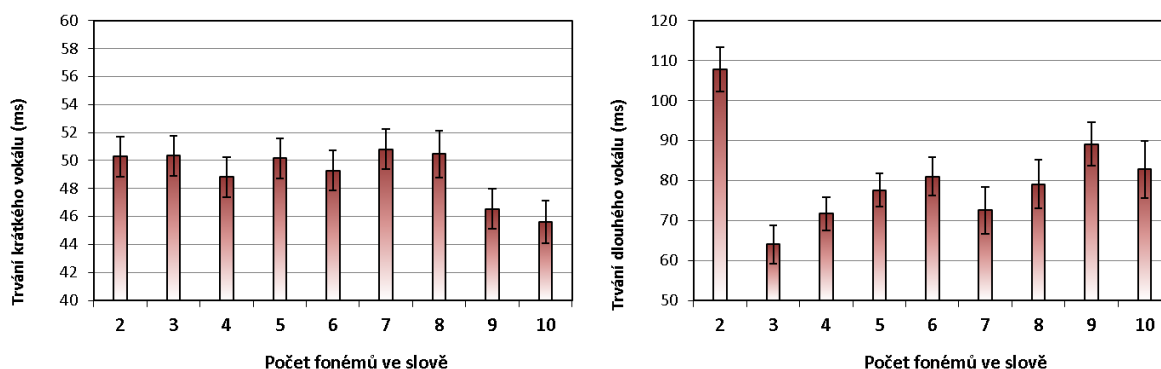
Abychom mohli prozkoumat vliv délky slova na trvání vokálů, seskupili jsme data podle mluvčího, hlásky a délky slova (vyjádřené nejprve ve slabikách, potom ve fonémech). Vyřazeny byly položky, které vznikly seskupením pěti a méně případů. V úvahu jsme vzali pouze mediální slova; krátké a dlouhé vokály (i s diftongem [ou̯]) byly analyzovány zvlášť. Pětislabičných slov bylo v korpusu poměrně málo, proto jsme je do analýzy nezařadili. Výsledky ukazují grafy 8.18a a b níže.



Obrázky 8.18a-b: Průměrné trvání krátkých (vlevo) a dlouhých (vpravo) vokálů ze seskupených dat podle délky slova vyjádřené ve slabikách. Svorky označují 95% interval spolehlivosti.

V grafech sice vidíme drobný trend krátkých vokálů se zkracovat se vzrůstajícím počtem slabik ve slově, u dlouhých vokálů je tento trend ale opačný. Efekt délky slova nicméně není ani pro jednu skupinu dat významný (jednofaktorová ANOVA:  $p > 0,05$ ).

Pro zkoumání vlivu délky slova ve fonémech jsme seskupené položky již nevyřazovali, protože by nebyl dostatek případů pro jednotlivé skupiny. Analyzována jsou slova se dvěma až desíti fonémami a průměry pro krátké a dlouhé vokály zvlášť zobrazují grafy 8.19a-b.



Obrázky 8.19a-b: Průměrné trvání krátkých (vlevo) a dlouhých (vpravo) vokálů ze seskupených dat podle délky slova vyjádřené ve fonémech. Svorky označují 95% interval spolehlivosti.

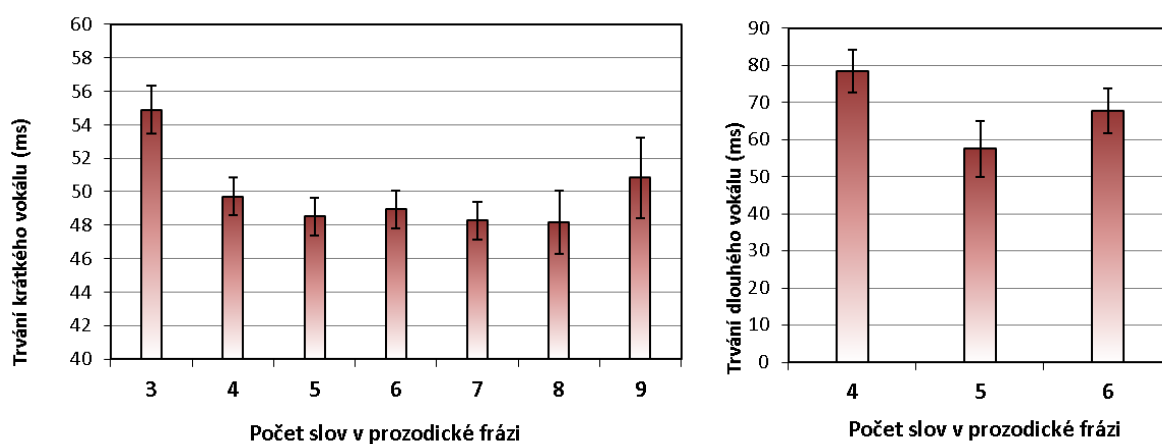
Efekt délky slova je statisticky významný pro obě skupiny vokálů – krátké:  $F(8, 1451) = 5,9$ ;  $p < 0,001$  i dlouhé:  $F(8, 806) = 21,9$ ;  $p < 0,001$ . U krátkých vokálů je trend zkracování dlouhých slov viditelný až u slov o délce devět a deset fonémů – rozdíly mezi průměry těchto dvou skupin a ostatních jsou signifikantní (Tukeyho post-hoc test:  $p < 0,02$ ), kromě vokálů ze slov se čtyřmi fonémami, které se od slov s devíti a desíti fonémami významně neliší. Také vokály ze šestifonémových slov se významně neliší od devítifonémových, od desíti- už ano.

U dlouhých vokálů je vidět zřetelý rozdíl mezi dvoufonémovými slovy a ostatními (Tukeyho post-hoc test:  $p < 0,005$ ) – nejdelší trvání vokálů ve dvoufonémových slovech (jde o slova jako *dá, já* nebo *ví*) může být způsobeno tím, že za těmito vokály nenásleduje konsonantická koda, zatímco u třífonémových ve velké většině případů ano. U krátkých tento efekt nepozorujeme, protože i dvoufonémová slova někdy obsahují kodu (*až, už*) a třífonémová slova často končí na vokál (např. *aby, ale, ani, jde*). Vliv struktury slabiky na trvání vokálu bude podrobněji prozkoumán dále.

Zkoumali jsme také vliv délky slova v milisekundách, tentokrát na neseskupených datech, ale stále pouze na mediálních slovech s vyřazenými instancemi *hm, jo, a no*. Protože jde v obou případech o spojité proměnné, měřena byla jejich korelace s pomocí Pearsonova  $r$ . Ukázalo se, že korelace trvání slova s trváním krátkého vokálu je velmi nízká až neexistující ( $r = 0,1$ ), u dlouhých vokálů je korelace o něco málo vyšší,  $r = 0,2$ . Obě korelace jsou statisticky

významné ( $p < 0,01$ ). Tento trend zůstává, i pokud jsou analyzována dvou-, tří- a čtyřslabičná slova zvlášť. Ve dvouslabičných je výraznější korelace u krátkých vokálů ( $r = 0,32$ ) než u dlouhých ( $r = 0,18$ ), obě jsou statisticky významné ( $p < 0,01$ ). Ve tříslabičných slovech je naopak o něco vyšší korelace u dlouhých vokálů ( $r = 0,46$ ) než u krátkých ( $r = 0,16$ ), obojí je opět statisticky signifikantní ( $p < 0,01$ ). Podobně je tomu i v případě čtyřslabičných slov, kde je korelační koeficient u krátkých vokálů  $r = 0,18$  ( $p < 0,001$ ) a u dlouhých  $r = 0,32$  ( $p < 0,001$ ).

Pro analýzu vlivu délky prozodické fráze na trvání vokálu jsme seskupili data opět pouze z mediálních slov podle mluvčího, hlásky a délky prozodické fráze vyjádřené nejprve ve slovech, potom ve slabikách. Položky vzniklé seskupením pěti a méně případů byly vyřazeny pouze v prvním případě (délka prozodické fráze ve slovech), protože ve druhém (délka prozodické fráze ve slabikách) bychom se zcela připravili o některé skupiny.

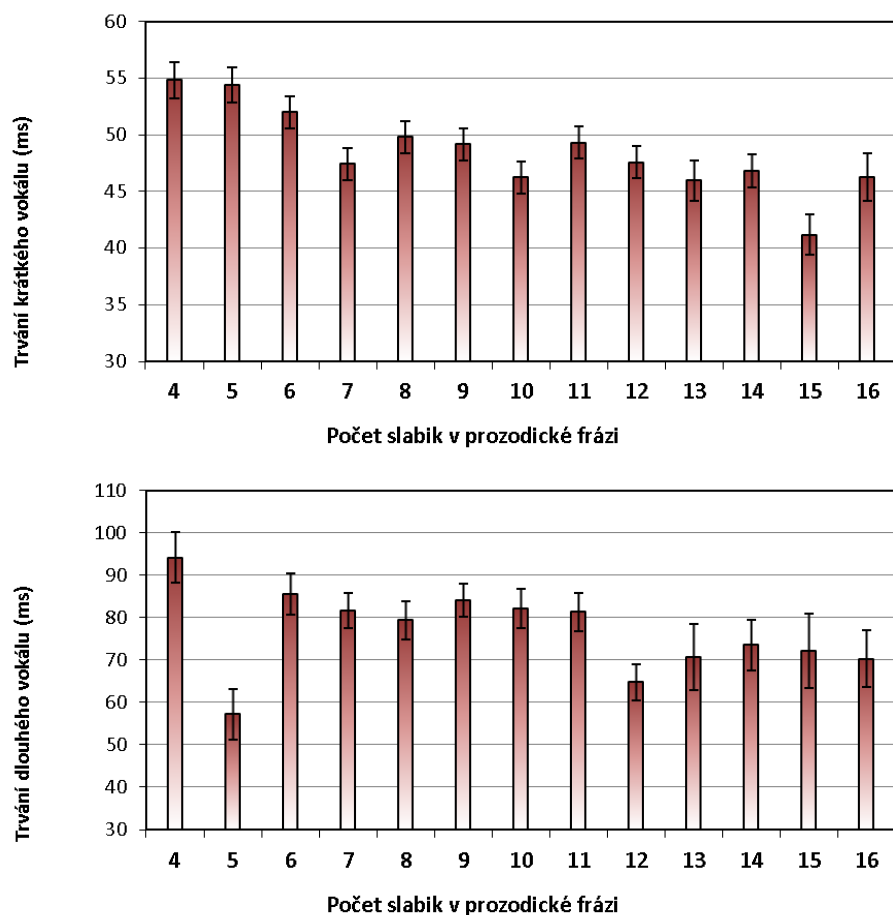


Obrázky 8.20a-b: Průměrné trvání krátkých (vlevo) a dlouhých (vpravo) vokálů ze seskupených dat podle délky prozodické fráze ve slovech. Svorky označují 95% interval spolehlivosti.

Na obrázcích 8.20a a b je znázorněno průměrné trvání vokálů v závislosti na délce prozodické fráze. U dlouhých vokálů byl dostatek dokladů pouze pro fráze sestávající ze čtyř, pěti a šesti slov. Efekt délky fráze je v obou případech významný, pro krátké vokály:  $F(6, 862) = 11,2$ ;  $p < 0,001$ , pro dlouhé vokály:  $F(2, 128) = 9,7$ ;  $p < 0,001$ .

Krátké vokály jsou průměrně nejdélejší ve frázích tvořených třemi slovy. Pak trvání klesá a o něco delší jsou až ve frázích tvořených devíti slovy. Statisticky významný je rozdíl mezi tříslavnými frázemi a všemi ostatními kromě devítislovných ( $p < 0,001$ ). Dlouhé vokály trvají nejdéle ve čtyřslovných frázích, kde jsou signifikantně delší než v pěti- a šestislovných ( $p \leq 0,03$ ).

Co se týče délky prozodické fráze vyjádřené ve slabikách, zde byly analyzovány čtyř- až šestnáctislabičné prozodické fráze. Výsledky zobrazují grafy 8.21a a b.



Obrázky 8.21a-b: Průměrné trvání krátkých (nahore) a dlouhých (dole) vokálů ze seskupených dat podle délky prozodické fráze ve slabikách. Svorcky označují 95% interval spolehlivosti.

Efekt délky prozodické fráze ve slabikách je významný pro krátké i dlouhé vokály – krátké:  $F(12, 1916) = 20$ ;  $p < 0,001$ , dlouhé:  $F(12, 1020) = 12,6$ ;  $p < 0,001$ . U krátkých vidíme jasný trend ke zkracování vokálů v dlouhých prozodických frázích. U dlouhých je tento trend poněkud méně zřetelný a u pěti- a dvanáctislabičných frází dochází k propadu hodnot. Podrobnější pohled na data odhaluje, že jde o vliv jednotlivých slov – propad u pětislabičných je způsoben tím, že většina případů je velmi krátké [i:] z gramatického slova *tím*. U dvanáctislabičných je většina případů také [i:], průměrně nejkratší z dlouhých vokálů.

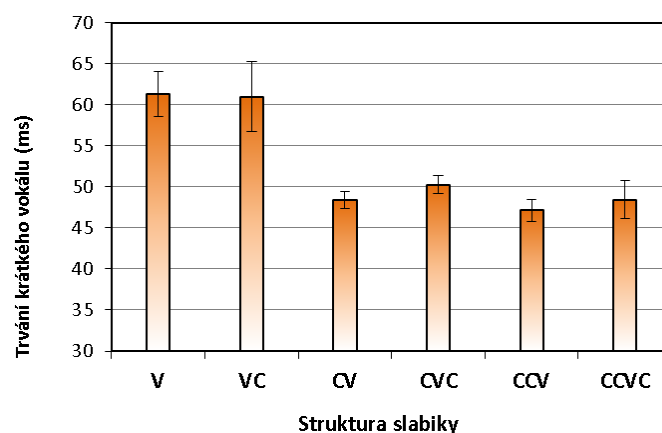
Co se týče délky prozodické fráze v milisekundách, u krátkých vokálů je korelace s trváním vokálu téměř nulová ( $r = 0,02$ ), a to statisticky významně ( $p = 0,005$ ). U dlouhých vokálů je lehce záporná ( $r = -0,09$ ) a také signifikantní ( $p < 0,001$ ). Můžeme tedy prohlásit, že jsme nenalezli souvislost mezi délkou prozodické fráze v milisekundách a trváním vokálů v mediálních slovech.



### 8.1.11 Vliv slabičné struktury na trvání vokálů

V tomto oddíle bude krátce analyzován vliv slabičné struktury na trvání vokálu v jádře slabiky – zda má přítomnost a složitost přetury nebo kody vliv i na vokalické jádro slabiky. Tato problematika by si zasloužila pojednat širěji a podrobněji, to však leží již mimo záběr předkládané práce.

Data z mediálních slov jsme opět seskupili, a to podle mluvčího, hlásky a typu slabiky. Vyřadili jsme případy, které vznikly seskupením pouze pěti a méně položek. Z důvodu nedostatku dokladů byly analyzovány pouze krátké vokály. Průměry pro jednotlivé typy slabik jsou zobrazeny v grafu 8.22.



Obrázek 8.22: Průměrné trvání krátkých vokálů ze seskupených dat podle struktury slabiky. Svorky označují 95% interval spolehlivosti.

Efekt typu slabiky je statisticky významný:  $F(5, 476) = 24,1; p < 0,001$ . Z obrázku je zřetelné, že ke zkracování vokálu v jádru slabiky dochází ve slabikách s přeturoou – přičemž na počtu konsonantů v přeture nezáleží. Přítomnost kody nemá na průměrné trvání žádný vliv. Vokály ve slabikách typu V a VC jsou signifikantně delší než ostatní skupiny (Tukeyho post-hoc test:  $p < 0,001$ ). Je ovšem pravděpodobné, že různé vokály se nebudou chovat stejně – tento faktor si jistě v budoucnu zaslouží bližší prozkoumání.

## 8.2 Průměrné trvání vokálů

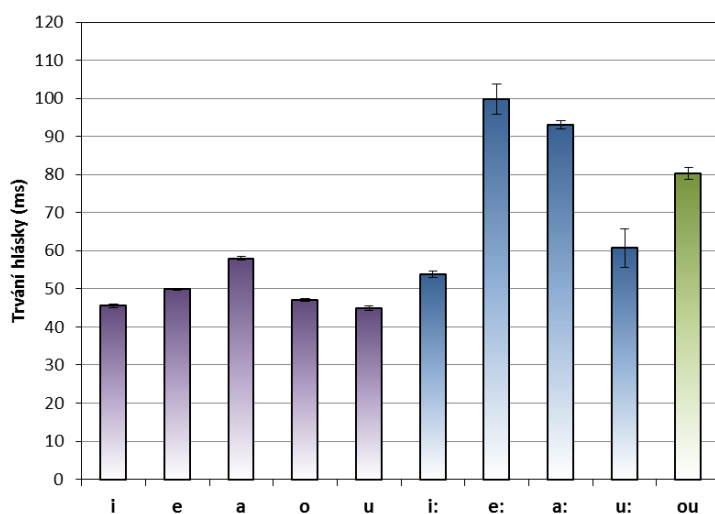
Pro zajištění co nejrepresentativnějšího průměrného trvání vokálů, které by bylo ovlivněno co nejméně dalšími faktory, jsme vyřadili vokály z finálních slov prozodických frází (ať už končily předělem hloubky 3 nebo 4), tedy i ze slov, která tvořila samostatnou prozodickou frázi, pak také vokály po rázu a v okolí elízí (tj. ty, které obsahovaly víc než jednu značku fonému – pro ilustraci viz obr. 7.2). Také byla opět vyřazena slova *hm*, *jo* a *no*.

Výsledný počet byl 26654 vokálů. Naměřené hodnoty zobrazuje tabulka 8.5.

	Počet	Průměr	Sm.odch.	Minimum	Maximum	Medián	10. percent.	90. percent.
<b>Krátké vokály</b>								
<b>a</b>	3151	<b>58</b>	<b>14,8</b>	12,6	131,1	56,9	39,9	77,1
<b>e</b>	9010	<b>49,9</b>	<b>14,6</b>	9	151,7	48,5	32	69
<b>i</b>	3624	<b>45,6</b>	<b>14,6</b>	7	116,3	44,1	28,4	64,5
<b>o</b>	4984	<b>47</b>	<b>13,9</b>	15,2	136	46	30,4	65,1
<b>u</b>	1632	<b>45</b>	<b>13,8</b>	8,8	105,2	43,9	28,8	63,2
<b>Dlouhé vokály</b>								
<b>a:</b>	1691	<b>93</b>	<b>23</b>	20,6	193,4	92,2	64,6	122
<b>e:</b>	89	<b>99,8</b>	<b>19</b>	66,4	166,3	96,6	80,2	126,4
<b>i:</b>	1824	<b>53,8</b>	<b>17,3</b>	7,7	145,4	51,8	33	76,6
<b>u:</b>	100	<b>60,7</b>	<b>25,5</b>	22,2	153,4	55,3	33,3	96,2
<b>Diftong</b>								
<b>ou</b>	549	<b>80,3</b>	<b>18,1</b>	32,5	157,7	80	56,3	102,6
<b>Celkem</b>	26654	<b>53,3</b>	<b>19,7</b>	7	193,4	50	32	78,2

Tabulka 8.5: Reprezentativní hodnoty trvání českých vokálů naměřené na vybraných částech korpusu Minidialogy-H. Hodnoty ve třetím až devátém sloupci jsou uvedeny v milisekundách.

Průměrná trvání jednotlivých vokálů jsou převedená do grafické podoby na obrázku 8.23.



Obrázek 8.23: Reprezentativní průměrná trvání krátkých a dlouhých vokálů a diftongu [ou]. Svorky označují 95% interval spolehlivosti.

U krátkých vokálů je zřetelně vidět tendence ke kratšímu trvání zavřených vokálů, oproti otevřenému [a], které je průměrně nejdelší. Středové vokály [e] a [o] pak spadají mezi ně. Srovnáme-li tyto údaje s dřívějšími výsledky z literatury (oddíl 5.1.1, obr. 5.1a), vidíme, že co se týče rozsahu hodnot trvání (krátké vokály trvají mezi 45 a 58 ms), nejvíce se blížíme hodnotám Podlipského et al. (2009), ostatní jsou výrazně delší. Rozdíly mezi trváním podle otevřenosti vokálu pak jsou nejpodobnější Chlumskému (1928) a Kaiserové (1964).

U dlouhých vokálů tuto tendenci porušuje [e:], které je průměrně nejdelší. Zavřené [i:] a [u:] však zůstává nejkratší. Porovnáme-li to opět se staršími daty (obr. 5.1b), rozsahem hodnot se pohybujeme opět poblíž výsledků Podlipského et al. (2009). Rozdíly mezi jednotlivými vokály jsou ovšem jiné než v citovaných studiích, nejkratší trvání [i:] a [u:] nicméně zůstává. Průměrné trvání diftongu [ou̯] je o více než 20 ms kratší než v ostatních studiích.

### 8.2.1 Dlouhé vs. krátké vokály

Tabulka 8.6 zobrazuje poměry a rozdíly reprezentativního trvání krátkých vokálů a jejich dlouhých protějšků.

Vokály	Poměr	Rozdíl (ms)
i:/i	1,2	8,2
e:/e	2,0	49,9
a:/a	1,6	35
u:/u	1,3	15,7

Tabulka 8.6: Poměr reprezentativního trvání dlouhého vokálu vůči jeho krátkému protějšku.

Vokál [e:] se vůči [e] prodlužuje dvakrát, což je nejvíce. Naopak nejméně je prodlouženo [i:], které je pouze 1,2krát delší než jeho krátký protějšek. Výsledky pro [ɪ], [a] a [u] jsou srovnatelné s novějšími studii citovanými v kap. 5.1.1 (tabulka 5.2), výsledky pro [e] jsou pravděpodobně zkreslené malým počtem případů a také slov, ze kterých je většina realizací (konkrétně zejména slovo *ne*). Nicméně poměr 2:1 vyšel u [e] vs. [e:] i Studenovskému (2012). Pro diskusi výsledků viz oddíl 8.5.

## 8.3 Konsonanty

Pro analýzu jednotlivých konsonantů jsme opět vyřadili slova *hm*, *jo* a *no*. Také byly vyřazeny neznělé explozivní následující po pauze, jejichž trvání je do jisté míry určeno arbitrárně anotátorem. Dále jsme vyřadili konsonanty zařazené jako slabičné, které zbyly samostatné ve slabice po elizi jejího jádra (viz např. obrázek 7.3). Celkem analyzujeme 64 550 konsonantů. V následujících sekcích budou popsány výsledky pro jednotlivé skupiny konsonantů rozdělených podle způsobu tvoření.

Na rozdíl od vokálů se budeme u konsonantů zabývat také pozicí ve slabice, tj. zda je konsonant umístěn v přetučce či v kodě.

### 8.3.1 Explozivy

V korpusu Minidialogy-H se vyskytuje celkem 21 917 exploziv (což je přibližně třetina všech konsonantů). Počty výskytů jednotlivých hlásek zobrazuje tabulka 8.7a a tabulka 8.7b pak rozděluje tyto četnosti podle pozice konsonantu ve slabice.

Neznělé	Poč. výskytů	Znělé	Poč. výskytů
p	3474	b	2095
t	7105	d	3614
ť	1166	d'	898
k	2781	g	784

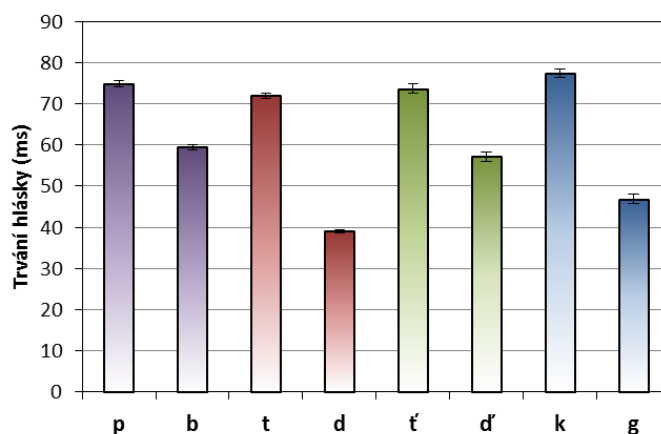
Tabulka 8.7a: Počty výskytů jednotlivých exploziv v korpusu Minidialogy-H.

Exploziva	Prétura	Koda
p	3406	68
b	2095	
t	5193	1912
d	3305	309
ť	1086	80
d'	612	286
k	1981	800
g	499	285
<b>Celkem</b>	<b>18177</b>	<b>3740</b>

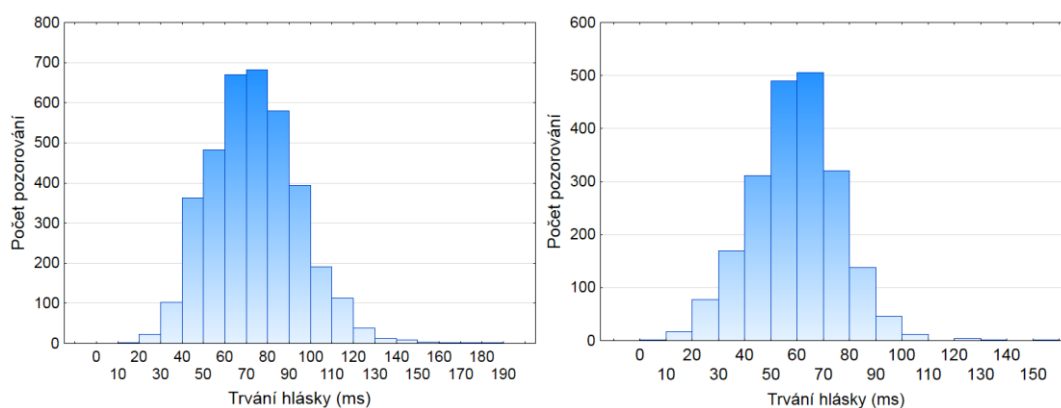
Tabulka 8.7b: Počty výskytů jednotlivých exploziv v korpusu Minidialogy-H podle pozice ve slabice (v préture či v kodě).

Ze druhé tabulky vidíme, že většina exploziv se v našem materiálu vyskytuje v préturách, [b] se dokonce v kodě nevyskytuje vůbec. To koreponduje s údaji o frekvenci slabičných typů ze starší literatury, která se shoduje na tom, že slabiky s nulovou kodou tvoří tři čtvrtiny ze všech slabik v mluvených projevech (Kučera & Monroe, 1968 a Ludvíková, 1987). Znělé explozivy v kodách jsou výsledkem asimilace znělosti.

Průměrné neseskupené trvání jednotlivých exploziv zobrazuje graf 8.24. Z něj je na první pohled zřetelné, že znělé explozivy jsou výrazně kratší než jejich neznělé protějšky (statistické ověření této hypotézy viz níže).

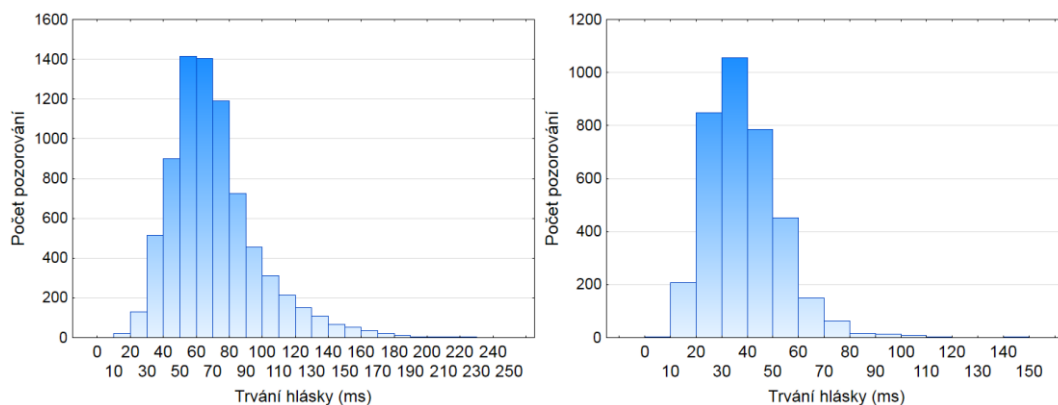


Obrázek 8.24: Průměrné neseskupené trvání znělých a neznělých exploziv. Svorky označují 95% interval spolehlivosti.



Obrázky 8.25a-b: Histogramy trvání bilabiálních exploziv [p] (vlevo) a [b] (vpravo) ve všech pozicích a všech hláskových okolích.

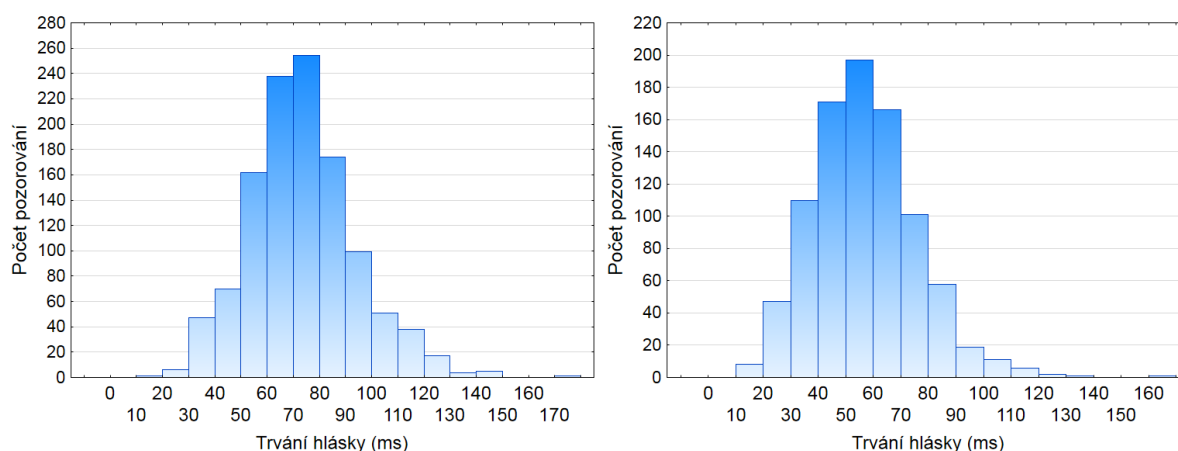
Obrázky 8.25a a b zobrazují rozložení hodnot trvání bilabiálních exploziv [p] a [b]. Neznělé [p] má průměrné trvání 74,9 ms se směr. odchylkou 20,4 ms a mediánem 73,4 ms. Minimální nalezené trvání je 19,9 ms, maximální 187 ms, 10. percentil 48,6 ms a 90. percentil 100,8 ms. [p] v préture trvá průměrně 75 ms se směr. odchylkou 20,4 ms a mediánem 73,7 ms. V kodě je pak o něco kratší, průměrně 63,7 ms, sm. odch. 18 ms a medián 61 ms. (Statistické ověření významnosti rozdílu není možné kvůli příliš nízkému počtu případů ve druhé skupině.) Znělé [b], které se vyskytuje jen v préture, trvá průměrně 59,4 ms se směrodatnou odchylkou 16,7 ms, mediánem 59,9 ms, minimem 8,7 ms, maximem 150,8 ms, 10. percentilem 37,6 ms a 90. percentilem 80 ms.



Obrázky 8.26a-b: Histogramy trvání exploziv [t] (vlevo) a [d] (vpravo) ve všech pozicích a všech hláskových okolicích.

Rozložení hodnot trvání alveolárních exploziv [t] a [d] zobrazují histogramy 8.26a a b. Neznělá exploziva má průměrné trvání 72 ms se sm. odchylkou 27,9 ms a mediánem 67,2 ms. Minimální nalezené trvání [t] je 11,6 ms, maximální 253,7 ms. 10. percentil je 41,2 ms a 90. percentil 107,9 ms. Podíváme-li se na jeho trvání v préturách, je průměrně 67,7 ms, směr. odchylka 20,9 ms a medián 65,9 ms. V kodách pak [t] trvá průměrně o něco déle, 83,6 ms, sm. odch. 39 ms a medián 74,7 ms. Výsledek t-testu na datech seskupených podle mluvčího, hlásky a pozice ve slabice ukazuje na významnost tohoto rozdílu:  $t(66) = 10,1; p < 0,001$ .

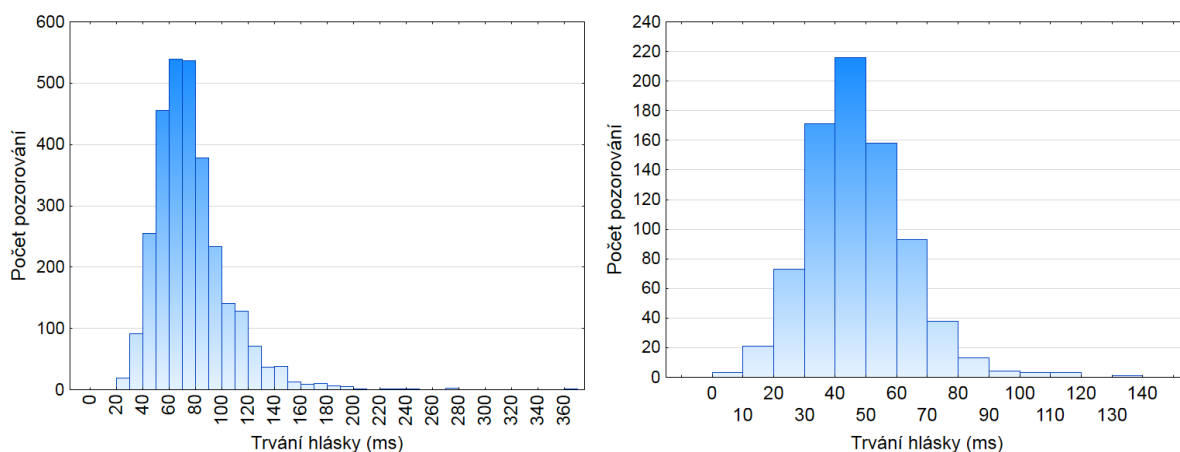
Znělé [d] má průměrné trvání 39 ms, směrodatnou odchylku 14,7 ms, medián 36,9 ms, minimum 8,8 ms, maximum 149,9 ms, 10. percentil 22,7 ms a 90. percentil 56,8 ms. [d] v prétuře trvá průměrně 38,5 ms se směr. odchylkou 14,7 ms a mediánem 36,5 ms. V kodě má průměrné trvání o něco delší, 44,2 ms, směr. odchylka 13,8 ms a medián 42,5 ms. (Statistickou významnost rozdílu nelze stejně jako u [b] ověřit kvůli příliš nízkému počtu případů ve druhé skupině – jen desetina z případů ve skupině první.)



Obrázky 8.27a-b: Histogramy trvání palatálních exploziv [tʃ] (vlevo) a [dʃ] (vpravo) ve všech pozicích a všech hláskových okolicích.

V histogramech 8.27a a b jsou vidět četnosti hodnot trvání palatálních exploziv [tʰ] a [dʰ]. [tʰ] trvá průměrně 73,6 ms, má směr. odchylku 20,5 ms a medián 72 ms. Nejkratší nalezené trvá 19,7 ms, nejdelší 170,4 ms. 10. percentil je 48,7 ms a 90. percentil 99,9 ms. [tʰ] v přetúře má průměrné trvání 73,8 ms, směr. odchylku 19,7 ms a medián 72 ms. [tʰ] v kodě je nepatrně kratší, průměr 71 ms, se směr. odchylkou 29,1 ms a mediánem 63,7 ms. Rozdíl mezi přetúrou a kodou není statisticky významný (t-test na datech seskupených podle mluvčího a pozice hlásky ve slabice:  $p = 0,37$ ).

Znělé [dʰ] má celkové průměrné trvání 57,1 ms, směr. odchylku 18,9 ms a medián 56 ms. Minimální změřené trvání je 13 ms a maximální 160,1 ms. 10. percentil je 33,4 ms a 90. percentil 81,2 ms. V přetúře je dlouhé průměrně 59,7 ms, směr. odchylka 18,7 ms a medián 58,3 ms. V kodě je pak [dʰ] o něco kratší, průměrně 51,4 ms, směr. odchylka 18 ms a medián 49,5 ms. Tento rozdíl je signifikantní podle t-testu na datech seskupených podle mluvčího a pozice ve slabice:  $t(66) = 4,7; p < 0,001$ .



Obrázky 8.28a-b: Histogramy trvání velárních exploziv [k] (vlevo) a [g] (vpravo) ve všech pozicích a všech hláskových okolích.

Rozložení hodnot trvání velárních exploziv [k] a [g] zobrazují histogramy 8.28a a b.

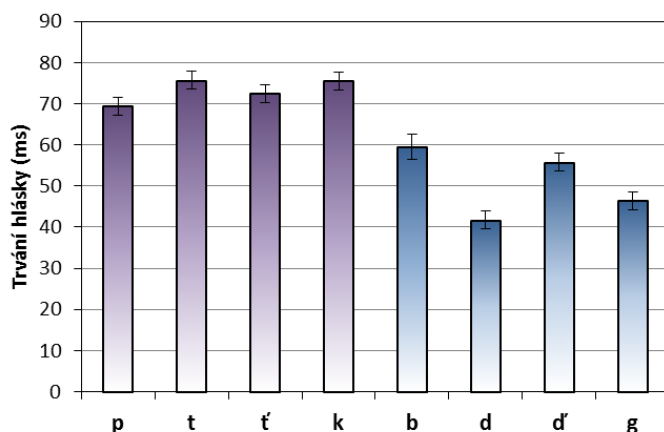
Vezmeme-li v potaz obě pozice ve slabice, trvá [k] průměrně 77,4 ms, má směr. odchylku 27,8 ms a medián 72,9 ms. Minimálně trvá 20,4 ms, maximálně 361,1 ms. 10. percentil je 48 ms a 90. percentil 113,2 ms. V přetúře má [k] průměrné trvání 80 ms, směr. odchylku 24,5 ms a medián 76,5 ms. V kodě je potom průměr 70,9 ms, směr. odchylka 33,8 ms a medián 62,4 ms. Po seskupení dat podle mluvčího, hlásky a pozice ve slabice a provedení t-testu se tento rozdíl ukazuje jako významný:  $t(66) = 5,3; p < 0,001$  – v kodě jsou [k] kratší než v přetúře.

Jeho znělý protějšek, [g], trvá průměrně 46,8 ms, směr. odchylka 16,1 ms a medián 45,8 ms. Minimální naměřené trvání je 7,1 ms, maximální 132,4 ms, 10. percentil je 28,7 ms a 90. percentil 66,2 ms. V přetúře trvá [g] průměrně 48,7 ms se směr. odchylkou 16,4 ms a

mediánem 48,2 ms. V kodě je průměrné trvání [g] o něco kratší, 43,5 ms, směr. odchylka 14,9 ms a medián 42,5 ms. Rozdíl je opět statisticky významný při seskupení dat podle mluvčího a pozice konsonantu ve slabice:  $t(66) = 3,3; p = 0,001$ .

Provedeme-li na seskupených datech (podle mluvčího, hlásky a pozice ve slabice) od všech hlásek najednou jednofaktorový test ANOVA, zjistíme, že efekt pozice konsonantu ve slabice (tj. zdali je v prětuře či v kodě) není statisticky významný:  $F(1, 508) = 1,2; p = 0,26$ . Je to způsobeno tím, že různé explozivny se chovají různě – zatímco [p], [dʰ], [k] a [g] jsou v kodách průměrně kratší, [t] a [d] jsou naopak delší.

Dále byl otestován také efekt znělosti – ověření hypotézy, že neznělé explozivny jsou delší než jejich znělé protějšky. Výsledek jednofaktorového testu ANOVA (na stejné seskupených datech jako výše) s faktorem ZNĚLOST ukazuje na vysoce významný rozdíl:  $F(1, 508) = 668,1; p < 0,001$ , který tuto hypotézu podporuje. Průměry skupin jednotlivých hlásek ukazuje obrázek 8.28. K ověření, zda jsou rozdíly významné pro všechny dvojice hlásek byla provedena ještě jedna jednofaktorová ANOVA pouze s faktorem HLÁSKA. Tukeyho post-hoc test pak ukazuje, že k rozdílu mezi znělými a neznělými přispívají všechny dvojice exploziv,  $p < 0,001$ . Zároveň je z něj možné vyčíst, že v případě neznělých exploziv se od sebe významně liší trváním [p] a [k] ( $p = 0,003$ ) a [p] a [t] ( $p = 0,002$ ), ostatní dvojice rozdílné nejsou. V případě znělých se od sebe trváním neliší dvojice [b]-[dʰ] a [d]-[g].



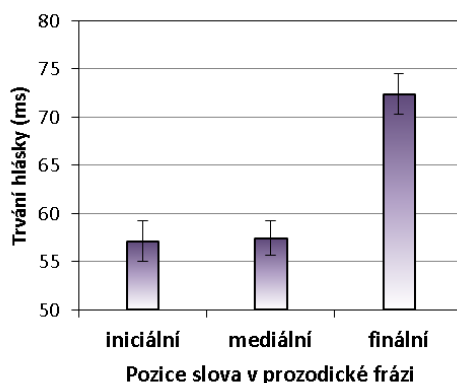
Obrázek 8.29: Průměrné seskupené trvání znělých a neznělých exploziv. Svorcky označují 95% interval spolehlivosti.

Podle trvání se tak explozivny seskupují do dvou skupin – bilabiální mají podobné trvání jako palatální a alveolární jako velární.



### 8.3.1.1 Vliv frázového zpomalování na trvání exploziv

V první analýze byly data seskupena podle hlásky, mluvčího a pozice slova v prozodické frázi. Finální slova byla vzata v úvahu jen před předělem hloubky 4 (tj. plným prozodickým předělem). Po seskupení jsme vyřadili případy, které vznikly seskupením pouze pěti a méně hodnot. Také byla vyřazena individuální slova, protože tvořila pouze malé procento případů. Průměrné trvání konsonantů v iniciálních, mediálních a finálních slovech zobrazuje graf 8.30.



Obrázek 8.30: Průměrné trvání exploziv ze seskupených dat rozdělených podle pozice slova v prozodické frázi. Svorcky označují 95% interval spolehlivosti.

Efekt pozice je vysoce významný:  $F(2, 678) = 71$ ;  $p < 0,001$  (jednofaktorová ANOVA), přičemž tento rozdíl je způsoben finálními slovy, v nichž je trvání exploziv výrazně delší (Tukeyho post-hoc test:  $p < 0,001$ ). Rozdíl mezi iniciálními a mediálními slovy je nevýznamný.

Tabulka 8.8 zobrazuje koeficienty závěrového zpomalování, které jsou spočítány jako poměr (multiplikativní koeficient), příp. rozdíl (aditivní koeficient) průměrného trvání exploziv v mediálních a finálních slovech (paralelně s vokály, viz tabulka 8.2 výše). Explozivy [tʰ], [dʰ] a [g] nejsou v seskupených datech zastoupeny v obou skupinách (konkrétně [dʰ] se vyskytuje pouze ve finálních slovech a [tʰ] a [g] se naopak ve finálních slovech nevyskytuje vůbec).

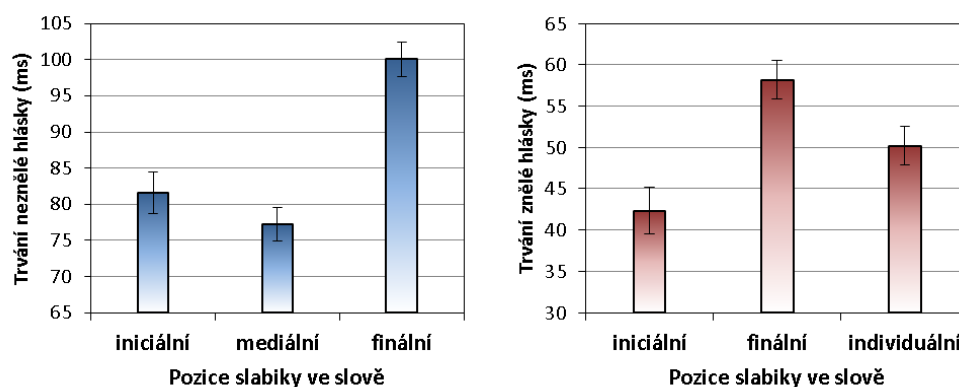
Exploziva	Multipl. koef.	Adit. koef. (ms)
p	1,09	6,5
b	1,2	11,5
t	1,53	31,7
d	1,44	14
k	1,34	22,3
průměr	1,32	17,2

Tabulka 8.8: Koeficienty průměrných rozdílů mezi explozivami z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

Z tabulky je vidět, že nejvíce se ve finálních slovech prodlužuje [t], a to na 1,5násobek svého trvání v mediálních slovech. Naopak [p] se neprodužuje téměř vůbec.

Stejně jako u vokálů, i zde bude prozkoumáno, nakolik je tento efekt způsoben pouze finálními konsonanty (tj. v kodě finální slabiky).

Data byla seskupena podle mluvčího, hlásky, pozice hlásky ve slabice, pozice slabiky ve slově a pozice slova v prozodické frázi. Individuální slova byla zahrnuta mezi finální. U finálních slov pak byla opět vzata v úvahu jen ta, která předcházela plnému prozodickému předělu. Vyřazeny byly opět položky, které vznikly seskupením méně než šesti případů.

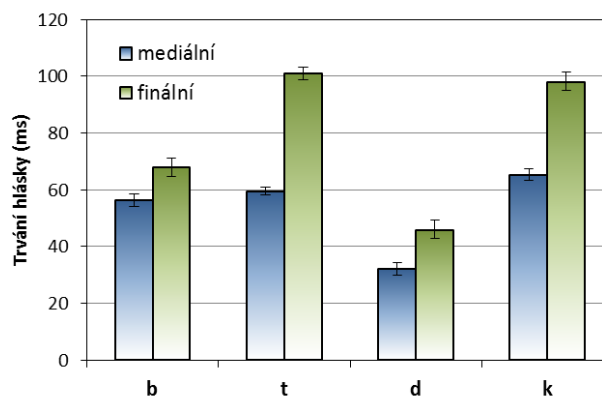


Obrázky 8.31a-b: Průměrné trvání exploziv ze seskupených dat ve finálních slovech podle pozice slabiky ve slově. Neznělé explozivy jsou zobrazeny v grafu vlevo a znělé vpravo. Svorčky označují 95% interval spolehlivosti.

Výsledky analýz ukazují obrázky 8.31a a b. V grafu vlevo jsou zobrazeny neznělé explozivy, kde vidíme jasnou tendenci k prodlužování hlásek ve finálních slabikách. Efekt pozice je signifikantní:  $F(2, 268) = 101,8; p < 0,001$ . Podle Tukeyho post-hoc testu k významnosti přispívají finální slabiky, rozdíl mezi iniciálními a mediálními slabikami je nesignifikantní.

Vpravo jsou pak zobrazeny znělé explozivy – protřídění dat způsobilo ztrátu konsonantů v mediálních slabikách, ale naopak na rozdíl od neznělých zde jsou zastoupeny i individuální slabiky (tzn. případy, kdy finální slovo v prozodické frázi je tvořeno pouze jednou slabikou). Stejně jako u neznělých jsou explozivy ve finálních slabikách výrazně nejdelší a efekt pozice je významný:  $F(2, 165) = 31,6; p < 0,001$ . Tukeyho post-hoc test prozrazuje, že rozdíly mezi všemi dvojicemi pozic jsou významné. Individuální slabiky se trváním nacházejí někde mezi iniciálními a finálními. Zde je ovšem nutné podotknout, že všechny případy exploziv v individuálních slabikách jsou tvořeny hláskou [d], která je sama o sobě výrazně nejkratší.

Porovnání mediálních exploziv a exploziv z finálních slabik finálních slov (jako u vokálů, viz oddíl 8.1.8 výše) není možné pro všechny hlásky, protože nejsou zastoupeny ve všech skupinách. Provedli jsme tuto analýzu tedy pouze pro [b], [k] a alveolární explozivy.



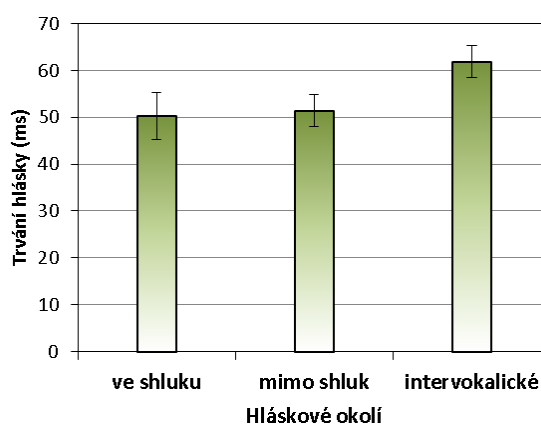
Obrázek 8.32: Průměrné trvání vybraných exploziv v mediálních slovech (modrá) a ve finálních slabikách finálních slov (zelená). Svorky označují 95% interval spolehlivosti.

Efekt pozice je dle předpokladů vysoce významný  $F(1, 570) = 702; p < 0,001$ , významná je také interakce HLÁSKA\*POZICE:  $F(3, 570) = 70,5; p < 0,001$ . Rozdíly u jednotlivých hlásek jsou také všechny signifikantní (Tukeyho post-hoc test:  $p < 0,001$ ).

### 8.3.1.2 Vliv hláskového okolí na trvání exploziv

Aby bylo možno prozkoumat, jak se liší trvání exploziv v závislosti na jejich hláskovém okolí, rozdělili jsme explozivy na tři skupiny. Intervokalické, tedy nacházející se mezi dvěma vokály, explozivy ve shlucích, tedy ty, které jsou součástí konsonantického shluku uvnitř slabiky, a všechny ostatní, což jsou explozivy, které mají v okolí další konsonant, ten ale není součástí stejné slabiky. Pro tuto analýzu jsme explozivy seskupili podle mluvčího, hlásky a hláskového okolí a využili pouze mediální slova. Položky, které vznikly seskupením méně než šesti případů, byly vyřazeny.

Obrázek 8.33 níže ukazuje průměry jednotlivých skupin.



Obrázek 8.33: Průměrné trvání exploziv v mediálních slovech podle hláskového okolí. Svorky označují 95% interval spolehlivosti.

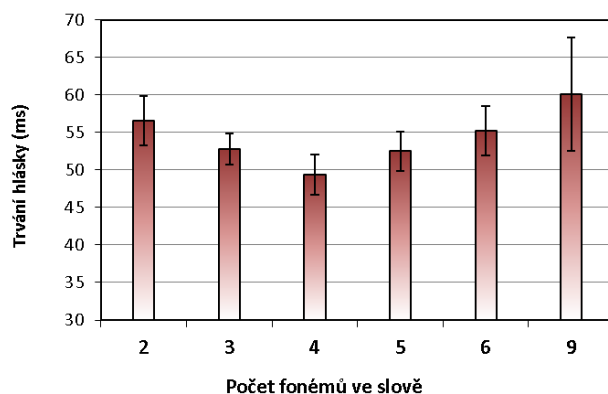
Podle jednofaktorového testu ANOVA je vliv hláskového okolí na trvání exploziv signifikantní – intervokální jsou signifikantně delší než explozivy v konsonantickém shluku uvnitř slabiky nebo mimo slabiku:  $F(2, 457) = 33,3; p < 0,001$ . Rozdíl mezi explozivami ve shluku a mimo shluk je nevýznamný. Dvoufaktorový test pro jednotlivé explozivy zvlášť není možné udělat z důvodu nedostatku případů v některých skupinách.

Nicméně zdá se, že jediná hláska, která neodpovídá tomuto vzorci, je [d], které je v intervokální pozici naopak kratší. Lze předpokládat, že se v některých případech artikulace redukuje na pouhý alveolární švih, a tím vlastně přestává být fonetickou explozivou.

### 8.3.1.3 Vliv délky vyšší jednotky na trvání exploziv

Data pouze z mediálních slov byla seskupena podle mluvčího, hlásky a délky slova (vyjádřené nejprve ve slabikách a následně ve fonémech), ve druhé analýze pak podle mluvčího, hlásky a délky prozodické fráze (vyjádřené ve slovech a slabikách), položky seskupené z pěti a méně případů byly vyřazeny.

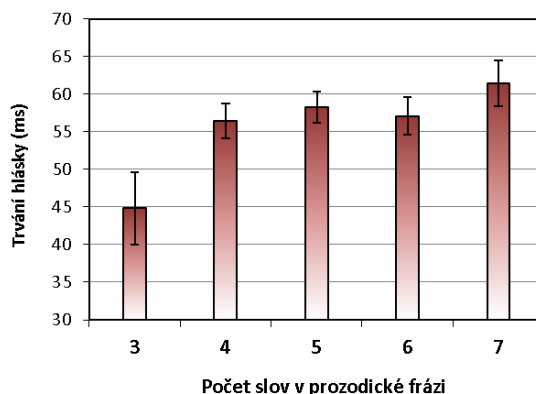
Efekt délky slova vyjádřené ve slabikách není významný. Vyjádříme-li délku slova počtem fonémů, získáme zajímavý nelineární trend, který ukazuje obr. 8.34. Explozivy jsou nejkratší ve slovech se čtyřmi fonémy, v kratších, ale i delších trvají déle. Efekt délky slova je zde významný:  $F(5, 531) = 3,3; p = 0,006$ .



Obrázek 8.34: Průměrné trvání exploziv ze seskupených dat podle délky slova vyjádřené ve fonémech. Svorky označují 95% interval spolehlivosti.

Neseskupená data ze všech exploziv byla následně korelována s trváním slova vyjádřeným v milisekundách. Korelace je mírně pozitivní (Pearsonovo  $r = 0,24$ ) a statisticky významná ( $p < 0,01$ ). Nalézáme tedy mírný trend prodlužování exploziv ve slovech, která trvají déle.

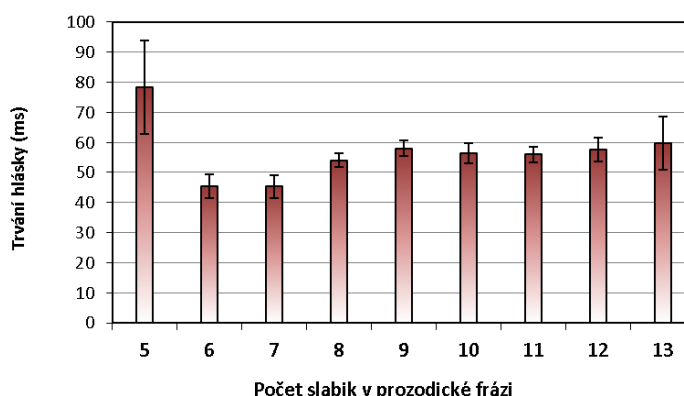
Délka prozodické fráze vyjádřená ve slovech vykazuje statisticky signifikantní vliv na trvání exploziv – jak je vidět z obrázku 8.35.



Obrázek 8.35: Průměrné trvání exploziv ze seskupených dat podle délky prozodické fráze vyjádřené ve slovech. Svorky označují 95% interval spolehlivosti.

Jednofaktorová ANOVA:  $F(4, 604) = 8,5; p < 0,001$ . Explozivy v krátkých prozodických frázích jsou nejkratší a zároveň nejvariabilnější, a posléze se prodlužují. Podle Tukeyho post-hoc testu je statisticky významný rozdíl mezi explozivami v tříslavných frázích a všemi ostatními skupinami ( $p < 0,001$ ), další rozdíly již významné nejsou.

Vyjádříme-li délku prozodické fráze ve slabikách, získáme podrobnější obrázek. Pětislabičných prozodických frází je málo a data mají velkou variabilitu, nicméně průměrně jsou zde explozivy nejdelší. Od šestislabičných frází dále vidíme opět trend k prodlužování s délkou fráze. Efekt délky fráze je statisticky významný:  $F(8, 564) = 8,3; p < 0,001$ .



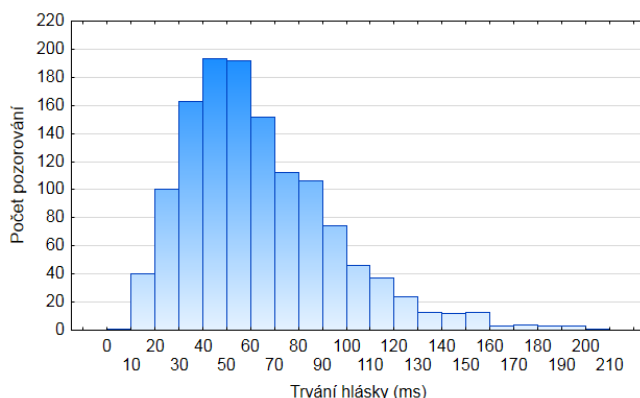
Obrázek 8.36: Průměrné trvání exploziv ze seskupených dat podle délky prozodické fráze vyjádřené ve slabikách. Svorky označují 95% interval spolehlivosti.

Tukeyho post-hoc test navíc ukazuje, že významný je rozdíl mezi šesti- a sedmislabičnými frázemi (v nichž jsou explozivy kratší) a všemi ostatními skupinami kromě třináctislabičných.

Vezmeme-li ovšem trvání explozivny a trvání prozodické fráze v milisekundách, zjistíme, že korelace (na neseskupených datech) je téměř nulová ( $r = 0,09$ ), což je nicméně statisticky významné ( $p < 0,05$ ).

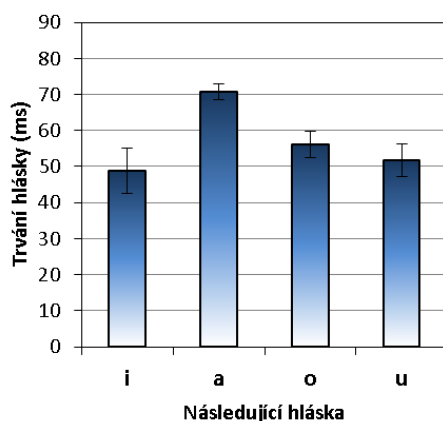
### 8.3.2 Ráz

V histogramu 8.37 je zobrazeno rozložení hodnot trvání rázu. V úvahu jsme vzali pouze rázy, které nenásledují po pauze, a jejich trvání tak nezávisí na libovůli anotátora.



Obrázek 8.37: Histogram trvání glotalizací nenásledujících po pauze.

Rázů nenásledujících po pauze je v našem korpusu 1292. Jejich průměrné trvání je 63,8 ms se směrodatnou odchylkou 31,7 ms a mediánem 57,7 ms. Nejkratší nalezený ráz trvá 9,6 ms, nejdelší 209,4 ms.<sup>6</sup> 10. percentil je 28,9 ms a 90. percentil 106,6 ms. Průměrným trváním leží ráz někde mezi znělými a neznělými explozivami.

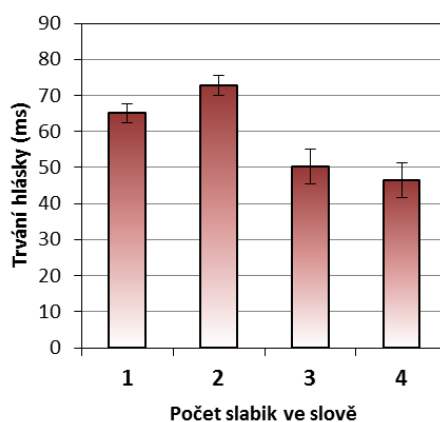


Obrázek 8.38: Průměrné trvání rázů nenásledujících po pauze před různými vokály. Svorky označují 95% interval spolehlivosti.

<sup>6</sup> Rázů delších než 120 ms je pouze necelých 6 %. V některých případech by možná bylo vhodnější označit pauzu a za ní ráz s kratším trváním, v jiných případech ale je glotalizace skutečně takto dlouhá.

Prozkoumali jsme v našem materiálu také vliv některých dalších faktorů na trvání rázu. Na obrázku 8.38 jsou průměry trvání rázu před různými vokály (rázy před [e] jsou pouze tři a nejsou do této analýzy zahrnuty). Jednofaktorová ANOVA potvrzuje, že efekt následujícího vokálu je významný:  $F(3, 1284) = 34,9$ ;  $p < 0,001$  a před [a] jsou rázy signifikantně delší než před ostatními vokály (Tukeyho post-hoc test:  $p < 0,001$ ).

Obrázek 8.39 pak znázorňuje trvání rázů nenásledujících po pauze podle délky slova vyjádřené ve slabikách. Efekt délky slova je opět významný:  $F(3, 1288) = 40$ ;  $p < 0,001$ , z Tukeyho post-hoc testu vyplývá, že signifikantní jsou všechny rozdíly ( $p < 0,001$ ) kromě tří a čtyřslabičných slov, které se od sebe neliší. Rázy na začátku dvouslabičných slov jsou tedy nejdelší a na začátku tří- a čtyřslabičných slov nejkratší.



Obrázek 8.39: Průměrné trvání rázů nenásledujících po pauze podle délky slova ve slabikách. Svorcky označují 95% interval spolehlivosti.

Dále byl nalezen také významný efekt typu slabiky:  $F(1, 1289) = 29,1$ ;  $p < 0,001$ . Rázy tvořící préturu otevřené slabiky, tedy CV, jsou signifikantně delší (průměrné trvání 66,5 ms) než rázy tvořící préturu slabiky CVC (průměrné trvání 55,5 ms).

### 8.3.3 Frikativy

Frikativ je v analyzované části korpusu Minidialogy-H celkem 15 549, což je přibližně čtvrtina všech analyzovaných konsonantů.

Neznělé	Poč. výskytů	Znělé	Poč. výskytů
f	874	v	2397
s	5420	z	1830
š	2146	ž	1083
ch	859	h	940

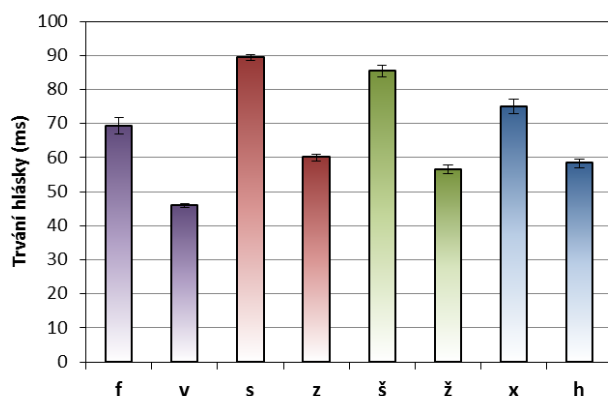
Tabulka 8.9a: Počty výskytů jednotlivých frikativ v korpusu Minidialogy-H.

V tabulce 8.9a jsou uvedeny celkové počty výskytů jednotlivých frikativ. V korpusu se též vyskytuje znělý alofon [ɣ] (ve slově *bůhvi*), nicméně realizací je pouze 16, takže jsme se rozhodli jej do analýz nezařadit. V tabulce 8.9b jsou pak počty výskytů rozdělené podle pozice hlásky ve slabice (tj. v prětuře či v kodě).

Exploziva	Prětura	Koda
f	642	232
v	2256	141
s	4501	919
z	1763	67
š	809	1337
ž	674	409
ch	370	489
h	937	3
<b>Celkem</b>	<b>11952</b>	<b>3597</b>

Tabulka 8.9b: Počty výskytů jednotlivých frikativ v korpusu Minidialogy-H podle pozice ve slabice (v prětuře či v kodě).

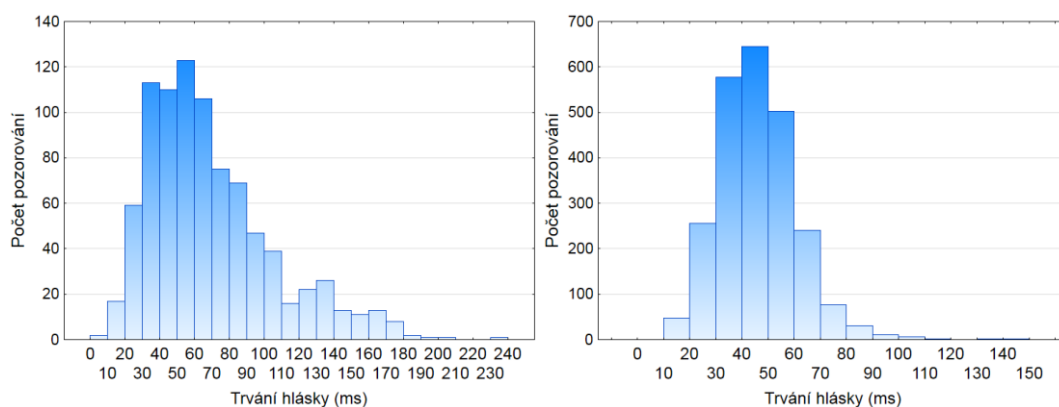
Stejně, jako je tomu v případě exploziv, se většina frikativ vyskytuje v prěturách, což je způsobeno větší preferencí otevřených slabik v češtině. V našem korpusu se v kodě nejčastěji objevuje [š], díky četnému zastoupení sloves s koncovkou druhé osoby jednotného čísla (*myslíš, zeptáš, zjistíš, dostaneš*, atd.). Naopak [h] se v kodě vyskytuje pouze třikrát, což jsou tři případy asimilace ve slově *bůhvi*, kde jinak mluvčí spíše asimilují na [ɣ] (viz výše), či případně neasimilují vůbec.



Obrázek 8.40: Průměrné neseskupené trvání znělých a neznělých frikativ. Svorky označují 95% interval spolehlivosti.

Průměrná trvání frikativ z neseskupených dat znázorňuje obrázek 8.40. Stejně jako u exploziv (obr. 8.24) je zřetelné kratší trvání znělých vůči jejich neznělým protějškům. Tento rozdíl bude statisticky ověřen níže.

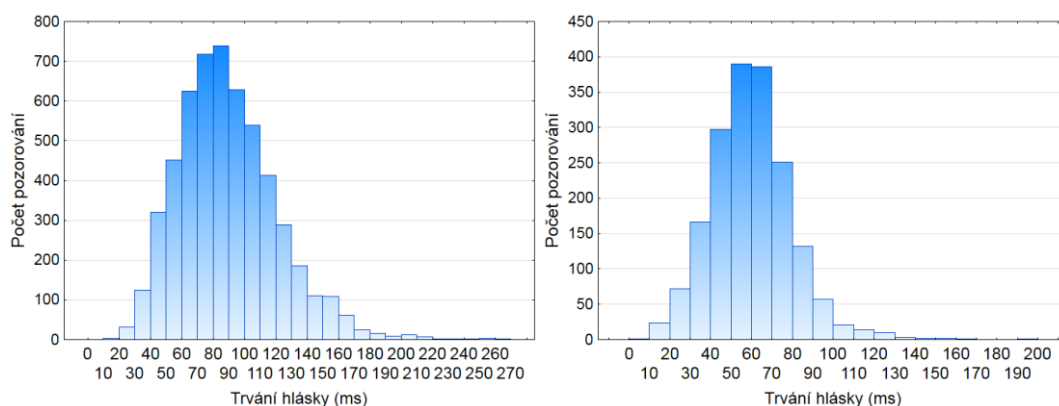




Obrázky 8.41a-b: Histogramy trvání labiodentálních frikativ [f] (vlevo) a [v] (vpravo) ve všech pozicích a všech hláskových okolích.

Neznělá labiodentální frikativa [f], jejíž histogram tvoří obrázek 8.41a, má v našich datech průměrné trvání 69,3 ms. Směr. odchylka je 36,2 ms, medián 61,5 ms, minimum 7,7 ms, maximum 233,2 ms, 10. percentil 31,1 ms a 90. percentil 123,4 ms. Změříme-li trvání pouze [f] v préturách, dostaneme se k průměru 58,7 ms se směr. odchylkou 25,3 ms a mediánem 56,1 ms. [f] v kodách je delší, průměrně 98,7 ms, směr. odchylka 44,7 ms a medián 95,4 ms. T-test na seskupených datech (podle mluvčího, hlásky a pozice hlásky ve slabice) ukazuje na statisticky významný rozdíl mezi těmito dvěma skupinami:  $t(66) = 15,8; p < 0,001$ .

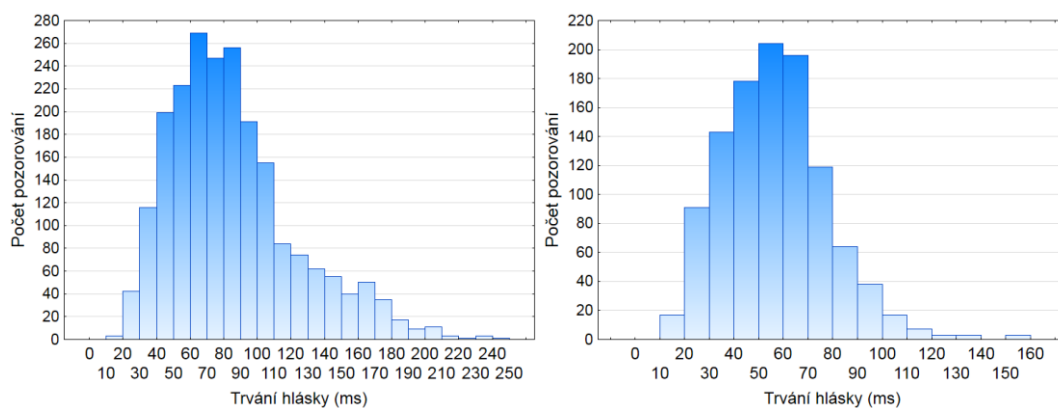
Znělé [v], zobrazené v histogramu 8.41b, trvá průměrně 46 ms se směrodatnou odchylkou 14,8 ms a mediánem 44,7 ms. Minimální změřené trvání je 10,9 ms, maximální 149 ms. 10. percentil je 28,3 ms a 90. percentil 64,1 ms. [v] v préturách má průměrnou délku 45,7 ms, se směr. odchylkou 14,8 ms a mediánem 44,3 ms. V kodách je pak [v] opět nepatrně delší, průměr 50,4 ms, směr. odchylka 13,3 ms a medián 50 ms. Tento rozdíl ovšem nelze statisticky otestovat vzhledem k malému počtu výskytů [v] v kodě.



Obrázky 8.42a-b: Histogramy trvání alveolárních frikativ [s] (vlevo) a [z] (vpravo) ve všech pozicích a všech hláskových okolích.

Neznělá alveolární frikativa [s] má průměrné trvání 89,4 ms se směr. odchylkou 31,5 ms a mediánem 86,2 ms. Minimum je 17,4 ms, maximum 263,1 ms, 10. percentil 52 ms a 90. percentil 130 ms. Rozdělení hodnot jejího trvání zobrazuje histogram 8.29a. [s] v přetúře trvá průměrně 86,5 ms, směr. odchylka 28,2 ms a medián 84,7 ms. V kodě je pak opět o něco delší, průměrně 103,5 ms, směr. odchylka 41,6 ms a medián 96,2 ms. Po seskupení dat podle mluvčího, hlásky a pozice ve slabice vychází tento rozdíl významný:  $t(66) = 9,3; p < 0,001$ .

Její znělý protějšek, [z], trvá průměrně 60 ms, směr. odchylka je 19,9 ms, medián 58,5 ms, minimum 9,2 ms, maximum 191,1 ms, 10. percentil 36,9 ms a 90. percentil 84,2 ms. Rozdělení hodnot viz obr. 8.29b. [z] v přetúře trvá průměrně téměř stejně dlouho, 59,7 ms, sm. odch. 19,6 ms a medián 58,3 ms. [z] v kodě trvá déle, průměr 68,9 ms, sm. odch. 26,2 ms a medián 64,7 ms. T-test opět není vhodné použít kvůli příliš malému počtu [z] v kodě.

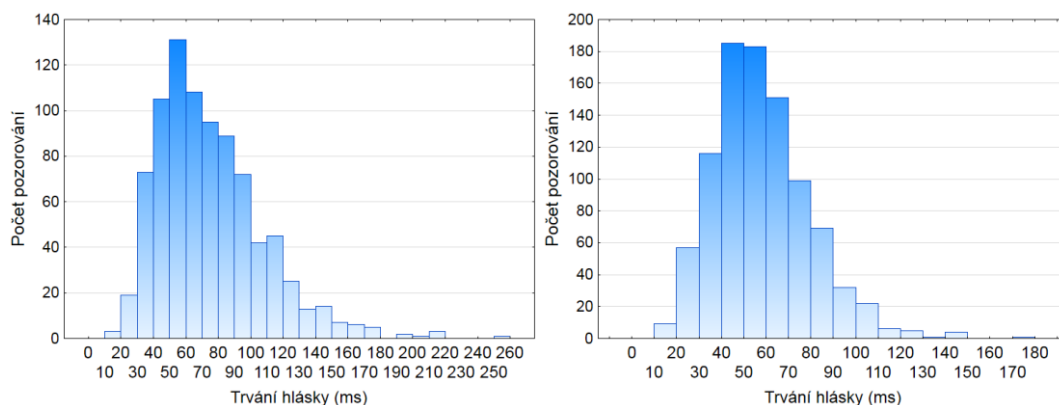


Obrázky 8.43a-b: Histogramy trvání postalveolárních frikativ [š] (vlevo) a [ž] (vpravo) ve všech pozicích a všech hláskových okolicích.

Obr. 8.43a a b znázorňují rozložení hodnot trvání postalveolárních frikativ [š] a [ž]. Průměrně trvá [š] 85,4 ms se směr. odchylkou 38,2 ms a mediánem 79,1 ms. Minimální změřené trvání je 14 ms a maximální 246,6 ms. 10. percentil je 43,2 ms a 90. percentil 141,4 ms. [š] v přetúře trvá průměrně 89,9 ms se směr. odchylkou 30,9 ms a mediánem 88 ms. [š] v kodě je kratší, průměrně 82,7 ms, směr. odchylka 41,7 ms a medián 72,8 ms. Tento rozdíl je statisticky významný – t-test na seskupených datech (podle mluvčího, hlásky a pozice ve slabice):  $t(66) = 4,5; p < 0,001$ .

Znělá frikativa [ž] má průměrné trvání 56,5 ms. Směr. odchylka je 21 ms a medián 55,6 ms. Minimum 11,7 ms, maximum 151,9 ms, 10. percentil 30 ms a 90. percentil 82,6 ms. Vezmeme-li v úvahu pouze [ž] v přetúrách, dostaneme se k průměrnému trvání 60,1 ms, směr. odchylce 20,9 ms a mediánu 58,8 ms. [ž] v kodách trvá průměrně 50,6 ms, se směr. odchylkou 20 ms a mediánem 48,2 ms. Stejně jako u [š] je tento rozdíl významný, což bylo

ověřeno t-testem na seskupených datech (podle mluvčího, hlásky a pozice ve slabice):  $t(66) = 6,6$ ;  $p < 0,001$ . [ž] v kodách je tedy průměrně kratší.



Obrázky 8.44a-b: Histogramy trvání velární frikativy [x] (vlevo) a laryngální frikativy [h] (vpravo) ve všech pozicích a všech hláskových okolích.

Neznělá velární frikativa [x], jejíž trvání je znázorněno histogramem 8.44a, má průměrné trvání 75 ms, směr. odchylku 32,8 ms a medián 69 ms. Minimální trvání je 11,8 ms, maximální 253,4 ms, 10. percentil 39,3 ms a 90. percentil 117,1 ms. [x] v préture trvá průměrně 75,3 ms, směr. odchylka 29,2 ms a medián 73,3 ms. [x] v kodě má průměrné trvání 74,7 ms, se směr. odchylkou 35,3 ms a mediánem 64,4 ms. Tento rozdíl není významný na seskupených ani na neskupených datech.

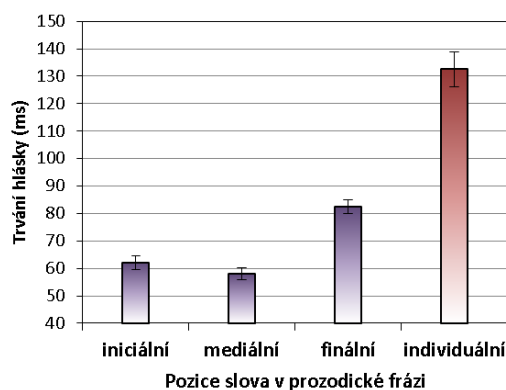
Znělá laryngální frikativa [h], zobrazená v obr. 8.44b, má průměrné trvání 58,3 ms, směr. odchylku 21,5 ms, medián 55,7 ms, minimum 11,3 ms, maximum 170,4 ms, 10. percentil 33,6 ms a 90. percentil 87,3 ms. [h] v kodě jsou pouze tři, nemá tedy smysl je analyzovat zvlášť.

Celkově pro všechny frikativy najednou (3 případy [h] v kodě byly vyřazeny) je efekt pozice ve slabice významný, což jsme ověřili jednofaktorovým testem ANOVA na datech seskupených podle mluvčího, hlásky a pozice ve slabice:  $F(1, 508) = 32,2$ ;  $p < 0,001$ . Výsledek je vysoce významný – v kodách jsou frikativy průměrně delší než v préturech – i přes to, že rozdíly u [š] a [ž] jsou opačné.

Stejně jako u exploziv i zde jsme otestovali efekt znělosti. Rozdíl mezi znělými a neznělými je vysoce významný – jednofaktorová ANOVA s faktorem ZNĚLOST na stejně seskupených datech:  $F(1, 508) = 444,6$ ;  $p < 0,001$ . Jednofaktorová ANOVA s faktorem HLÁSKA a post-hoc test pak ukazují, že k tomuto rozdílu přispívají všechny čtyři dvojice frikativ,  $p < 0,001$ . V rámci neznělých se pak od sebe trváním významně neliší pouze dvojice [f] a [x], v rámci znělých se od sebe neliší pouze [ž] a [h].

### 8.3.3.1 Vliv frázového zpomalování na trvání frikativ

Frikativy jsme seskupili podle hlásky, mluvčího a pozice slova v prozodické frázi, u finálních slov jsme vyřadili ta, která nestojí před plným prozodickým předělem (index 4), dále byly vyřazeny případy, které vznikly seskupením 5 a méně hodnot. Obr. 8.45 ukazuje průměrná trvání konsonantů v iniciálních, mediálních, finálních a individuálních slovech v proz. frázi.



Obrázek 8.45: Průměrné trvání frikativ ze seskupených dat rozdělených podle pozice slova v prozodické frázi. Svorky označují 95% interval spolehlivosti.

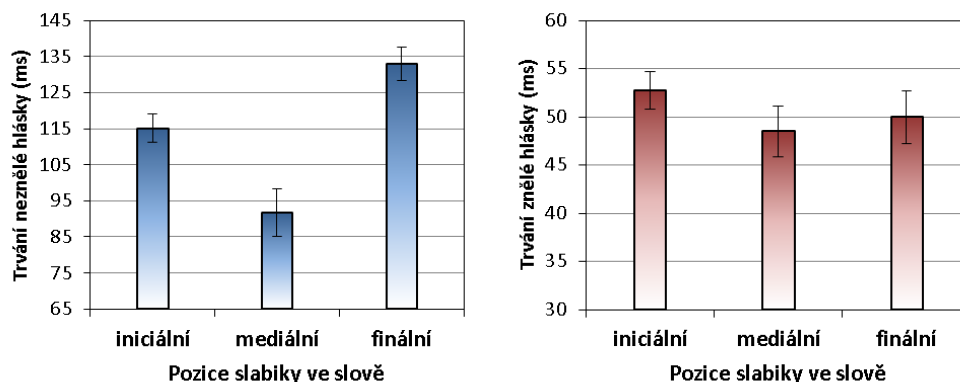
Efekt pozice slova je vysoce významný, což potvrzuje jednofaktorová ANOVA:  $F(3, 719) = 199,5; p < 0,001$ . Tukeyho post-hoc test zároveň ukazuje, že rozdíly mezi průměry iniciálních a mediálních slov významné nejsou, ostatní ano na úrovni  $p < 0,001$ . Velmi dlouhé trvání frikativ v individuálních slovech je pravděpodobně artefaktem toho, že velkou většinu z této skupiny tvoří [s], které je průměrně nejdelší frikativou.

I zde můžeme pro frikativy spočítat koeficienty závěrového zpomalování, což je obsahem tabulky 8.10. Chybí v ní pouze [ž] a [x], které se po seskupení a vyřazení některých dat (viz výše) nevyskytují ve finálních slovech. Nejvíce se prodlužuje [š], až na 1,7násobek svého trvání v mediálních slovech. Naopak nejméně je prodlouženo [v]. Obecně lze říci, že neznělé frikativy se prodlužují více než znělé.

Frikativa	Multipl. koef.	Adit. koef. (ms)
f	1,62	34,6
v	1,18	7,4
s	1,33	25,6
z	1,26	13,4
š	1,70	51,2
h	1,27	13,4
<b>Průměr</b>	<b>1,39</b>	<b>24,3</b>

Tabulka 8.10: Koeficienty průměrných rozdílů mezi frikativami z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

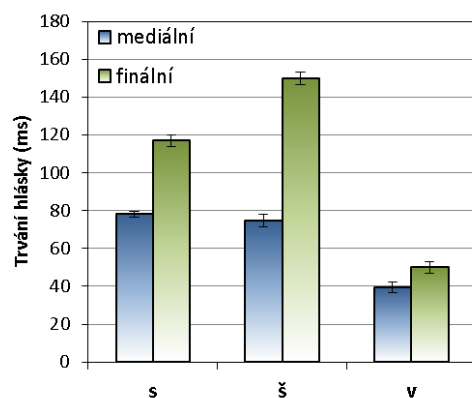
Pro porovnání efektu pozice slabiky ve finálním slově prozodické fráze byla data opět seskupena podle mluvčího, hlásky, pozice hlásky ve slabice, pozice slabiky ve slově a pozice slova v prozodické frázi. Individuální slova byla zahrnuta mezi finální. Průměry jednotlivých skupin zobrazují grafy 8.46a a b.



Obrázky 8.46a-b: Průměrné trvání frikativ ze seskupených dat ve finálních slovech podle pozice slabiky ve slově. Neznělé frikativy jsou zobrazeny v grafu vlevo a znělé vpravo. Svorcky označují 95% interval spolehlivosti. (Měřítka grafů nejsou stejná.)

Co se týče neznělých hlásek (graf vlevo), vidíme zde podobný trend jako u exploziv s tím rozdílem, že všechny tři dvojice se od sebe významně liší:  $F(2, 193) = 53,2; p < 0,001$ ; Tukeyho post-hoc test:  $p < 0,001$ . Hlásky ve finálních slabikách finálních slov jsou nejdelší, v mediálních pak nejkratší. U znělých hlásek vidíme výrazně menší rozdíly, což může být způsobeno tím, že jsou celkově kratší. Efekt pozice slabiky je významný,  $F(2, 135) = 3,7; p = 0,03$ , Tukeyho post-hoc testy ukazují, že významně se od sebe liší pouze mediální slabiky od iniciálních. Co se finálních slabik týče, ve skupině je zastoupeno pouze [v] v přetvářkách, tj. nejkratší frikativa, což způsobuje zmizení efektu frázového zpomalování. Individuální slabiky (tj. jednoslabičná slova) se v protříděném materiálu nevyskytovaly.

Porovnání mediálních exploziv a exploziv z finálních slabik finálních slov jsme provedli pouze pro hlásky [s], [š] a [v], pro které byl dostatek případů. Efekt pozice je vysoce významný:  $F(1, 322) = 1198; p < 0,001$ . Interakce HLÁSKA\*POZICE je také významná:  $F(2, 322) = 211,1; p < 0,001$ , Tukeyho post-hoc testy pro všechny tři hlásky:  $p < 0,001$ . Průměry skupin jsou zobrazené v grafu 8.47.

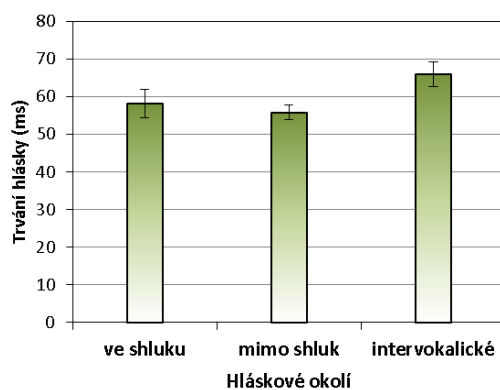


Obrázek 8.47: Průměrné trvání vybraných frikativ v mediálních slovech (modrá) a ve finálních slabikách finálních slov (zelená). Svorky označují 95% interval spolehlivosti.

Je zřetelné, že i v tomto případě zůstává trend viditelný v tabulce 8.10, kdy [v] se prodlužuje nejméně a [š] nejvíce.

### 8.3.3.2 Vliv hláskového okolí na trvání frikativ

V této analýze byly frikativy seskupeny podle mluvčího, hlásky a hláskového okolí. V úvahu byla vzata pouze mediální slova, vyřazeny byly položky seskupené z pěti a méně případů. Hláskové okolí bylo stejně jako u exploziv vyjádřeno třemi kategoriemi – intervokálníké frikativy, frikativy v konsonantickém shluku v rámci slabiky a ostatní (mimo shluk).



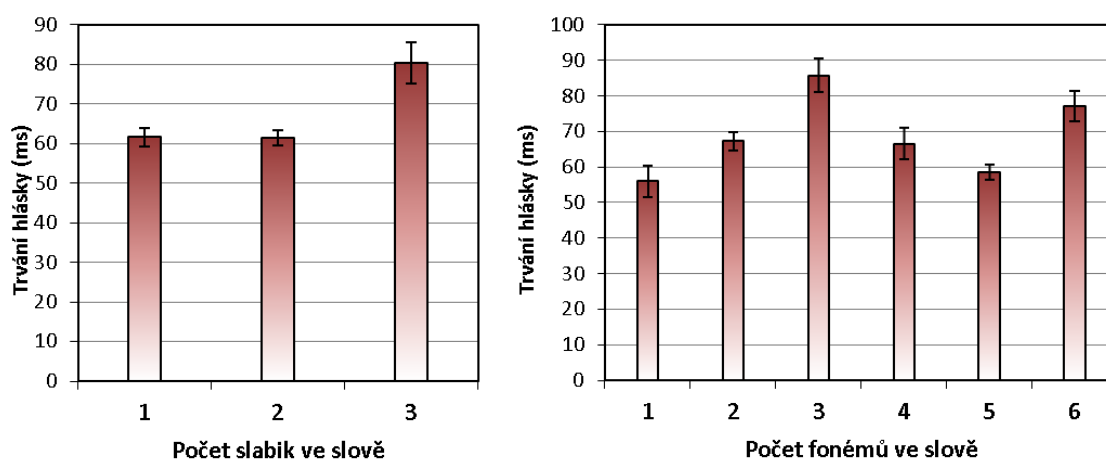
Obrázek 8.48: Průměrné trvání frikativ v mediálních slovech podle hláskového okolí. Svorky označují 95% interval spolehlivosti.

Jednofaktorová ANOVA potvrzuje stejný trend jako u exploziv – efekt hláskového okolí je významný:  $F(2, 388) = 13,9$ ;  $p < 0,001$ , přičemž intervokálníké frikativy jsou významně delší než frikativy ve shluku ( $p = 0,006$ ) a mimo shluk ( $p < 0,001$ ). Druhé dvě skupiny se od sebe významně neliší.

Rozdělíme-li frikativy na znělé a neznělé, zjistíme, že rozdíl mezi intervokalickými a neintervokalickými je výrazně větší u neznělých (přibližně 27 ms), zatímco znělé se v neintervokalických pozicích zkracují výrazně méně, jen o jednotky milisekund.

### 8.3.3.3 Vliv délky vyšší jednotky na trvání frikativ

Zkoumán byl vliv také délky slova a délky prozodické fráze – data byla tedy opět seskupena, vždy podle zkoumaného faktoru. Vyřazeny byly položky vzniklé seskupením pěti a méně případů, v úvahu byla vzata pouze mediální slova.



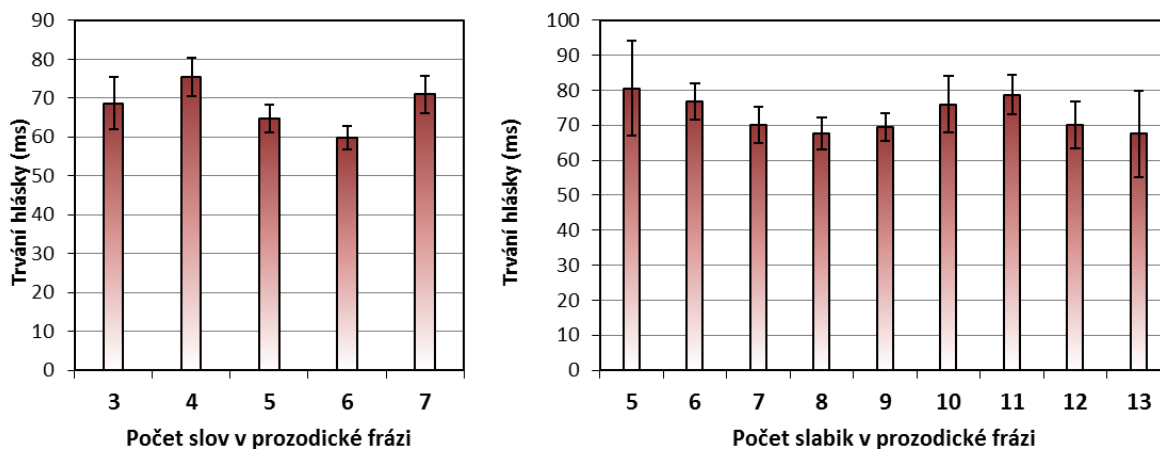
Obrázky 8.49a-b: Průměrné trvání frikativ v mediálních slovech podle délky slova vyjádřené ve slabikách (vlevo) a ve fonémech (vpravo). Svorčky označují 95% interval spolehlivosti.

Efekt délky slova vyjádřené ve slabikách je významný,  $F(2, 367) = 23,3$ ;  $p < 0,001$  a průměry skupin zobrazuje graf 8.49a. Podle Tukeyho post-hoc testu jsou signifikantně delší frikativy ve tříslabičných slovech, ty v jedno- a dvouslabičných se vzájemně neliší.

Obrázek vpravo, 8.49b, ukazuje trvání frikativ v závislosti na délce slova vyjádřené ve fonémech. Efekt je opět významný:  $F(5, 387) = 33,1$ ;  $p < 0,001$  a není lineární. Tukeyho post-hoc test rozděluje průměry do tří skupin: slova s jedním a pěti fonémy, slova se dvěma a čtyřmi fonémy a slova se třemi a šesti fonémy se od sebe signifikantně neliší. Ostatní rozdíly významné jsou ( $p < 0,02$ ). Rozpor mezi oběma obrázky je pouze zdánlivý – za prvé, data jsou jinak seskupena, za druhé, ve skupině tříslabičných slov z obrázku 8.48a jsou i více než šestifonémová slova, která na obrázku 8.48b již analyzována nebyla z důvodu nedostatku dokladů pro jednotlivé skupiny. Nejdelší třífonémová slova jsou jedno- i dvouslabičná, takže skupiny se překrývají jen částečně.

Mezi trváním frikativy a trváním slova v ms (z neseskupených dat) opět není téměř žádná korelace (Pearsonovo  $r = 0,1$ ;  $p < 0,001$ ).

Trvání frikativ v závislosti na délce prozodické fráze vyjádřené ve slovech (obr. 8.50a) i ve slabikách (obr. 8.50b) je opět signifikantně různé a nelineární.



Obrázky 8.50a-b: Průměrné trvání frikativ v mediálních slovech podle délky prozodické fráze vyjádřené ve slovech (vlevo) a ve slabikách (vpravo). Svorky označují 95% interval spolehlivosti.

V obou případech je efekt délky prozodické fráze významný, ve slovech:  $F(4, 270) = 8,9; p < 0,001$ , ve slabikách:  $F(8, 232) = 2,2; p = 0,03$ . Nejdelší jsou frikativy ve čtyř- a sedmislovných prozodických frázích, případně prozodických frázích, které jsou tvořeny pěti, šesti, deseti nebo jedenácti slabikami.

Protože ovšem analyzujeme data od všech frikativ najednou, je možné, že zde hraje roli rozdělení jednotlivých frikativ v daných slovech či prozodických frázích. Náš korpus není tak textově bohatý, abychom tento vliv mohli vyloučit. Výsledky by si tak zasloužily podrobnější prozkoumání, které však už nemůže být součástí této práce.

Korelace mezi trváním frikativ z neskupených dat a trváním prozodických frází v milisekundách je téměř nulová ( $r = 0,03$ ) a statisticky významná ( $p = 0,013$ ).

### 8.3.4 Afrikáty

V analyzované části korpusu Minidialogy-H se vyskytuje 2392 realizací afrikát. Většinu tvoří neznělé [c] a [č], znělých variant [d̥z] a [d̥ʒ] je nemnoho a objevují se pouze na místech asimilace znělosti. Při analýzách byly vyřazeny realizace neznělých afrikát po pauze, kde je délka počátečního závěru stanovená anotátorem.

Neznělé	Poč. výskytů	Znělé	Poč. výskytů
c	1427	dz	99
č	843	dž	23

Tabulka 8.11a: Počty výskytů jednotlivých afrikát v korpusu Minidialogy-H.

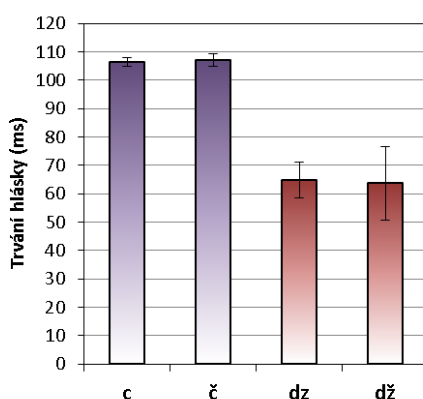


V tabulce 8.11b jsou pak celkové počty rozdělené mezi prétury a kody slabik. Znělé afrikáty se vyskytují pouze v kodách, neznělé častěji v préturách.

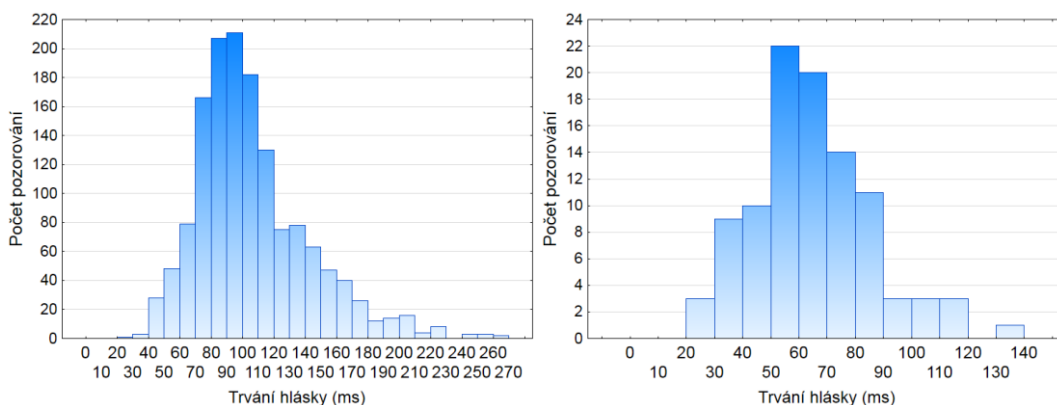
Afrikáta	Prétura	Koda
c	1133	294
č	544	299
dz		99
dž		23
<b>Celkem</b>	<b>1677</b>	<b>715</b>

Tabulka 8.11b: Počty výskytů jednotlivých afrikát v korpusu Minidialogy-H podle pozice ve slabice (v prétuře či v kodě).

Na obrázku 8.51 níže je zobrazeno průměrné neseskupené trvání všech čtyř afrikát. Opět je zřetelně vidět velký rozdíl mezi znělými a neznělými – tento rozdíl však nemůžeme statisticky ověřit kvůli nestejně velikosti vzorků.



Obrázek 8.51: Průměrné neseskupené trvání znělých a neznělých afrikát. Svorky označují 95% interval spolehlivosti.

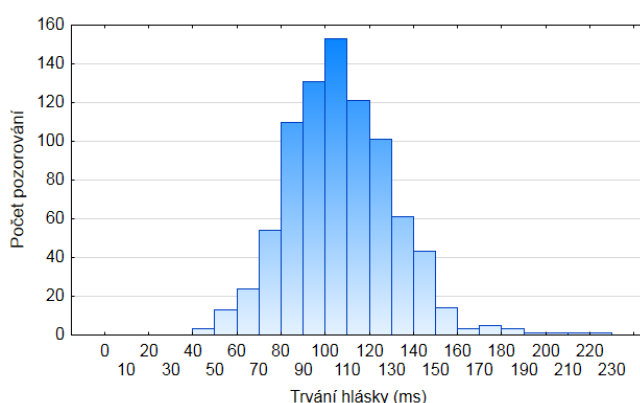


Obrázky 8.52a-b: Histogramy trvání alveolárních afrikát [c] (vlevo) a [dž] (vpravo) ve všech pozicích a všech hláskových okolicích.

Na obrázku 8.52a je znázorněno rozložení hodnot trvání neznělé alveolární afrikáty [c]. Tato hláska má průměrné trvání 106,4 ms se směr. odchylkou 36,4 ms a mediánem 98,4 ms.

Minimální nalezené trvání je 29,4 ms, maximální 263,7 ms, 10. percentil je 67,3 ms a 90. percentil 156,7 ms. [c] v přetúře trvá průměrně 102,3 ms, má směr. odchylku 30,7 a medián 96,5 ms. [c] v kodě je o něco delší, s průměrem 122,2 ms, směr. odchylkou 50 ms a mediánem 107,5 ms. T-test na datech seskupených podle mluvčího a pozice ve slabice ukazuje, že tento rozdíl je statisticky významný:  $t(66) = 7,4; p < 0,001$ .

Znělá alveolární afrikáta [d͡z] se vyskytuje pouze v kodě, má průměrné trvání 64,7 ms, směr. odchylku 21,2 ms, medián 63,2 ms, minimum 27,1 ms, maximum 138,6 ms, 10. percentil 39,1 a 90. percentil 91,7 ms.



Obrázky 8.53a-b: Histogramy trvání postalveolární afrikáty [č] ve všech pozicích a všech hláskových okolích.

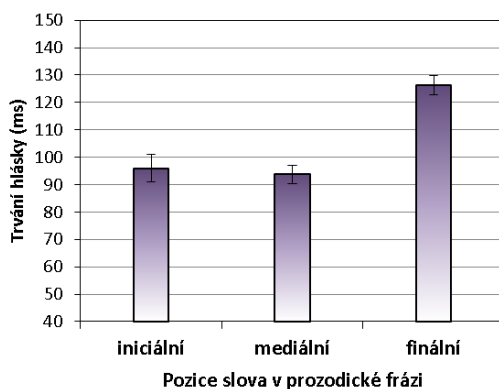
Trvání neznělé postalveolární afrikáty je znázorněno histogramem 8.53. Neznělé [č] trvá průměrně 107,2 ms, se směrodatnou odchylkou 24,1 ms a mediánem 105,1 ms. Minimální trvání bylo naměřeno 43,6, maximální 229,4 ms, 10. percentil 79,1 ms a 90. percentil 137,5 ms. Vezmeme-li v úvahu pouze [č] v přetúře, je jeho průměrné trvání 107,6 ms se směr. odchylkou 23,7 ms a mediánem 107 ms. [č] v kodě trvá průměrně 106,2 ms, směr. odchylka 24,5 ms a medián 103,3 ms. Rozdíl mezi přetúrou a kodou není statisticky významný, a to ani na neseskupených datech (t-test pro nezávislá měření:  $p = 0,42$ ).

Znělé [d͡z] se v našem materiálu vyskytuje pouze ve 23 případech, naměřené hodnoty tedy nejsou zcela srovnatelné s ostatními hláskami a nemá smysl vykreslovat histogram. Nicméně naměřené hodnoty odpovídají našim očekáváním. Průměrné trvání je 63,7 ms, směrodatná odchylka 23,2 ms a medián 52,9 ms.

Efekt pozice ve slabice lze ověřit pouze u neznělých hlásek – jednofaktorová ANOVA na datech seskupených podle mluvčího, hlásky a pozice ve slabice ukazuje na významný rozdíl:  $F(1, 134) = 20,6; p < 0,001$ . Hlásky v kodě jsou průměrně o 10 ms delší, což způsobuje zejména [c], protože jak už bylo zmíněno, rozdíl mezi [č] v přetúře a v kodě významný není.

### 8.3.4.1 Vliv frázového zpomalování na trvání afrikát

Vzhledem k nízkému počtu případů a omezenému výskytu znělých afrikát budeme v této podkapitole porovnávat pouze neznělé. Nejprve jsme seskupili případy podle mluvčího, hlásky a pozice slova v prozodické frázi. Vyřazena byla finální slova před předělem hloubky 3 a vyřadili jsme také položky, které vznikly seskupením pěti a méně případů. Jednofaktorová ANOVA ukazuje na signifikantní efekt pozice slova:  $F(2, 165) = 103,4$ ;  $p < 0,001$ , z Tukeyho post-hoc testu je vidět, že rozdíl mezi iniciálními a mediálními slovy významný není, zato finální jsou přibližně o 30 ms delší než obě skupiny ( $p < 0,001$ ).



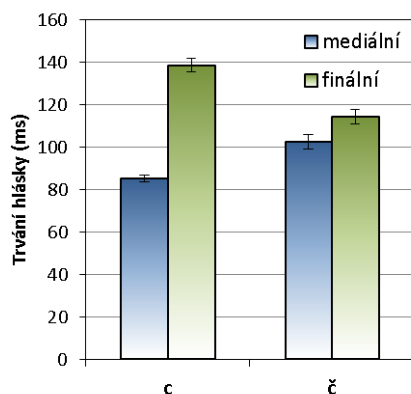
Obrázek 8.54: Průměrné trvání neznělých afrikát ze seskupených dat rozdělených podle pozice slova v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Koeficienty závěrového zpomalování pro obě afrikáty ukazuje tabulka 8.12 níže.

Frikativa	Multipl. koef.	Adit. koef. (ms)
c	1,63	53,3
č	1,12	11,8
Průměr	1,37	32,5

Tabulka 8.12: Koeficienty průměrných rozdílů mezi afrikátami z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

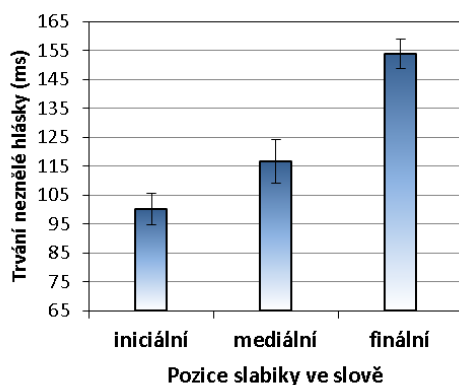
Z tabulky 8.12 vidíme výrazný rozdíl mezi oběma afrikátami, [c] se prodlužuje více (na 1,6násobek svého trvání) než [č], které je prodlouženo jen na 1,1násobek svého trvání. Průměry skupin jsou zobrazeny v grafu 8.55.



Obrázek 8.55: Průměrné trvání neznělých afrikát ze seskupených dat v mediálních a finálních slovech v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Doufaktorová ANOVA potvrzuje statisticky signifikantní efekt pozice:  $F(1, 132) = 311,2; p < 0,001$  a interakci HLÁSKA\*POZICE:  $F(1, 132) = 126,5; p < 0,001$ ; Tukeyho post-hoc test potvrzuje, že všech šest rozdílů je vysoce signifikantních ( $p < 0,001$ ).

Následně analyzujeme opět výhradně finální slova v prozodických frázích zakončených plným předělem. Data byla seskupena podle mluvčího, hlásky a pozice slabiky ve slově. Individuální slabiky (zejména z jednoslabičného slova „co“) byly počítány mezi finální. Vzhledem k nedostatku případů jsme tentokrát nevyřazovali žádné položky. Naměřené průměry pro obě afrikáty zobrazuje graf 8.56.



Obrázek 8.56: Průměrné trvání neznělých afrikát ze seskupených dat ve finálních slovech podle pozice slabiky ve slově. Svorky označují 95% interval spolehlivosti.

Efekt pozice slabiky ve slově je statisticky signifikantní:  $F(2, 174) = 107,5; p < 0,001$ , přičemž podle Tukeyho post-hoc testu jsou významné všechny tři rozdíly,  $p \leq 0,001$ .

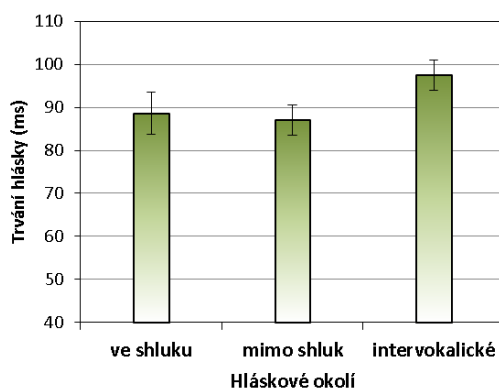
K porovnávání trvání afrikát z mediálních slov s finálními afrikátami ve finálních slovech používáme data seskupená podle mluvčího, hlásky, pozice slabiky ve slově a pozice slova

v prozodické frázi. Opět bereme u finálních slov v úvahu jen ta před předělem hloubky 4 a vyřazujeme položky, které vznikly seskupením 5 a méně případů.

Efekt pozice je významný, afrikáty z mediálních slov jsou kratší než finální afrikáty z finálních slov:  $F(1, 66) = 248,9$ ;  $p < 0,001$ . Mediální afrikáty trvají průměrně 84 ms, zatímco finální 129,6 ms (což je 1,54krát a o 45,6 ms více). Koeficienty pro každou hlásku zvlášť ovšem nelze spočítat, protože [č] se po protřídění nevyskytuje v mediálních slovech.

### 8.3.4.2 Vliv hláskového okolí na trvání afrikát

Vliv hláskového okolí jsme opět z důvodu počtu dokladů zkoumali jen na neznělých afrikátách. Položky z mediálních slov byly seskupeny podle mluvčího, hlásky a hláskového okolí. Položky vzniklé seskupením pěti a méně případů jsme tentokrát nevyřazovali, protože bychom se tím zcela připravili o [c] ve shluku ([č] se ve shluku nevyskytuje vůbec). Průměry skupin jsou znázorněny na obrázku 8.57.

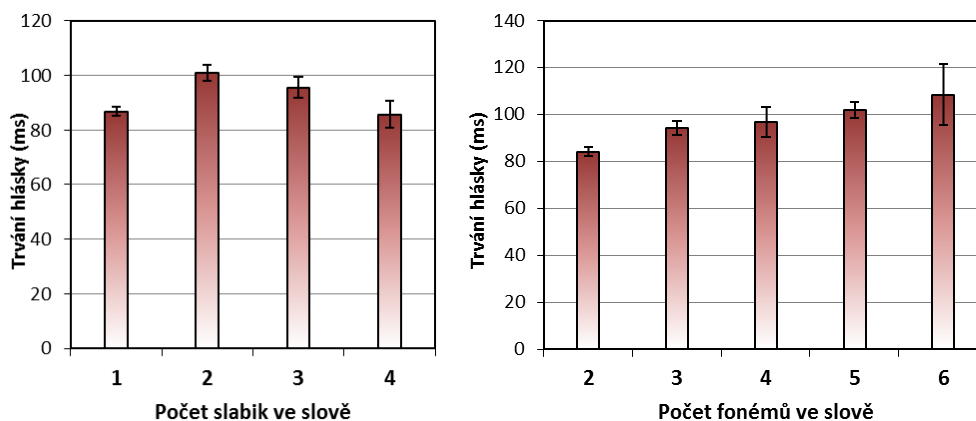


Obrázek 8.57: Průměrné trvání afrikát v mediálních slovech podle hláskového okolí. Svorky označují 95% interval spolehlivosti.

Jednofaktorová ANOVA ukazuje na významný efekt hláskového okolí:  $F(2, 167) = 9,7$ ;  $p < 0,001$ , přičemž podle Tukeyho post-hoc testu je významný rozdíl mezi afrikátami ve shluku a intervokalickými ( $p = 0,01$ ) a mimo shluk a intervokalickými ( $p < 0,001$ ). Když se podíváme na jednotlivé afrikáty, zjistíme, že se [c] v neintervokalických pozicích zkracuje více než [č].

### 8.3.4.3 Vliv délky vyšší jednotky na trvání afrikát

Pro zkoumání efektu délky slova a prozodické fráze na trvání afrikát data nebyla kvůli malému počtu dokladů seskupována. V úvahu ovšem stále bereme pouze mediální slova. Z důvodu nedostatečného počtu realizací znělých afrikát analyzujeme opět pouze neznělé.

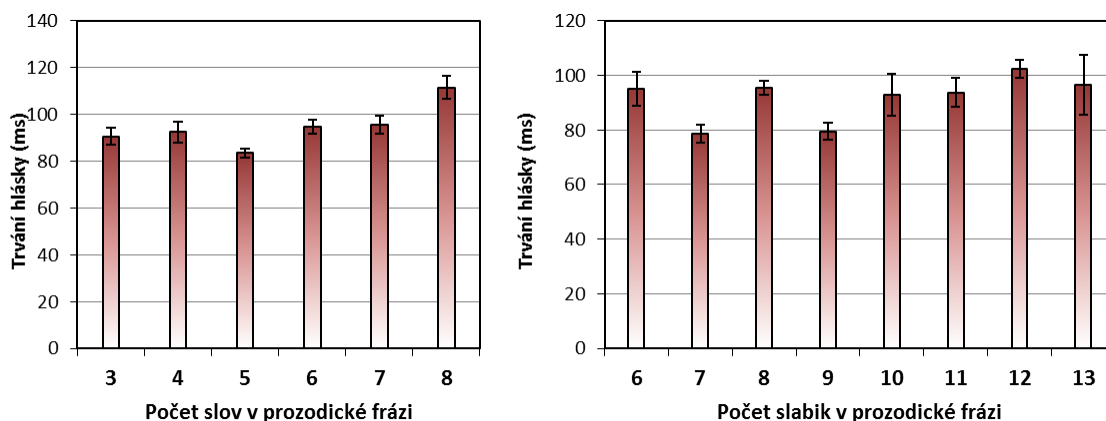


Obrázky 8.58a-b: Průměrné trvání neznělých afrikát v mediálních slovech podle délky slova vyjádřené ve slabikách (vlevo) a ve fonémech (vpravo). Svorky označují 95% interval spolehlivosti.

Na obr. 8.58a a b je zobrazeno průměrné trvání neznělých afrikát podle délky slova ve slabikách a fonémech. Efekt délky slova ve slabikách je významný:  $F(3, 948) = 26,2$ ;  $p < 0,001$  (jednofaktorová ANOVA pro nezávislá měření), přičemž podle Tukeyho post-hoc testu se od sebe signifikantně neliší jedno- a čtyřslabičná slova, která jsou kratší, a dvou- a tříslabičná slova. Ostatní rozdíly významné jsou ( $p < 0,01$ ). Také efekt délky slova ve fonémech je statisticky významný:  $F(4, 788) = 26,6$ ;  $p < 0,001$ . Přispívají k tomu zejména slova o dvou fonémech, kde jsou afrikáty signifikantně kratší než v delších slovech (Tukeyho post-hoc test:  $p < 0,003$ ). Protože nejdelší analyzované slovo má 6 fonémů, lze předpokládat, že tato skupina odpovídá dvou- až tříslabičným slovům z grafu vlevo.

Mezi trváním afrikát a trváním slova v milisekundách je mírná pozitivní korelace (Pearsonovo  $r = 0,28$ ), která je statisticky významná ( $p < 0,001$ ).

Co se týče vlivu trvání prozodické fráze, obr. 8.59a a b ukazují průměry skupin podle délky fráze ve slovech a ve slabikách. Efekt je v obou případech významný,  $F(5, 889) = 27,3$ ;  $p < 0,001$  pro délku fráze ve slovech a  $F(7, 774) = 24,4$ ;  $p < 0,001$ , ale nesnadno interpretovatelný. Pravděpodobně zde hraje roli to, že dat je méně než u ostatních konsonantů, a tedy se zde projevují i další vlivy (identita hlásky, pozice ve slabice, atp.).



Obrázky 8.59a-b: Průměrné trvání afrikát v mediálních slovech podle délky prozodické fráze vyjádřené ve slovech (vlevo) a ve slabikách (vpravo). Svorky označují 95% interval spolehlivosti.

Mezi trváním prozodické fráze v ms a trváním afrikát nalézáme opět jen velmi mírnou až neexistující, ale statisticky významnou korelaci ( $r = 0,16$ ;  $p < 0,001$ ).

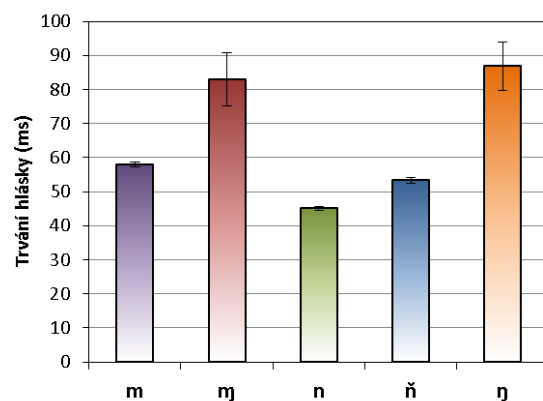
### 8.3.5 Nazály

Počet nazál vyskytujících se v analyzované části korpusu Minidialogy-H je 12 191. V textu se neobjevují slabičné nazály, kromě slova *hm*, které je ale z analýz vyřazeno. Stejně jako u předchozích typů konsonantů se jich většina vyskytuje v prětuře. Alofonické varianty [m] ze slova *trumf* a [ŋ] ze slova *novinka* mají jen velmi malé zastoupení (27, respektive 34 případů).

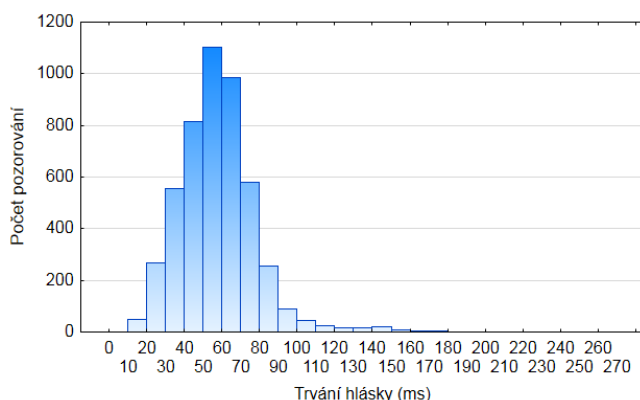
Nazála	Prětura	Koda	Celkem
m	3309	1537	4846
ŋ		27	27
n	4272	508	4780
ň	2403	101	2504
ŋ		34	34
<b>Celkem</b>	<b>9984</b>	<b>2207</b>	<b>12191</b>

Tabulka 8.13: Počty výskytů jednotlivých nazál v korpusu Minidialogy-H, celkové a podle pozice ve slabice (v prětuře či v kodě).

Obrázek 8.60 ukazuje průměry trvání jednotlivých nazál z neseskupených dat. Nejméně zastoupené hlásky jsou průměrně nejdělsí (na což může mít vliv i to, že jsou obě umístěné ve finálních slovech prozodické fráze) a vykazují největší rozptyl.



Obrázek 8.60: Průměrné neseskupené trvání nazál. Svorky označují 95% interval spolehlivosti.

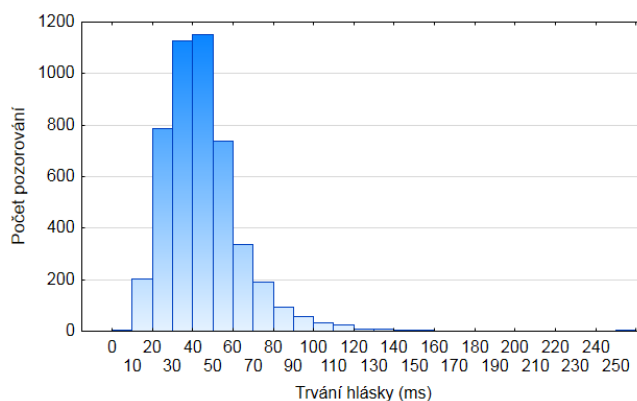


Obrázek 8.61: Histogram trvání bilabiální nazály [m] ve všech pozicích a všech hláskových okolích.

Rozdělení hodnot bilabiálního [m] zobrazuje histogram 8.61. Jeho průměrné trvání je 58 ms, směrodatná odchylka 20,9 ms a medián 56,9 ms. Minimální změřené trvání [m] je 10,7 ms, maximální pak 264 ms. 10. percentil je 33,6 ms a 90. percentil 80,2 ms. [m] v préturách trvá průměrně 56,1 ms se sm. odchylkou 17,9 ms a mediánem 57 ms. [m] v kodách trvá průměrně 62,1 ms, sm. odchylka 25,9 ms a medián 56,9 ms. Rozdíl mezi préturou a kodou je statisticky významný,  $t(66) = 4,6$ ;  $p < 0,001$  (t-test pro nezávislá měření na datech seskupených podle mluvčího, hlásky a pozice ve slabice), [m] v kodách jsou signifikantně delší.

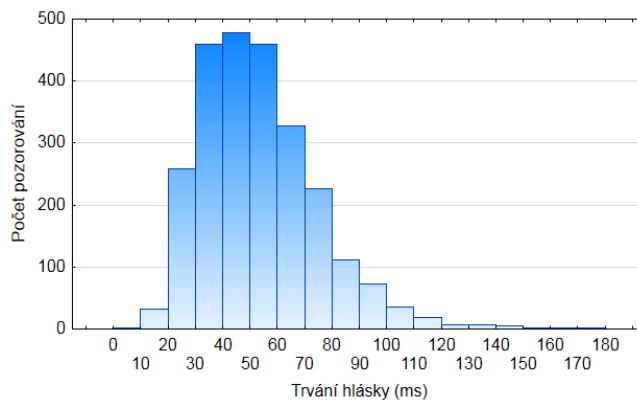
Labiodentální [m̥] se vyskytuje pouze v kodě jediného slova a má průměrné trvání 83,1 ms, se sm. odchylkou 23,8 ms a mediánem 83,5 (tj. delší než [m] v kodě). Vzhledem k malému počtu případů nemá smysl vykreslovat histogram.





Obrázek 8.62: Histogram trvání alveolární nazály [n] ve všech pozicích a všech hláskových okolích.

Na obrázku 8.62 jsou zobrazeny hodnoty trvání alveolárního [n] ve všech pozicích. [n] je nejkratší nazála, trvá průměrně 45,1 ms, se směr. odchylkou 20,4 ms a mediánem 42,1 ms. Minimální nalezené [n] v korpusu trvá 7,1 ms, maximální 327,6 ms. 10. percentil je 24,1 ms a 90. percentil 67,6 ms. Vezmeme-li v úvahu pouze [n] v přetůře, trvá průměrně 43,9 ms, se sm. odchylkou 20,2 ms a mediánem 41,2 ms. [n] v kodě jsou opět o něco delší, průměrně 55,7 ms, sm. odchylka 19,4 ms a medián 53,7 ms. Tento rozdíl je statisticky vysoce významný, t-test na datech seskupených podle mluvčího, hlásky a pozice ve slabice vychází  $t(66) = 9,6$ ;  $p < 0,001$ .



Obrázek 8.63: Histogram trvání palatální nazály [ň] ve všech pozicích a všech hláskových okolích.

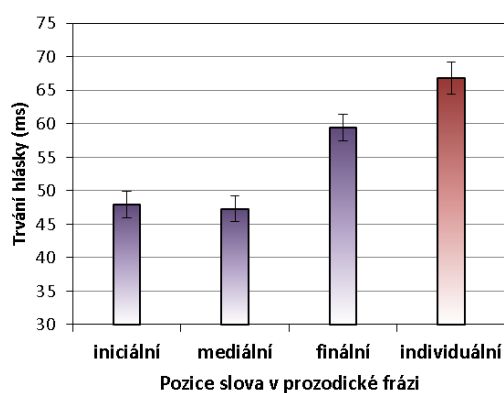
Palatální [ň] je zobrazeno na histogramu 8.63. Průměrně trvá 53,3 ms, směrodatná odchylka je 21,5 ms, medián 50,5 ms, minimum 4 ms a maximum 176 ms. 10. percentil je 29 ms a 90. percentil 80,7 ms. [ň] vyskytující se v přetůřích trvá průměrně 52,9 ms se směr. odchylkou 21,2 ms a mediánem 50,1 ms. [ň] v kodách je podobně jako [m] a [n] delší, průměrně 63,9 ms, směr. odchylka 25,3 ms a medián 59,3 ms. T-test pro ověření významnosti tohoto rozdílu není vhodné provádět vzhledem k výtazné nerovnosti velikostí vzorků.

Podobně jako [ŋ] se i velární alofon [ŋ] vyskytuje jen ve velmi omezeném počtu případů (34) a pouze v kodě. Průměrně trvá 86,9 ms se směrodatnou odchylkou 16,4 ms a mediánem 88,6. Je tedy průměrně nejdelší ze všech nazál, i pokud bereme v úvahu jen ty vyskytující se v kodě. Rozdíl v trvání [ŋ] a [ŋ] není statisticky významný (t-test pro nezávislá měření na neseskupených datech).

Efekt pozice ve slabice ověříme pouze pro [m], [n] a [ŋ], protože [ŋ] a [ŋ] by ovlivňovaly výsledek ve prospěch delší kody. Data byla seskupena podle mluvčího, hlásky a pozice ve slabice a položky vytvořené z pěti a méně případů byly opět vyřazeny. Efekt je významný:  $F(1, 168) = 55,8; p < 0,001$ , nazály v kodě jsou přibližně o 8 ms delší než v přetúře.

### 8.3.5.1 Vliv frázového zpomalování na trvání nazál

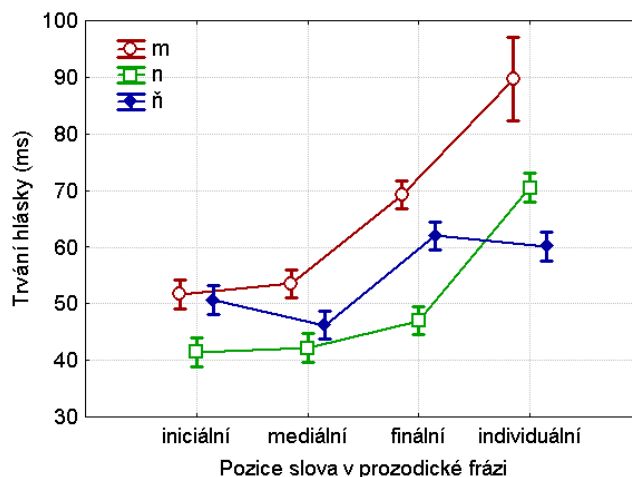
Z této analýzy byly vyřazeny hlásky [ŋ] a [ŋ]. Zbývající data byla seskupena podle hlásky, mluvčího a pozice slova v prozodické frázi, položky vzniklé seskupením méně než šesti případů byly vyřazeny. U finálních slov byla vzata v úvahu jen ta, vyskytující se před předělem hloubky 4. Jako první byl prozkoumán vliv pozice slova v prozodické frázi na trvání hlásky. Průměry jednotlivých skupin ukazuje obrázek 8.64.



Obrázek 8.64: Průměrné trvání nazál ze seskupených dat rozdělených podle pozice slova v prozodické frázi.

Svorky označují 95% interval spolehlivosti.

Faktor POZICE je vysoce významný:  $F(3, 372) = 75,4; p < 0,001$ . Podle Tukeyho post-hoc testu jsou významné všechny rozdíly ( $p < 0,001$ ) kromě rozdílu mezi iniciálními a mediálními. Nazály v individuálních slovech v prozodické frázi jsou nejdelší, následují nazály ve finálních slovech, v iniciálních a mediálních jsou přibližně stejně dlouhé. Také interakce HLÁSKA\*POZICE ve dvoufaktorovém testu ANOVA je signifikantní:  $F(6, 364) = 22,5; p < 0,001$ . Průměry pro jednotlivé hlásky zobrazuje graf 8.65.



Obrázek 8.65: Průměrné trvání jednotlivých nazál ze seskupených dat podle pozice slova v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Jak je vidět z obrázku 8.65, nejvýrazněji má pozice slova v prozodické frázi vliv na [m], u nějž jsou rozdíly mezi mediálními, finálními a individuálními slovy vysoce signifikantní (Tukeyho post-hoc test:  $p < 0,001$ ). U [n] je významný rozdíl pouze mezi individuálními slovy a ostatními skupinami ( $p < 0,001$ ). [ň] se chová ještě trochu jinak – významně se od sebe neliší iniciální a mediální slova, ani finální a individuální, ale jen tyto dvě skupiny mezi sebou ( $p < 0,001$ ).

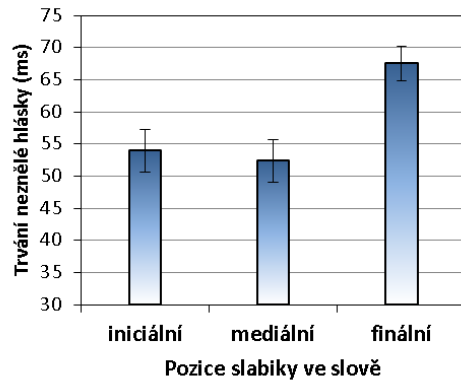
Nazála	Multipl. koef.	Adit. koef. (ms)
m	1,29	15,7
n	1,12	4,9
ň	1,34	15,8
Průměr	1,25	12,1

Tabulka 8.14: Koeficienty průměrných rozdílů mezi nazálami z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

Porovnáním mediálních a finálních slov vznikly koeficienty závěrového zpomalování v tabulce 8.14. Nejvíce se prodlužuje [m] a [ň] (přibližně o třetinu svého trvání), naopak [n] zůstává téměř stejné.

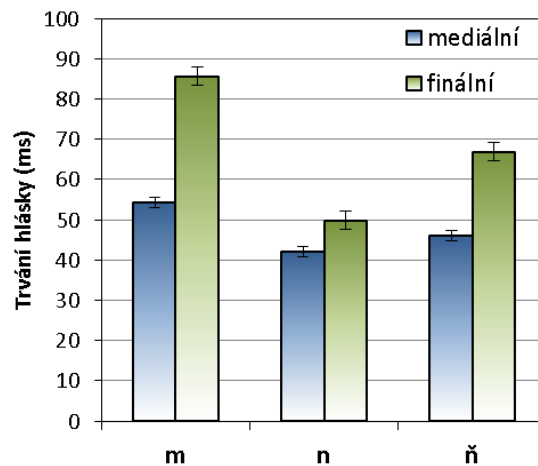
Před zkoumáním efektu pozice slabiky ve finálních slovech jsme seskupili data podle mluvčího, hlásky, pozice slabiky ve slově a pozice slova v prozodické frázi. U finálních slov byla vytříděna jen ta nacházející se před předělem hloubky 4. Opět jsme vyřadili položky seskupené z méně než šesti případů. Individuální slova byla také vyřazena z důvodu malého počtu případů, a protože obsahovala pouze [ň].

Průměrné trvání v jednotlivých skupinách je zobrazeno na obr. 8.66.



Obrázek 8.66: Průměrné trvání nazál ze seskupených dat ve finálních slovech podle pozice slabiky ve slově. Svorky označují 95% interval spolehlivosti.

Jednofaktorová ANOVA ukazuje na významnost faktoru POZICE:  $F(2, 232) = 31,9; p < 0,001$  a podle Tukeyho post-hoc testu není významný jen rozdíl mezi iniciálními a mediálními slabikami, ostatní významné jsou ( $p < 0,001$ ). Opět zde vidíme tendenci k prodlužování finální slabiky ve finálních slovech, přičemž iniciální a mediální jsou přibližně stejně dlouhé.



Obrázek 8.67: Průměrné trvání nazál ze seskupených dat v mediálních slovech (modrá) a finálních slabikách finálních slov (zelená) v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Porovnání nazál z mediálních slov a finálních slabik finálních slov před plným prozodickým předělem (obr. 8.67) ukazuje na stejný trend, jaký jsme viděli v tabulce 8.14 – [m] a [ň] se prodlužují výrazně více než [n], nicméně v tomto případě je rozdíl mezi [m] a [ň] větší. Interakce faktorů HLÁSKA\*POZICE je vysoce signifikantní:  $F(2, 402) = 77,9; p < 0,001$ , Tukeyho post-hoc testy naznačují, že rozdíly mezi mediálními a finálními jsou významné pro všechny tři hlásky (dokonce i pro [n]):  $p < 0,001$ .

Protože v případě nazál je k dispozici dostatečný počet dokladů ve všech skupinách, lze spočítat spolehlivé koeficienty prodlužování pro finální slabiky ve finálních slovech v prozodické frázi, které shrnuje tabulka 8.15.

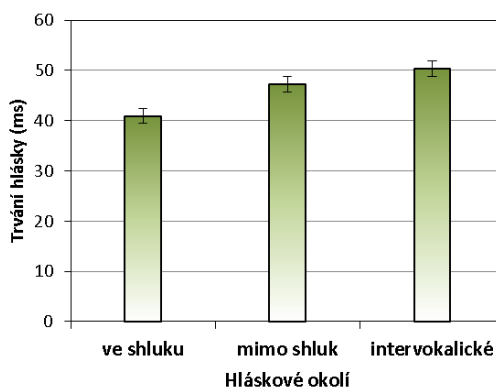
Nazála	Multipl. koef.	Adit. koef. (ms)
m	1,58	31,4
n	1,18	7,7
ň	1,45	20,8
Průměr	1,4	19,9

Tabulka 8.15: Koeficienty průměrných rozdílů mezi nazálami z mediálních slov a finálních slabik z finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

Z tabulky je vidět, že nejvíce se prodlužuje [m], které je ve finálních slabikách finálních slov průměrně 1,58krát delší. Naopak [n] je prodlouženo jen na 1,18násobek svého trvání.

### 8.3.5.2 Vliv hláskového okolí na trvání nazál

Efekt hláskového okolí zkoumáme opět na datech z mediálních slov seskupených podle mluvčího, hlásky a hláskového okolí. [ŋ] ani [ɲ] se v mediálních slovech nevyskytují. Protože byl v tomto případě dostatek dokladů ve všech skupinách, nebylo nutné žádné případy vyřazovat. Průměry jsou zobrazeny v grafu 8.68 – intervokální jsou nazály z obou stran obklopené vokály, ve shluku znamená, že nazála je součástí konsonantického shluku v rámci slabiky. Mimo shluk je skupina, kam spadají všechny ostatní případy.



Obrázek 8.68: Průměrné trvání nazál v mediálních slovech podle hláskového okolí. Svorky označují 95% interval spolehlivosti.

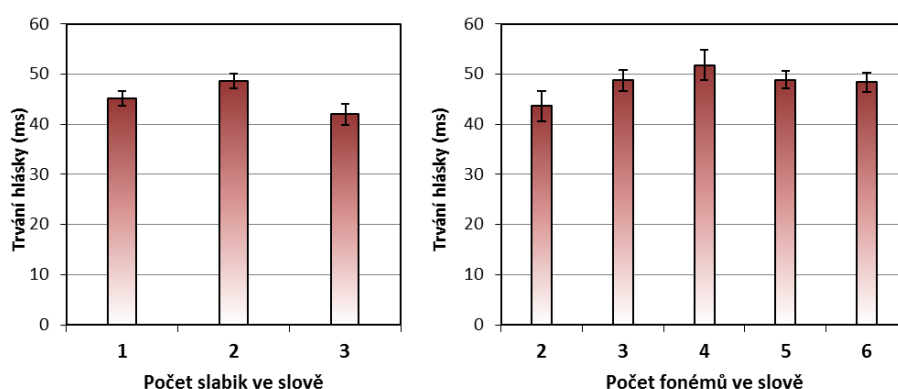
Jednofaktorová ANOVA potvrzuje významnost efektu hláskového okolí:  $F(2, 303) = 36,6$ ;  $p < 0,001$ . Rozdíly mezi všemi třemi skupinami jsou významné (Tukeyho post-hoc test:  $p \leq 0,02$ ), přičemž intervokální nazály jsou nejdelší a nazály ve shlucích nejkratší.

Dvoufaktorová ANOVA pak prozrazuje, že k tomuto trendu přispívají všechny nazály podobnou měrou, pouze u [n] nenalzáme významný rozdíl mezi pozicemi mimo shluk a intervokalickou a u [m] není signifikantní rozdíl mezi oběma neintervokalickými pozicemi.

### 8.3.5.3 Vliv délky vyšší jednotky na trvání nazál

V této analýze byly nazály opět seskupeny, a to podle mluvčího, hlásky a délky slova, případně prozodické fráze. Z důvodu nedostatku dokladů pro [ŋ] a [ɲ] analyzujeme stále pouze nazály [m], [n] a [ň]. I zde se zabýváme pouze mediálními slovy a položky vzniklé seskupením pěti a méně případů vyřazujeme.

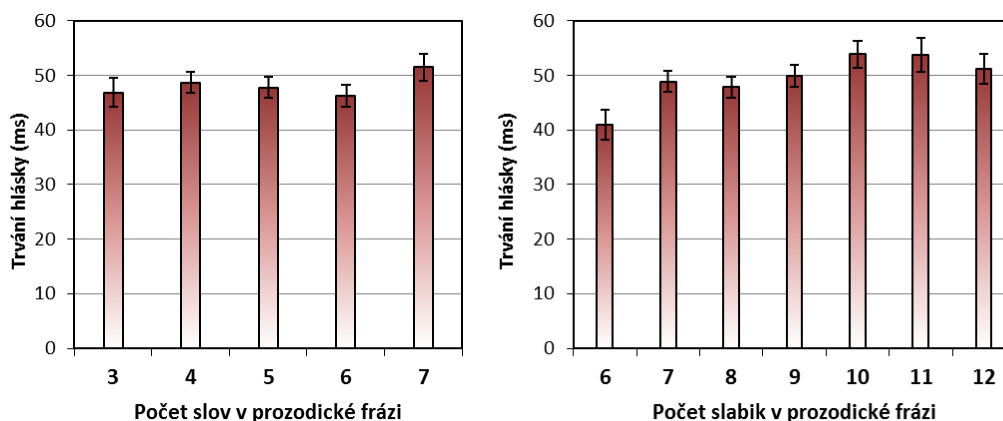
Jak se mění trvání nazál v závislosti na délce slova, ukazují obrázky 8.69a a b.



Obrázky 8.69a-b: Průměrné trvání nazál v mediálních slovech podle délky slova vyjádřené ve slabikách (vlevo) a ve fonémech (vpravo). Svorčky označují 95% interval spolehlivosti.

V obrázku vlevo je zobrazeno průměrné trvání nazál podle počtu slabik ve slově. Efekt délky slova je statisticky významný:  $F(2, 252) = 13; p < 0,001$ . Ve dvouslabičných slovech jsou nazály průměrně nejdelší a v tříslabičných nejkratší. Všechny tři rozdíly mezi skupinami jsou také statisticky významné (Tukeyho post-hoc test:  $p < 0,05$ ). I efekt délky slova vyjádřené ve fonémech je významný:  $F(4, 316) = 3,8; p = 0,005$ . Nazály jsou průměrně nejdelší ve slovech o čtyřech fonémech a pak jejich trvání klesá. Post-hoc test ukazuje na významnost rozdílu mezi dvou- a čtyřfonémovými slovy a dvou- a pětifonémovými slovy ( $p < 0,03$ ).

Mezi trváním nazály a trváním slova v milisekundách (neseskupená data) je mírná pozitivní korelace (Pearsonovo  $r = 0,2$ ), která je statisticky významná ( $p < 0,01$ ).



Obrázky 8.70a-b: Průměrné trvání nazál v mediálních slovech podle délky prozodické fráze vyjádřené ve slovech (vlevo) a ve slabikách (vpravo). Svorky označují 95% interval spolehlivosti.

Průměrné trvání nazál podle délky prozodické fráze ukazují obrázky 8.70a a b. Efekt délky fráze je v obou případech významný – měřeno ve slovech:  $F(4, 325) = 3,1$ ;  $p = 0,02$  a ve slabikách:  $F(6, 328) = 10,6$ ;  $p < 0,001$ . Ačkoliv dlouhé prozodické fráze složené ze sedmi slov obsahují průměrně nejdelší nazály, tento rozdíl je statisticky významný pouze ve srovnání s šestislovnými frázemi, které obsahují kratší nazály (Tukeyho post-hoc test:  $p = 0,01$ ). Zároveň nejkratší prozodické fráze (složené z šesti slabik) obsahují i signifikantně nejkratší nazály ( $p < 0,001$ ). Nejdelší nazály se objevily v prozodických frázích složených z deseti a jedenácti slabik.

Korelace mezi neseskupenými hodnotami trvání nazál a trváním prozodické fráze v milisekundách je velmi slabá ( $r = 0,12$ ) a zároveň statisticky významná ( $p < 0,001$ ).

### 8.3.6 Vibranty a aproximanty

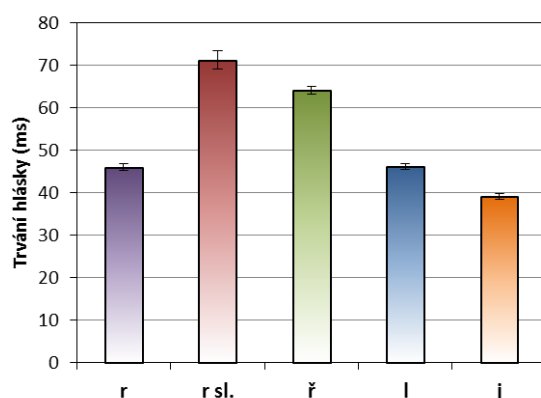
Do poslední skupiny analýz byly zařazeny vibranty [r] a [ř] a aproximanty [l] a [j]. Rozhodli jsme se v této práci nerozlišovat znělé a neznělé realizace postalveolární frikativní vibranty [ř] na úrovni hlásek. [ř] je nejčastěji desonorizovaná frikativa z důvodu neexistence neznělého protějšku s distinktivní platností (viz Machač, 2008 či Skarnitzl, 2011) a hranice mezi znělými a neznělými se nese snadno určuje. Při anotaci korpusu tedy nebyl tento rozdíl brán v úvahu a představuje tematiku pro další výzkum.

Celkem je v této skupině k dispozici 10 294 realizací. Pro slabičné [r] je v našem materiálu celkem 273 dokladů ze 7 různých slov (*krmivo*, *mrknout*, *opatrně*, *opatrnost*, *prskat*, *první*, *zvrtnout*), v jednom případě vzniklo elizí vokálu ve slově *promysleli*. Slabičné [l] ani [ɱ] se v korpusu nevyskytují. Jediný případ slabičného [ř], které vzniklo elizí jádra slabiky ve slově *přišli*, nebyl do analýz zahrnut.

Hláška	Prétura	Koda	Celkem
r nesl.	1833	408	2241
r slab.			273
ř	1561	68	1629
l	2461	613	3074
j	2556	521	3077
<b>Celkem</b>	<b>8411</b>	<b>1610</b>	<b>10294</b>

Tabulka 8.16: Počty výskytů jednotlivých vibrant a aproximant v korpusu Minidialogy-H, celkové a podle pozice ve slabice (v prétuře či v kodě).

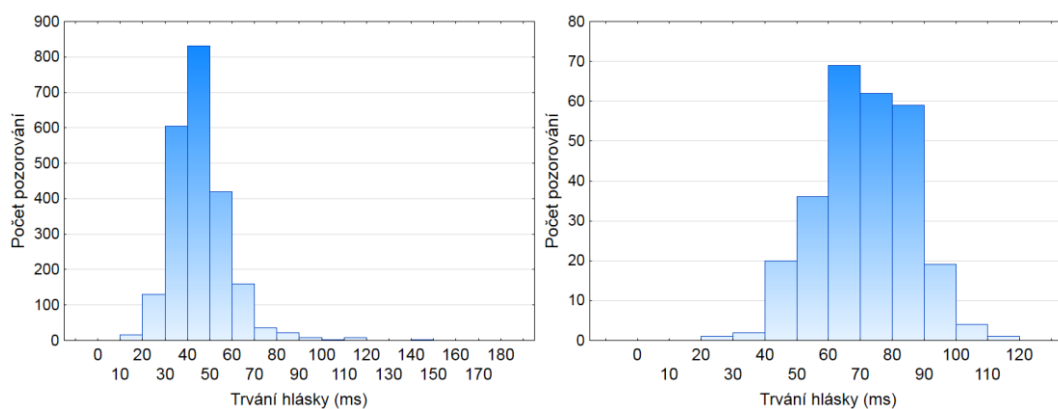
Tabulka 8.16 uvádí počty výskytů jednotlivých vibrant a aproximant. Opět je vidět, že výrazná většina těchto konsonantů (84 %) se vyskytuje v préturách.



Obrázek 8.71: Průměrné nesekupené trvání vibrant a aproximant. Svorcky označují 95% interval spolehlivosti.

Obrázek 8.71 ukazuje průměry trvání jednotlivých vibrant a aproximant z nesekupených dat. Nejdelší je slabičné [r], které svým trváním přesahuje jak znělé explozivny a frikativy, tak i nazály [m], [n] a [ň]. Je dokonce průměrně delší i než krátké vokály.

O něco kratší je pak [ř], následované neslabičným [r] a [l]. Průměrně nejkratší je palatální aproximanta [j], která se tak vedle [d] řadí mezi nejkratší české hlásky.

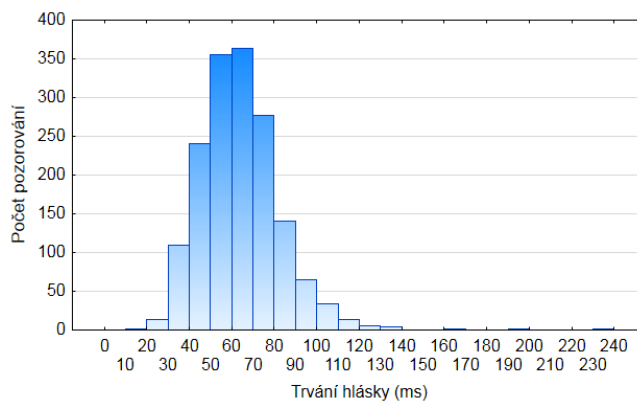


Obrázek 8.72a-b: Histogramy trvání alveolární vibranty [r] (vlevo) a její slabičné varianty [r] (vpravo) ve všech pozicích a všech hláskových okolích.



Na obrázku 8.72a je znázorněno rozložení hodnot trvání neslabičné alveolární vibranty [r]. Její průměrné trvání je 45,9 ms, směrodatná odchylka 13,6 ms a medián 44,5 ms. Minimální naměřené trvání je 15,6 ms, maximální 223 ms, 10. percentil 32 ms a 90. percentil 60,5 ms. [r] v préturách trvá průměrně 45,3 ms se směr. odchylkou 11 ms a mediánem 44,5 ms. [r] v kodách má průměrné trvání 48,7 ms se směr. odchylkou 21,5 ms a mediánem také 44,5 ms. T-test na datech seskupených podle mluvčího a pozice ve slabice ukazuje na statisticky signifikantní rozdíl mezi trváním [r] v prétuře a v kodě:  $t(66) = 3$ ;  $p = 0,004$ , přičemž [r] v kodách jsou o něco delší a mají také výrazně větší rozptyl.

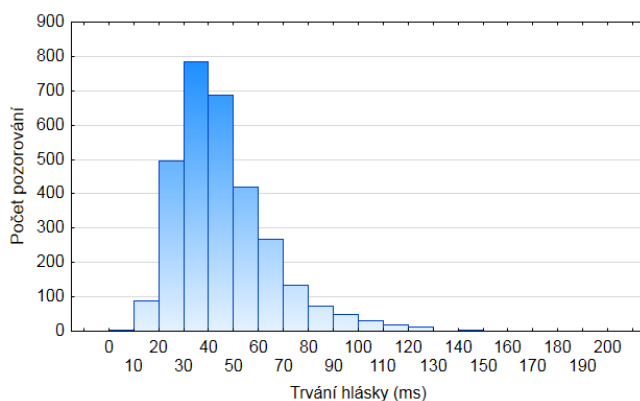
Obrázek 8.72b zobrazuje histogram hodnot trvání slabičného [r̥]. Jeho průměrné trvání je 71,1 ms, směr. odchylka 14,6 ms, medián 71,9 ms, minimální trvání 23,9 ms, maximální 117,8 ms, 10. percentil 51,6 ms a 90. percentil 88,4 ms.



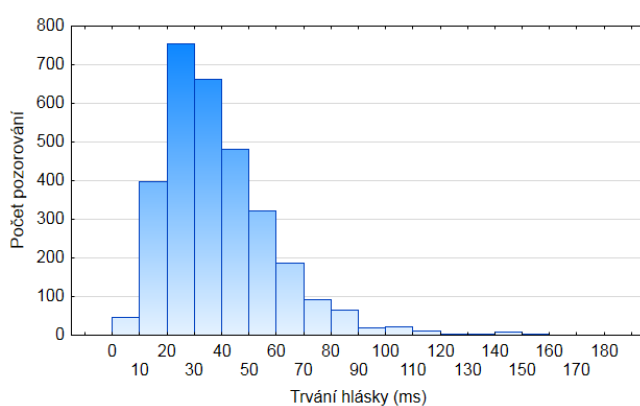
Obrázek 8.73: Histogram trvání postalveolární frikativní vibranty [ʀ] ve všech pozicích a všech hláskových okolích.

Postalveolární frikativní vibranta [ʀ] trvá průměrně 64,1 ms, se směrodatnou odchylkou 19 ms a mediánem 62,4 ms. Minimální naměřené trvání je 16,5 ms, maximální 237,3 ms, 10. percentil je 42 ms a 90. percentil 87,1 ms. Rozložení hodnot viz histogram 8.52. Vezmeme-li v úvahu pouze [ʀ] v préturách, dostaneme průměrné trvání 63,2 ms, směrodatnou odchylku 17,2 ms a medián 62,1 ms. [ʀ] v kodách jsou výrazně delší, průměrně trvají 84,8 ms, se směrodatnou odchylkou 37,3 ms a mediánem 76,1. Tento rozdíl ale nelze statisticky ověřit kvůli malému počtu vzorků [ʀ] v kodě.

Rozložení hodnot trvání laterální alveolární aproximanty [l] zobrazuje histogram 8.74. [l] trvá průměrně 46,1 ms, sm. odchylka 19,9 ms a medián 42,2 ms. Minimum je 7,1 ms, maximum 192 ms. 10. percentil je 25,4 ms, 90. percentil 71,1 ms. [l] v préturách vykazuje průměrné trvání 41,9 ms, sm. odchylku 15,1 ms a medián 40 ms. [l] v kodách pak má opět vyšší průměr, 63 ms, sm. odchylku 26,7 ms a medián 58,4 ms. Rozdíl mezi préturami a kodami je statisticky významný na datech seskupených podle mluvčího a pozice ve slabice:  $t(66) = 16,5$ ;  $p < 0,001$  (t-test pro nezávislá měření).



Obrázek 8.74: Histogram trvání laterální alveolární aproximanty [l] ve všech pozicích a všech hláskových okolích.



Obrázek 8.75: Histogram trvání palatální aproximanty [j] ve všech pozicích a všech hláskových okolích.

Na obr. 8.75 je vidět rozložení hodnot trvání palatální aproximanty [j] v analyzovaném materiálu. [j] průměrně trvá 39 ms a má sm. odchylku 20,6 ms a medián 34,5 ms. Minimální změřené trvání [j] je 7 ms, maximální 179,3 ms, 10. percentil 18 ms a 90. percentil 64,9 ms. [j] v préturách je opět o kratší než v kodách, trvá průměrně 36,4 ms se směr. odchylkou 19,2 ms a mediánem 32 ms. [j] v kodách má průměrné trvání 51,7 ms, směr. odchylku 22,5 ms a medián 47,7 ms. Statistickou významnost rozdílu mezi oběma skupinami jsme ověřili t-testem pro nezávislá měření:  $t(66) = 13,6$ ;  $p < 0,001$ .

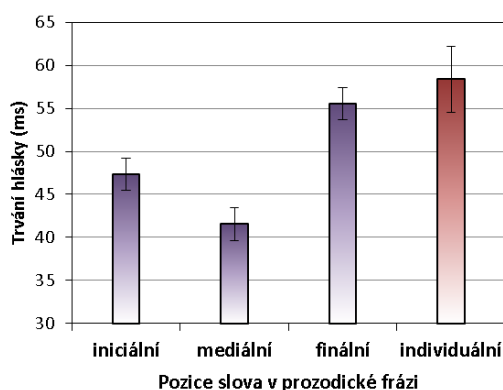
Efekt pozice ve slabice pro všechny hlásky dohromady je významný – jednofaktorová ANOVA na datech seskupených podle mluvčího, hlásky a pozice:  $F(1, 270) = 72$ ;  $p < 0,001$ . Zkoumané hlásky jsou v prétuře průměrně o 15 ms kratší než v kodě.

### 8.3.6.1 Vliv závěrového zpomalování na trvání vibrant a aproximant

Pro zjištění změn trvání vibrant a aproximant podle pozice slova v prozodické frázi byly případy opět seskupeny podle mluvčího, hlásky a pozice slova v prozodické frázi, přičemž u

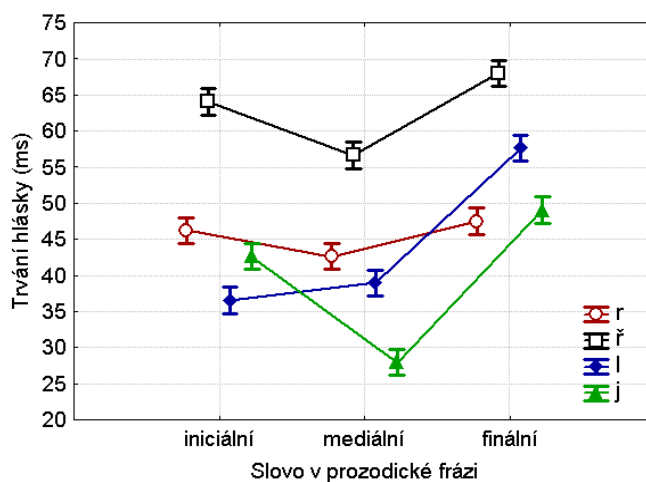
finálních jsme vyřadili případy před předělem hloubky 3. Po vyřazení položek vzniklých seskupením pěti a méně případů vypadlo slabičné [r], které budeme analyzovat zvlášť.

Průměry jednotlivých skupin jsou ukázány na obrázku 8.76:



Obrázek 8.76: Průměrné trvání vibrant a aproximant ze seskupených dat rozdělených podle pozice slova v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Efekt pozice slova v prozodické frázi je významný:  $F(3, 437) = 44,3; p < 0,001$  (jednofaktorová ANOVA pro nezávislá měření). Podle Tukeyho post-hoc testu není významný jen rozdíl mezi finálními a individuálními slovy, ostatní signifikantní jsou ( $p < 0,001$ ). Skupinu individuálních slov však zastupují v drtivé většině pouze hlásky [j], které jsou inherentně kratší. Dvoufaktorový test ANOVA jsme tedy provedli jen na iniciálních, mediálních a finálních slovech.



Obrázek 8.77: Průměrné trvání jednotlivých vibrant a aproximant ze seskupených dat podle pozice slova v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Z obr. 8.77 je zřetelné, že ne všechny konsonanty se chovají stejně. Zatímco [l] má přibližně stejné trvání v iniciálních a mediálních slovech, ztateně se prodlužuje ve finálních (Tukeyho post-hoc test:  $p < 0,001$ ). [r] a [ř] mají podobný vzorec – v mediálních slovech je hláška

nejkratší, zatímco v iniciálních a finálních delší. Rozdíl mediálních oproti finálním je v obou případech významný ( $p = 0,01$ , respektive  $p < 0,001$ ). Podobný trend se objevuje i u [j], které je opět v mediálních slovech nejkratší a v iniciálních a finálních výrazně delší ( $p < 0,001$ ).

Pro slabičné [r] jsou doklady pouze v mediálních, finálních a individuálních slovech – rozdíl mezi trváním [r] v těchto pozicích je statisticky nevýznamný ( $p = 0,52$ ).

Hláška	Multipl. koef.	Adit. koef. (ms)
r	1,11	4,9
r sl.	1,01	1
ř	1,2	11,4
l	1,48	18,6
j	1,75	21,1
Průměr	1,31	11,4

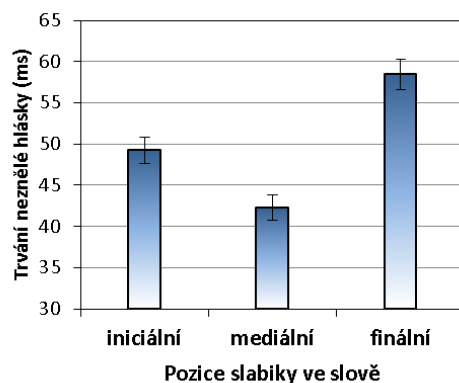
Tabulka 8.17: Koeficienty průměrných rozdílů mezi vibranti a aproximanti z mediálních a finálních slov v prozodické frázi ukončené předělem hloubky 4. Druhý sloupec obsahuje podíl finálních a mediálních, třetí sloupec jejich rozdíl.

Koeficienty v tabulce 8.17 jsou spočítány jako poměr, respektive rozdíl průměrného trvání hlásky v mediálních a finálních slovech. Vidíme, že nejvíce se prodlužuje [j], které dosahuje ve finálních slovech až 1,75násobek své délky, což je nejvíce ze všech konsonantů (druhé je [š], které se prodlužuje 1,7krát, viz tabulka 8.10). Naopak [r] dosahuje ve finálních slovech pouze 1,1násobku svého mediálního trvání. Slabičné [r] se neprodlužuje vůbec.

Pro změření efektu pozice slabiky ve finálních slovech jsme seskupili data podle mluvčího, hlásky, pozice slabiky ve slově a pozice slova v prozodické frázi. Provedli jsme stejné protřídění dat jako výše, díky čemuž jsme vyřadili slabičné [r] (vyskytuje se pouze v iniciálních a mediálních slabikách) a [ř] z důvodů příliš malého počtu dokladů v jednotlivých skupinách. Individuální slova jsme zařadili mezi finální.

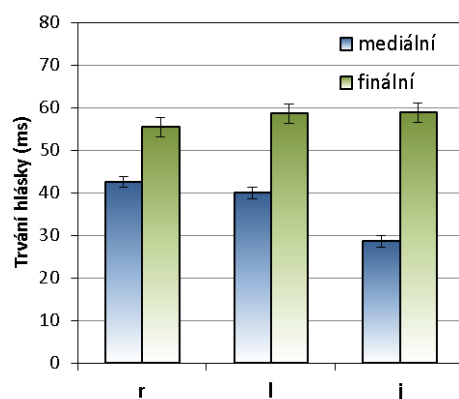
Efekt pozice slabiky je signifikantní, viz obr. 8.78 ( $F(2, 269) = 88,6; p < 0,001$ ), přičemž hlásky ve finálních slabikách finálních slov jsou nejdelší, následované slabikami iniciálními. Všechny tři rozdíly jsou statisticky významné (Tukeyho post-hoc test:  $p < 0,001$ ).

[ř] bylo následně porovnáno na neseskupených datech. V iniciálních a mediálních slovech je rozdíl v trvání [ř] nevýznamný, drobně se prodlužuje v individuálních slovech a významně ve slovech finálních (jednofaktorová ANOVA na neseskupených datech:  $F(3, 625) = 70,8; p < 0,001$ ).



Obrázek 8.78: Průměrné trvání [r], [l] a [j] ze seskupených dat ve finálních slovech podle pozice slabiky ve slově. Svorky označují 95% interval spolehlivosti.

Porovnání hlásek v mediálních slovech a finálních hlásek z finálních slov (ze stejné seskupených dat) je možné pouze pro [r], [l] a [j], kde máme dostatek dokladů pro oba případy. Průměry skupin zobrazuje graf 8.79.

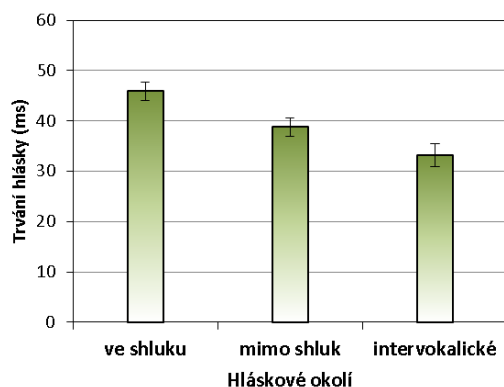


Obrázek 8.79: Průměrné trvání [r], [l] a [j] ze seskupených dat v mediálních slovech (modrá) a finálních slabikách finálních slov (zelená) v prozodické frázi. Svorky označují 95% interval spolehlivosti.

Efekt pozice je pro tyto tři hlásky vysoce významný,  $F(1, 301) = 476$ ;  $p < 0,001$ . Interakce HLÁSKA\*POZICE je také signifikantní:  $F(2, 301) = 42,8$ ;  $p < 0,001$ . Tukeyho post-hoc test naznačuje, že i rozdíly u všech tří hlásek jsou významné ( $p < 0,001$ ).

### 8.3.6.2 Vliv hláskového okolí na trvání vibrant a aproximant

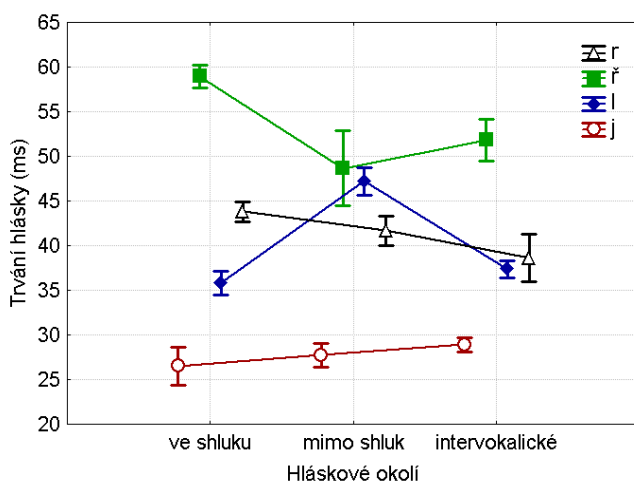
Pro tuto analýzu jsme opět využili pouze konsonanty z mediálních slov, seskupené podle mluvčího, hlásky a hláskového okolí. Seskupené položky byly protříděny podle počtu případů – ty, které obsahovaly pět a méně, byly vyřazeny. Výsledné průměry skupin hláskového okolí zobrazuje graf 8.80.



Obrázek 8.80: Průměrné trvání vibrant a aproximant v mediálních slovech podle hláskového okolí. Svorky označují 95% interval spolehlivosti.

Jednofaktorová ANOVA ukazuje na významný efekt hláskového okolí:  $F(2, 270) = 38,3; p < 0,001$ , tentokrát však ve zcela opačném směru než u ostatních konsonantů – intervokalické jsou nejkratší a vibranty a aproximanty ve shluku naopak nejdelší.

Podrobnější analýza odhalila, že se v tomto ohledu chová každá hláska jinak. Obrázek 8.81 ukazuje průměry pro všechny čtyři zkoumané hlásky z neseskupených dat.



Obrázek 8.81: Průměrné trvání jednotlivých vibrant a aproximant v mediálních slovech podle hláskového okolí z neseskupených dat. Svorky označují 95% interval spolehlivosti.

Z grafu je zřetelné, že všechny čtyři konsonanty mají svůj vlastní vzorec chování podle hláskového okolí. Dvoufaktorová ANOVA a Tukeyho post-hoc test odhaluje, že rozdíly u [j] v jednotlivých okolích jsou nesignifikantní. U [r] je významný pouze rozdíl mezi realizacemi ve shluku a intervokalickými, které jsou kratší ( $p = 0,02$ ). [ř] se neliší trváním v intervokalické pozici a mimo shluk, ve shluku je signifikantně delší ( $p < 0,001$ ). Co se týče [l], rozdíl mezi intervokalickými a realizacemi ve shluku je nevýznamný. Významně delší jsou ale realizace mimo shluk ( $p < 0,001$ ).

### 8.3.6.3 Vliv délky vyšší jednotky na trvání vibrant a aproximant

Vibranty a aproximanty z mediálních slov jsme rovněž seskupili podle mluvčího, hlásky a délky vyšší jednotky, tedy slova (ve slabikách a ve fonémech) nebo prozodické fráze (ve slovech a slabikách) a vyřadili položky, které vznikly seskupením pěti a méně případů.

Efekt délky slova je v obou případech statisticky významný, ve slabikách:  $F(3, 222) = 26,1$ ;  $p < 0,001$  i ve fonémech:  $F(3, 201) = 20,7$ ;  $p < 0,001$ . Nicméně ani jedna skupina nevykazuje jasný trend a výsledek není snadno interpretovatelný. Post-hoc analýzy ukazují, že různé aproximanty se vyskytují v různých skupinách, a tedy je zde velký vliv identity hlásky. Rovněž také počty případů jsou velmi nízké.

Korelace mezi trváním slova v ms a trváním vibrant a aproximant z neseskupených dat je nicméně nejvyšší ze všech skupin konsonantů (Pearsonovo  $r = 0,36$ ) a statisticky významná ( $p < 0,01$ ). Korelační koeficient je pozitivní, což znamená, že s prodlužováním délky slova se prodlužuje i trvání vibrant a aproximant. Analýzy po jednotlivých hláskách ukazují, že k tomu přispívají všechny vibranty a aproximanty kromě [r], jehož délka s trváním slova nekoreluje,  $r = 0,03$  – tento výsledek ovšem není statisticky signifikantní.

Co se týče trvání nazál v závislosti na trvání prozodické fráze, výsledky jsou opět problematické jako v předchozím případě. Efekt délky prozodické fráze vyjádřené ve slovech je sice významný,  $F(4, 244) = 4,65$ ;  $p = 0,001$ , nicméně opět se zde neukazuje žádný jasný trend a jednotlivé skupiny mají silně nestejně zastoupení. Délka prozodické fráze vyjádřená ve slabikách vykazuje také signifikantní vliv:  $F(5, 149) = 6$ ;  $p < 0,001$ , zde je vidět mírný trend ke zkracování vibrant a aproximant v delších frázích, nicméně opět nejsou jednotlivé skupiny srovnatelně zastoupené.

Korelace mezi trváním vibrant a aproximant a trváním prozodických frází v ms je mírně pozitivní ( $r = 0,16$ ) a statisticky signifikantní ( $p < 0,01$ ).

## 8.4 Průměrné trvání konsonantů

Abychom získali co nejrepresentativnější hodnoty trvání konsonantů, vyřadili jsme konsonanty z finálních slov prozodických frází (a ze slov, která tvořila samostatnou prozodickou frázi), iniciální hlásky po pauze a hlásky ze slov *hm*, *jo* a *no*. Také byly vyřazeny konsonanty v okolí elizí, tj. takové, které obsahovaly více než jednu značku fonému (k této metodice viz obrázek 7.2).

Konsonant	Zdroj	Počet	Průměr	Sm.odch.	Minimum	Maximum	Medián	10. percent.	90. percent.
explozivny									
<b>p</b>	bez shluků	708	<b>78,1</b>	<b>18,7</b>	21,6	178,9	78,4	55,4	99,9
<b>b</b>	intervokal.	1123	<b>58,7</b>	<b>15,3</b>	8,7	128,3	58,6	39,9	77,2
<b>t</b>	intervokal.	2387	<b>68,1</b>	<b>17,5</b>	16,8	189,8	67,2	48,0	88,4
<b>d</b>	intervokal.	1089	<b>30,7</b>	<b>10,4</b>	9,0	94,2	29,3	19,1	45,5
<b>ť</b>	bez shluků	694	<b>71,9</b>	<b>20,4</b>	19,7	148,3	71,3	47,5	96,5
<b>ď</b>	bez shluků	376	<b>52,1</b>	<b>18,3</b>	13,0	120,0	50,5	30,5	76,2
<b>k</b>	bez shluků	1439	<b>64,4</b>	<b>19,6</b>	20,4	274,0	63,0	42,3	85,5
<b>g</b>	bez shluků	270	<b>43,4</b>	<b>15,5</b>	7,9	114,1	42,2	25,9	64,0
frikativy									
<b>f</b>	bez shluků	396	<b>59,6</b>	<b>18,5</b>	21,8	134,0	58,1	35,6	83,9
<b>v</b>	bez shluků	576	<b>41,6</b>	<b>13,1</b>	11,6	96,8	40,8	25,7	58,5
<b>s</b>	intervokal.	996	<b>92,7</b>	<b>19,1</b>	32,0	160,2	91,0	70,3	118,2
<b>z</b>	bez shluků	535	<b>52,9</b>	<b>17,8</b>	9,2	128,0	53,2	28,3	74,1
<b>š</b>	bez shluků	1358	<b>71,8</b>	<b>26,2</b>	19,7	188,8	69,8	39,7	105,2
<b>ž</b>	bez shluků	988	<b>54,9</b>	<b>19,5</b>	11,7	134,6	54,5	29,8	80,0
<b>ch</b>	bez shluků	398	<b>65,8</b>	<b>21,7</b>	22,2	135,0	63,2	38,8	95,8
<b>h</b>	bez shluků	151	<b>55,4</b>	<b>21,6</b>	11,3	143,3	53,0	32,6	79,8
afrikáty									
<b>c</b>	bez shluků	809	<b>87,9</b>	<b>21,2</b>	29,4	180,7	87,9	60,8	112,0
<b>dz</b>	bez shluků	82	<b>66,7</b>	<b>21,2</b>	27,1	138,6	65,9	41,1	95,0
<b>č</b>	bez shluků	431	<b>101,5</b>	<b>21,1</b>	46,4	173,3	101,4	76,1	127,6
<b>dž</b>	vše	23	<b>63,7</b>	<b>23,2</b>	33,4	127,5	52,9	37,8	98,2
nazály									
<b>m</b>	intervokal.	1202	<b>58,0</b>	<b>15,4</b>	11,5	159,9	58,2	39,0	75,4
<b>ŋ</b>	vše	27	<b>83,1</b>	<b>23,8</b>	32,0	130,1	83,5	56,2	117,3
<b>n</b>	bez shluků	1979	<b>43,4</b>	<b>18,4</b>	7,1	274,7	40,8	24,0	64,0
<b>ň</b>	bez shluků	1094	<b>50,5</b>	<b>19,0</b>	9,3	146,4	47,9	29,1	72,7
<b>ŋ</b>	vše	34	<b>86,9</b>	<b>16,4</b>	50,2	114,4	88,6	68,3	111,5
vibranty									
<b>r nesl.</b>	bez shluků	330	<b>41,5</b>	<b>10,6</b>	16,7	83,4	40,5	29,1	55,1
<b>r slabičné</b>	vše	273	<b>71,1</b>	<b>14,6</b>	23,9	117,8	71,9	51,6	88,4
<b>ř</b>	bez shluků	189	<b>53,8</b>	<b>16,6</b>	23,6	121,1	49,5	35,5	77,3
aproximanty									
<b>l</b>	intervokal.	1015	<b>37,7</b>	<b>11,2</b>	8,4	100,8	36,8	24,2	52,6
<b>j</b>	intervokal.	916	<b>28,6</b>	<b>12,3</b>	7,2	146,3	26,4	15,7	43,7

Tabulka 8.18: Reprezentativní trvání českých konsonantů naměřené na vybraných částech korpusu Minidialogy-H. Sloupec Zdroj uvádí, ze kterých hláskových okolí byly průměry vypočítány. Hodnoty ve čtvrtém až desátém sloupci jsou uvedeny v milisekundách.

V ideálním případě by bylo žádoucí pracovat pouze s intervokalickými konsonanty, kde lze předpokládat nejmenší ovlivnění hláskovým okolím. Bohužel v takovém případě nebudeme



mít pro některé hlásky k dispozici dostatek dokladů (pokud vůbec nějaké). Rozhodli jsme se tedy pro konsonanty, které mají kolem 1000 intervokalických realizací nebo více, použít pouze intervokalickou pozici a pro ostatní alespoň vyřadit realizace z konsonantických shluků v rámci slabiky. Co se týče slabičného [r], pokud bychom využili jen realizace z nefinálních slov, šlo by pouze o jediné slovo, první. Protože se ale ukázalo, že [r] se v závislosti na pozici v prozodické frázi příliš neprodlužuje (viz tabulka 8.17), použili jsme pro spočítání reprezentativního průměru všechny realizace. Pro [d̃ž], [m̃] a [ŋ] jsme také museli využít všechny realizace, jelikož se jich v korpusu nachází celkově velmi málo.

Výsledné hodnoty z celkem 21 888 realizací konsonantů udává tabulka 8.18.

## 8.5 Diskuse k deskriptivním analýzám

V předchozí kapitole bylo zevrubně popsáno trvání jednotlivých českých hlásek měřené na semispontánním korpusu Minidialogy-H. Prozkoumán byl vliv pozice hlásky ve slabice (prétura či koda), pozice slova v prozodické frázi (iniciální, mediální či finální), pozice slabiky ve finálním slově (iniciální, mediální či finální), hláskového okolí (intervokalické, ve shluku v rámci slabiky nebo ve shluku mimo rámec slabiky, u vokálů přítomnost rázu), délky slova a prozodické fráze, u vokálů také struktury slabiky.

Uvedené výsledky jsou všeobecné a každý jednotlivý jev by si zasloužil ještě další prozkoumání, které však již nemůže být součástí této práce a čeká tedy na budoucí výzkum.

Celkovým průměrným datům bohužel nemůžeme říkat referenční a doporučit je s čistým svědomím pro veškeré forenzní účely, protože nejsou měřeny na dostatečně bohatém vzorku populace. Nicméně množství materiálu zaručuje reprezentativnost alespoň v populaci mladých dospělých mluvčích z Čech.

Histogramy trvání jsou obecně pravostranně sešikmené, liší se však ve špičatosti a rozptylu rozdělení. Kratší obstruenty, zejména [b], [g] či [v] mají rozdělení téměř symetrické, u ostatních hlásek se blíží lognormálnímu.

Hodnoty trvání vokálů jsou oproti starším studiím – tj. Chlumskému (1928), Kaiserové (1964) a Borovičkové a Maláčovi (1967) o polovinu a více kratší. To je důsledkem především povahy našeho materiálu – tyto zmíněné studie pracovaly s izolovanými zřetelně vyslovovanými slovy (v pořádku Borovičkové a Maláče dokonce pseudoslovy), zatímco v korpusu Minidialogy jde sice o řeč čtenou, nicméně plynulou a v rámci dialogu.

Stejně tak jsou vokály poloviční i ve srovnání s materiálem Psutky et al. (2006), a to jak čteným, tak dokonce spontánním. Zde je třeba příčinu hledat jednak v segmentaci – automatické zarovnávání, které autoři použili, pravděpodobně bonifikovalo trvání vokálů na

úkor konsonantů – ale také v povaze materiálu. Ve spontánní řeči se vyskytuje velké množství hezitací, dysfluencí a jiných fenoménů, které mohou mít vliv na trvání hlásek – v naší studii byly tyto vlivy důsledně kontrolovány.

Srovnáme-li výsledky se studii Podlipského et al. (2009) a Studenovského (2012), dostáváme se již řádově do stejných hodnot, v přechodí literatuře opět o něco delších. Tyto drobné rozdíly budou jednak opět souviset s mluvním stylem ve využitém materiálu – obě studie měly k dispozici čtené monology – ale možnou příčinou je také to, že v naší práci byl kladen důraz na odfiltrování prozodických vlivů, především závěrového zpomalování.

Krátké vokály v naší práci kopírují hypotézu o inherentním trvání vokálů v závislosti na vertikální poloze jazyka – otevřený vokál [a] je nejdelší, následován [e] a [o] a nejkratší jsou zavřené [i] a [u]. V žádné z citovaných studií se tento efekt ovšem neprojevil, nejvíce se tomu blíží hodnoty Chlumského (1928), kde je však [e] stejně dlouhé jako [u]. Ve většině novějších studií (Psutka et al., 2006; Studenovský, 2012 a do jisté míry i u Podlipského et al., 2009) jsou středové vokály dokonce delší než otevřený, viz obrázek 5.1a. Tato diskrepance by mohla být způsobena prozodickými (či jinými) faktory, které ve zmiňovaných studiích nebyly odfiltrovány.

Dalo by se očekávat, že trvání dlouhých vokálů bude kopírovat trvání krátkých. V našem materiálu se tak neděje, a to především z důvodů nedostatečného (a nereprezentativního) zastoupení vokálů [e:] a [o:]. [o:] je k dispozici poze v částicích *no* a *jo* a 53 % realizací [e:] pochází se slova *ne*. Není tedy nijak překvapivé, že tyto dvě hlásky vykazují suverénně nejdelší trvání. Co se týče zbylých tří, ty kopírují relativní vztah, který se projevil ve všech ostatních citovaných zdrojích – průměrně nejkratší je [i:], dále [u:] a nejdelší z této trojice pak [a:] (viz obrázek 5.1b).

Poměr dlouhých vokálů a jejich krátkých protějšků (pomineme-li opět [e:] a [o:]) přibližně odpovídá údajům z novější literatury. Tradičně uváděný údaj, že české dlouhé vokály jsou přibližně dvojnásobně delší než krátké (Palková, 1994: 179) tedy podle těchto výsledků neodpovídá stavu v současné mluvené češtině. V našich datech se projevila tendence k malému prodlužování [i:] oproti [i] (přibližně 1,2krát), které je doprovázeno kontrastem v kvalitě. Poměrně málo se v našem materiálu prodlužovalo také [u:] vzhledem k [u], pouze 1,3krát, [a:] potom bylo průměrně 1,6krát delší než jeho krátký protějšek. Vztahy mezi těmito třemi hláskami přitom kopírují výsledky novějších studií (Psutka et al., 2006; Podlipský et al., 2009 a Studenovský, 2012).

Lze tedy uzavřít, že otevřené samohlásky jsou v češtině průměrně delší než zavřené, a u jejich dlouhých protějšků se tento efekt projevuje ještě zřetelněji. Pokud bychom měli

reprezentativní vzorek dlouhých středových vokálů, očekávali bychom výsledky někde mezi hodnotami pro otevřené a zavřené.

Diftong [ou̯] trvá v našem materiálu průměrně 80,3 ms, což je méně než v ostatních studiích, které zkoumaly diftongy (Psutka et al., 2006; Podlipský et al., 2009 a Studenovský, 2012). Jeho trvání je kratší než součet průměrného trvání monoftongů [o] a [u]:  $47 + 45 = 92$  ms, což koresponduje se všemi výše citovanými studiemi. Poměr trvání [ou̯] a součtu monoftongů je v této práci 0,87, v literatuře se pohybuje mezi 0,79 (Psutka et al., 2006, čtená řeč) až 0,94 (Podlipský et al., 2009). V korpusu Minialogy-H se další dva české diftongy, [au̯] a [eu̯], bohužel nevyskytují.

Rovněž hodnoty trvání konsonantů jsou vesměs nižší než v literatuře citované v oddílu 5.1.2, a to dokonce kratší než hodnoty uváděné pro spontánní a semispontánní řeč (Psutka et al., 2006; Machač, 2006; Šimek, 2010 nebo Ondrušková, 2011 – s výjimkou [j]). Zde opět můžeme hypotetizovat, že jde o žádoucí výsledek odfiltrování vlivu závěrového zpomalování na koncích prozodických frází.

Výrazně nižší než v citované literatuře je v našem materiálu průměrné trvání [d]. Důvodem je pravděpodobně to, že naše [d] byla měřena v intervokalické pozici, kde se mohou zkracovat na alveolární švih. Průměrné trvání frikativ se od citovaných zdrojů liší více v neznělých než znělých. Pozoruhodně konzistentní je nicméně [s], které vykazuje velmi podobné trvání ve všech studiích včetně této. Zajímavý je vztah mezi [s] a [š] – zatímco u Chlumského (1928), Borovičkové a Maláče (1967) a v této práci je [s] delší, u Psutky et al. (2006) a Homolkové (2009) je delší [š], i když v některých případech jsou tyto rozíly jen v řádu jednotek milisekund. Zdálo by se rozumné tedy vyvodit, že tyto hlásky mají v češtině velmi podobné trvání, což potvrzují i studie o angličtině (Maniwa et al., 2009). Nesoulad v našem materiálu je opět pravděpodobně způsoben tím, že [s] bylo měřeno v intervokalických pozicích, kde nabývá delšího trvání než v neintervokalických. Podobnou diskrepanci můžeme najít i u dvojice [c] a [č] – zatímco starší studie uvádějí jejich trvání také víceméně stejné, zde vyšlo [č] delší než [c]. Příčinou může být rozdílné zastoupení v materiálu a vliv některých dalších prozodických faktorů.

Znělé explozivy jsou v našem materiálu obecně kratší než neznělé, i když ne o tolik jako například v materiálu Chlumského (1928) nebo Machače (2006). Nejkratší jsou přitom explozivy alveolární, což také odpovídá většině citovaných zdrojů. Nejdelší potom bilabiální, kde již výsledky v ostatních studiích nejsou zajedno, nicméně podobný trend vykazuje také Šimek (2010). Chlumský (1928) uvádí jako nejdelší velární explozivy, u ostatních to nelze jednoznačně říci (liší se pro znělé a neznělé).

Také neznělé frikativy jsou v našem materiálu obecně delší než znělé, nicméně ne o tolik jako ve starší literatuře (Chlumský, 1928; Borovičková & Maláč, 1967; Homolková, 2009). Nejkratší neznělá, [f], je téměř stejně dlouhá jako nejdelší znělá, [h]. Hodnoty nazál, vibrant i aproximant se také shodují se starší literaturou, dokonce průměrné trvání slabičného a neslabičného [r] je velmi podobné jako u Chlumského (1928) a Vernerové (2006).

Ukazuje se, že obecné tendence trvání konsonantů uvedené v literatuře (viz oddíl 5.1.2) jsou poněkud zjednodušující a často je nelze přijmout bez výhrad (kromě výsledku, že neznělé konsonanty jsou obecně delší než jejich znělé protějšky, to se potvrdilo bez výjimek). Nepotvrdil se vztah, že by neznělé afrikáty a frikativy byly obecně nejdelšími českými konsonanty – platí to pouze pro afrikáty. Neznělé explozivy jsou často podobně dlouhé jako neznělé frikativy. Tvrzení, že nejkratší explozivy jsou alveolární, také není zcela neproblematické, [t] je kratší než ostatní neznělé explozivy jen nepatrně a v seskupených datech dokonce vůbec. Lze přijmout tvrzení, že nejdelší frikativy jsou alveolární a postalveolární, nicméně trvání [h] je téměř stejné jako [z] nebo [ž]. Co naše data ovšem potvrzují, je fakt, že afrikáty jsou kratší než součet trvání jim odpovídajících dvou obstruentů. Vliv pozice konsonantu ve slabice na jeho trvání byl celkově statisticky nevýznamný, protože u různých konsonantů se projevuje různě – a to i v rámci jedné skupiny konsonantů. Z toho můžeme usoudit, že ostatní faktory (např. znělost, hláskové okolí, pozice v rámci vyšší jednotky) budou pro modifikace trvání důležitější.

Z našich výsledků vyplývá, že závěrové zpomalování nepostihuje všechny hlásky stejně – vokály se prodlužují více než konsonanty a neznělé obstruenty více než znělé. Sonory se prodlužují přibližně stejně jako znělé obstruenty.

Změna trvání vlivem hláskového okolí také není konstantní. Obstruenty (s výjimkou [ř] a již zmíněného [d]) jsou nejdelší v intervokalické pozici a mají-li ve svém okolí alespoň jeden konsonant, pak se významně zkracují. Nazály jsou ještě citlivější, ve shluku v rámci slabiky se zkracují ještě více, než pokud sousedící konsonant není ve stejné slabice. U vibrant a aproximant je tento efekt přesně opačný než u nazál – v intervokalické pozici jsou nejkratší. Podrobnější pohled ale napovídá, že jednotlivé hlásky v této kategorii se chovají rozdílně.

Vliv hláskového okolí na trvání segmentů by bylo záhodno v budoucnu prozkoumat podrobněji, v souvislosti s velikostí shluku a pozicí shluku ve slabice či slově – zdá se, že tyto faktory by také mohly přispívat ke změně trvání konsonantů (Churaňová et al., v recenzním řízení).

Z našich výsledků je zřetelné, že vliv délky vyšší jednotky (ať už slova nebo prozodické fráze) na trvání jednotlivých hlásek není lineární. Po uvážení výsledků Dankovičové (2001:

51 nn.) bychom očekávali, že hlásky se budou se zvyšující délkou slova a fráze zkracovat, nicméně tento efekt nacházíme jen u krátkých vokálů (viz sekce 8.1.10). U všech ostatních tříd hlásek je změna trvání nelineární. I tento faktor by si zasloužil blíže prozkoumat, především se podrobněji podívat na jednotlivé hlásky a kontrolovat nerovnoměrné rozložení případů v korpusu.

Mezi faktory, které by dále mohly ovlivňovat trvání hlásek a nebyly v této sekci zkoumány, patří zejména sémantická třída slova – již v úvodu bylo zmíněno, že synsémantická slova by se měla z hlediska trvání segmentů chovat jinak než autosémantická. K této problematice bude nicméně ještě přihlíženo v další části práce.

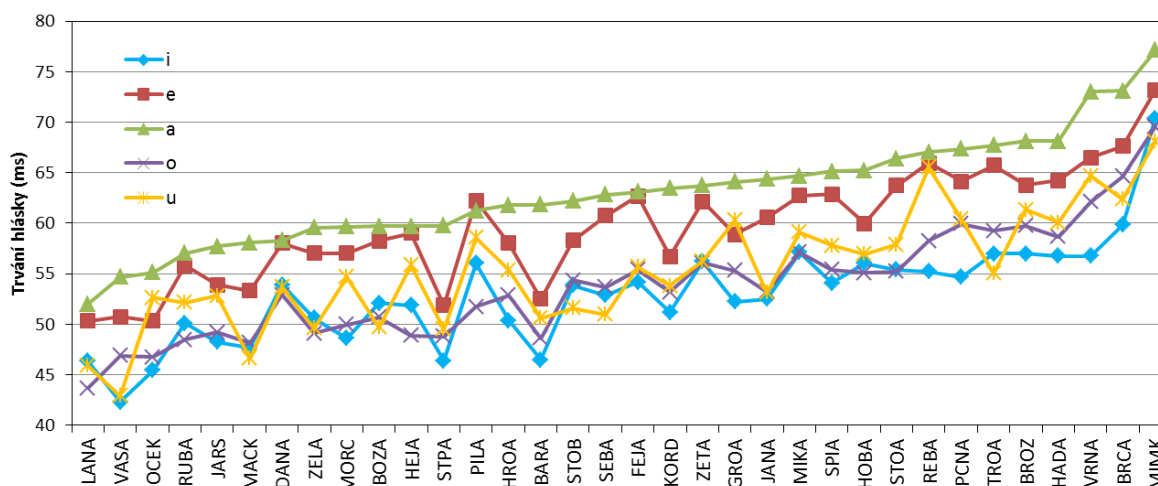
## 9 Jevy specifické pro mluvčího

### 9.1 Trvání segmentů

Tato část práce se bude zabývat specifičností trvání hlásek, případně jejich skupin pro jednotlivé mluvčí. Při těchto analýzách již data nebudou seskupována, protože cílem není generalizovatelnost výsledků, ale to, zda z celkové variability všech dat lze nějakou část vysvětlit idiosynkratickým chováním mluvčích.

Úspěšnost jednotlivých ukazatelů pro rozpoznání mluvčího budeme navzájem porovnávat s pomocí výsledků Tukeyho post-hoc testu. Při 34 mluvčích je možných 561 různých párů mluvčích. Úspěšnost ukazatele bude pak procento párů, které Tukeyho post-hoc test označí za signifikantně odlišné. Je zde nutno podotknout, že tato procedura je dosti přísná, protože test bere v úvahu variabilitu v celé skupině všech mluvčích – kdybychom prováděli např. t-testy na jednotlivých párech mluvčích, ke statistické významnosti by byl dostatečný menší rozdíl.

Jako první bude analyzováno trvání vokálů. Obrázek 8.82 ukazuje průměrné trvání pěti krátkých vokálů pro každého mluvčího. Pro větší přehlednost jsou odstraněny svorky s intervalem spolehlivosti (které se pohybují řádově mezi 3-5 ms na obě strany) a data jsou seřazena podle trvání nejdelšího vokálu, čili [a]. Mluvčí jsou kódováni čtyřpísmennými zkratkami (viz oddíl 7.1).



Obrázek 8.82: Průměry trvání krátkých vokálů z neseskupených dat pro jednotlivé mluvčí, seřazené vzestupně podle trvání [a].

Z obrázku je vidět, že mluvčí mají poměrně konzistentní průměrné trvání všech pěti krátkých vokálů – hodnoty se pohybují v pásmu širokém průměrně 11 ms a maximálně 16 ms (u mluvčí VRNA). Extrémem je mluvčí MIMK (zcela vpravo), jehož krátké vokály trvají mezi

66 a 77 milisekundami. Na druhé straně grafu naopak můžeme vidět mluvčí, jejichž trvání vokálů spadá pod 55 ms.

Také si lze všimnout forenzně užitečné informace, že průměrné trvání [ɪ], [o] a [u] není u žádného mluvčího delší než průměrné trvání [a]. Ve většině případů (kromě dvou mluvčích) je také [e] delší než tyto tři vokály. [e] trvá déle než [a] pouze u jednoho z mluvčích. Tento vztah lze přímo využít v praxi – pokud bude na jedné z porovnávaných nahrávek [a] trvat průměrně např. 57 ms (jako u mluvčí RUBA) a na druhé z nahrávek bude [ɪ] trvat průměrně 60 ms (jako u mluvčí BRCA), ukazuje to s vysokou mírou jistoty na rozdílnost mluvčích. Tento vztah je obzvláště výhodný, pokud v porovnávaných nahrávkách nejsou k dispozici srovnatelné počty jednotlivých vokálů. Ovšem je třeba přísně kontrolovat, zda rozdíly nejsou způsobeny odlišným mluvním stylem, afektivním stavem nebo jiným externím faktorem. Trvání krátkých vokálů napříč všemi mluvčími se tedy zdá být na sobě nějakým způsobem závislé, z průměrného trvání jednoho lze do jisté míry předpovědět i průměrné trvání ostatních vokálů. Korelační matici průměrů z grafu 8.82 zobrazuje tabulka 8.19.

	i	e	a	o	u
i		0,94	0,87	0,91	0,82
e	0,94		0,91	0,93	0,86
a	0,87	0,91		0,96	0,85
o	0,91	0,93	0,96		0,86
u	0,82	0,86	0,85	0,86	

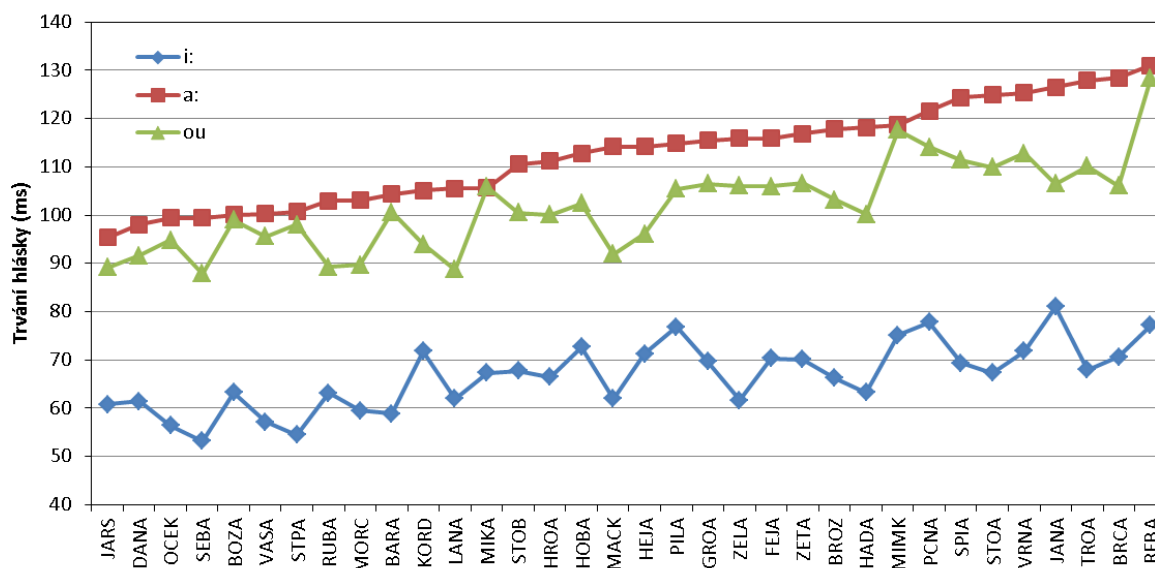
Tabulka 8.19: Korelační matice průměrů trvání krátkých vokálů u jednotlivých mluvčích.

V tabulce jsou uvedeny Pearsonovy korelační koeficienty a jak je vidět, ve všech případech jde o velmi silnou korelaci,  $r \geq 0,82$ . Ačkoliv můžeme vidět individuální odchylky od tohoto trendu (kupříkladu u mluvčích STPA a BARA), hlavní specifičnost je třeba hledat v relativním umístění celého vektoru průměrů vokálů.

Prozkoumali jsme též možnost, že na toto umístění má vliv artikulační tempo (viz také další kapitola). Korelace trvání [a] a tempa (hláskového) je  $r = -0,67$  (středně silná korelace, statisticky významná:  $p < 0,001$ ), čili tempo nějakou část této variability vysvětlí, ale nikoliv všechno. Například nejvyšší artikulační tempo má mluvčí MORC, který je v grafu až devátý zleva; mluvčí LANA, která má nejnižší průměrné trvání krátkých vokálů, je až šestá nejrychlejší. Totéž platí opačně, mluvčí MIMK s výrazně nejdelším průměrným trváním vokálů je až 10. nejpomalejší podle artikulačního tempa.

Lze tedy říci, že trvání krátkých vokálů závisí kromě faktorů, které byly podrobně probrány v kapitole 8.1 (identita vokálu, hláskové okolí, struktura slabiky, pozice v rámci slova a prozodické fráze) a artikulačním tempu, které bude obsahem kapitoly 9.2, také na mluvčím.

Dlouhé vokály nevykazují tak silný trend vzájemné závislosti. Jejich průměry pro jednotlivé mluvčí jsou zobrazeny v grafu 8.83. [e:], [o:] a [u:] nejsou do analýzy zařazeny, protože jejich počet výskytů je řádově nižší, ostatní dlouhé vokály se v korpusu vyskytují více než 1000krát, což zajišťuje minimálně 30 realizací na mluvčího.



Obrázek 8.83: Průměry trvání dlouhých vokálů z neseskupených dat pro jednotlivé mluvčí, seřazené vzestupně podle trvání [a:].

Korelace je středně silná,  $r$  se pohybuje mezi 0,7 a 0,8, a statisticky signifikantní ( $p < 0,001$ ). Vidíme zde více individuálních rozdílů, které nicméně nejsou tak systematické jako u krátkých vokálů.

V tabulce 8.20 jsou uvedeny korelační koeficienty pro všechny zde analyzované vokály (tedy pět krátkých, [i:], [a:] a diftong [ou]).

	i	e	a	o	u	i:	a:	ou
i		0,94	0,87	0,91	0,82	0,10	0,59	0,62
e	0,94		0,91	0,93	0,86	0,11	0,73	0,74
a	0,87	0,91		0,96	0,85	0,00	0,71	0,72
o	0,91	0,93	0,96		0,86	-0,01	0,65	0,70
u	0,82	0,86	0,85	0,86		0,14	0,63	0,72
i:	0,10	0,11	0,00	-0,01	0,14		0,13	0,29
a:	0,59	0,73	0,71	0,65	0,63	0,13		0,81
ou	0,62	0,74	0,72	0,70	0,72	0,29	0,81	

Tabulka 8.20: Korelační matice průměrů trvání vokálů u jednotlivých mluvčích. Koeficienty nižší než 0,3 jsou označeny šedě.



Z tabulky je vidět, že všechna trvání spolu u jednotlivých mluvčích středně silně až velmi silně korelují ( $r \geq 0,59$ ), kromě [i:], jehož všechny korelační koeficienty jsou nižší než 0,3 a výrazně se tak liší od všech ostatních. Také [i:] je jediný vokál, u něhož nejsou korelace statisticky významné ( $p \geq 0,1$ ). Tento dlouhý vokál se tedy chová napříč mluvčími výrazně jinak než ostatní. To může souviset i s upozadováním kontrastu délky mezi [i] a [i:] ve prospěch kontrastu v kvalitě – trvání [i:] pak není tolik vázáno fonologickou délkou a může být citlivější na jiné faktory.

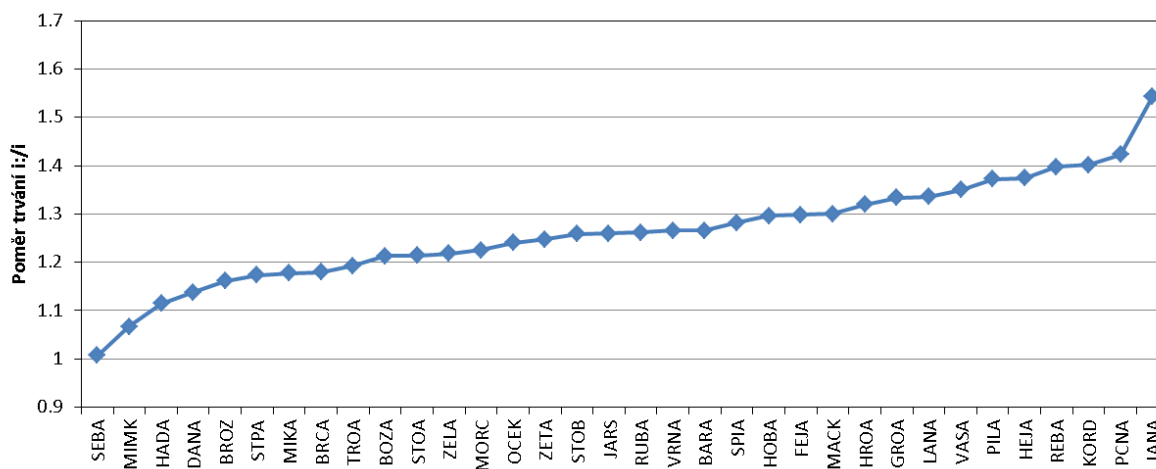
Abychom prozkoumali specifičnost trvání jednotlivých vokálů, provedli jsme sérii jednofaktorových testů ANOVA na neseskupených datech. Opět jsme využili jen vokály zmíněné výše, abychom měli minimální zastoupení 30 realizací na jednoho mluvčího. Výsledky jsou uvedeny v tabulce 8.21.

Hláška	ANOVA	Úspěšnost
i	$F(33,5444) = 7,6; p < 0,001$	14,3 % (80)
e	$F(33,13840) = 17,9; p < 0,001$	37,8 % (212)
a	$F(33,7743) = 11,5; p < 0,001$	25 % (140)
o	$F(33,8014) = 12,8; p < 0,001$	28,2 % (158)
u	$F(33,2983) = 4,3; p < 0,001$	6,8 % (38)
i:	$F(33,3239) = 5,5; p < 0,001$	10,2 % (57)
a:	$F(33,2659) = 5,5; p < 0,001$	11,8 % (66)
ou	$F(33,985) = 1,7; p = 0,007$	0,9 % (5)

Tabulka 8.21: Výsledky jednofaktorových testů ANOVA s faktorem MLUVČÍ a závislou proměnnou trvání dané hlásky. Ve třetím sloupci je uvedena úspěšnost testu jako procentuální podíl počtu dvojic, které Tukeyho post-hoc test označil za signifikantně odlišné (z 561 možných kombinací). V závorce následuje počet těchto dvojic.

Při počtu případů vyšším než tisíc a 34 mluvčích není vysoce významný výsledek u všech testů nijak překvapivý. Rovněž hodnota testového kritéria  $F$  odráží spíše velikost vzorku než specifičnost trvání jednotlivé hlásky. Zajímavější je v tomto ohledu informace ve třetím sloupci. Zde vidíme, že [e], u něhož je k dispozici více než deset tisíc případů, vykazuje úspěšnost nejvyšší, jako signifikantně odlišné bylo označeno 212 dvojic mluvčích z 561 možných. Naopak [ou], které se v korpusu vyskytuje 1019krát, rozlišilo pouze 5 dvojic. Nemáme-li tedy k dispozici opravdu velké množství dat, samotné trvání jednotlivých hlásek k rozpoznání mluvčího příliš nepřispěje. Důvodem je i to, že je ovlivněno mnoha dalšími faktory.

Protože vokál [i:] se chová jinak než ostatní, byl analyzován též poměr trvání [i:] k jeho krátkému protějšku u jednotlivých mluvčích. Na obrázku 8.84 je zobrazen výsledek.



Obrázek 8.84: Poměr trvání vokálů [i:]/[i] pro jednotlivé mluvčí, seřazené od nejnižšího po nejvyšší.

Graf zobrazuje poměr průměrného trvání [i:] a [i] pro jednotlivé mluvčí. Poměry se pohybují mezi 1 (mluvčí SEBA, jejíž dlouhá [i:] jsou průměrně stejně dlouhá jako krátká [i]) a 1,54 (mluvčí JANA), přičemž většina mluvčích leží na přímce mezi 1,15 a 1,4. Stejně jako u trvání krátkých vokálů se tedy mluvčí nacházejí v bodech kontinua, jehož extrémy tvoří právě mluvčí SEBA a JANA.

Stejná série jednofaktorových testů ANOVA byla provedena i na konsonantech, opět pouze těch, které dosahovaly počtu výskytů větších než 1000. U neznělých exploziv a afrikát byly vyřazeny realizace následující po pauze. Výsledky jsou uvedeny v tabulce 8.22.

U konsonantů je možné pozorovat větší rozdíly než u vokálů. Rovněž zde není tak přímočará závislost na počtu případů. Například [p] a [d] mají podobný počet realizací, ale výsledek testu je výrazně lepší pro [p]. [t], přestože má zdaleka nejvyšší počet případů, je jen o něco málo lepší v rozpoznání než [b], kde je počet případů k dispozici mnohem menší. U [š] není dokonce výsledek statisticky významný vůbec. Dále nenalzáme mnoho rozdílů mezi mluvčími (úspěšnost menší než 10 %) u [tʰ], [k], rázu, [v], [ž], [c], [ň], [r], [l] a [j].

Obecně se zdá, že konsonanty vykazují ještě méně variability mezi mluvčími než vokály, i když i zde lze najít zajímavé individuální odchylky. Například mluvčí BROZ vykazuje zdaleka nejdelší rázu ze všech mluvčích a mluvčí BOZA má abnormálně dlouhá [ř].

Hláška	ANOVA	Úspěšnost
<b>p</b>	$F(33,3440) = 15; p < 0,001$	32,4 % (182)
<b>b</b>	$F(33,2061) = 8,3; p < 0,001$	18,7 % (105)
<b>t</b>	$F(33,7071) = 9; p < 0,001$	19,4 % (109)
<b>d</b>	$F(33,3580) = 7,4; p < 0,001$	17,8 % (100)
<b>ť</b>	$F(33,1132) = 3,5; p < 0,001$	4,5 % (25)
<b>k</b>	$F(33,2750) = 4,9; p < 0,001$	9,3 % (52)
<b>ráz</b>	$F(33,1258) = 3,2; p < 0,001$	5,9 % (33)
<b>v</b>	$F(33,2363) = 4,4; p < 0,001$	6,2 % (35)
<b>s</b>	$F(33,5392) = 8; p < 0,001$	18,2 % (102)
<b>z</b>	$F(33,1798) = 5,6; p < 0,001$	12,1 % (68)
<b>š</b>	nevýznamné	0 % (0)
<b>ž</b>	$F(33,1049) = 2,5; p < 0,001$	1,4 % (8)
<b>c</b>	$F(33,1395) = 2,8; p < 0,001$	0,7 % (4)
<b>m</b>	$F(33,4812) = 6,7; p < 0,001$	16,4 % (92)
<b>n</b>	$F(33,4746) = 6; p < 0,001$	11,4 % (64)
<b>ň</b>	$F(33,2470) = 3,7; p < 0,001$	3,6 % (20)
<b>r</b>	$F(33,2480) = 3,3; p < 0,001$	3,9 % (22)
<b>ř</b>	$F(33,1596) = 6; p < 0,001$	11,9 % (67)
<b>l</b>	$F(33,3040) = 2,1; p < 0,001$	0,5 % (3)
<b>j</b>	$F(33,3043) = 2,7; p < 0,001$	2,3 % (13)

Tabulka 8.22: Výsledky jednofaktorových testů ANOVA s faktorem MLUVČÍ a závislou proměnnou trvání dané hlásky. Ve třetím sloupci je uvedena úspěšnost testu jako procentuální podíl počtu dvojic, které Tukeyho post-hoc test označil za signifikantně odlišné (z 561 možných kombinací). V závorce následuje počet těchto dvojic.

## 9.2 Tempo řeči<sup>7</sup>

Tempo řeči je často uváděno mezi charakteristikami typickými pro jednotlivé mluvčí. Zde měříme tempo artikulační, tj. bez zahrnutí pauz. Repliky, které obsahovaly dysfluence nebo hezitace, byly celé vyřazeny. Průměrné artikulační tempo daného mluvčího je spočítáno jako průměr artikulačních temp ve všech jím vyslovených replikách. Je třeba také podotknout, že mluvčí při experimentu nedostali žádné instrukce týkající se tempa jejich řeči, tedy lze předpokládat, že zvolili tempo jim nejpřirozenější.

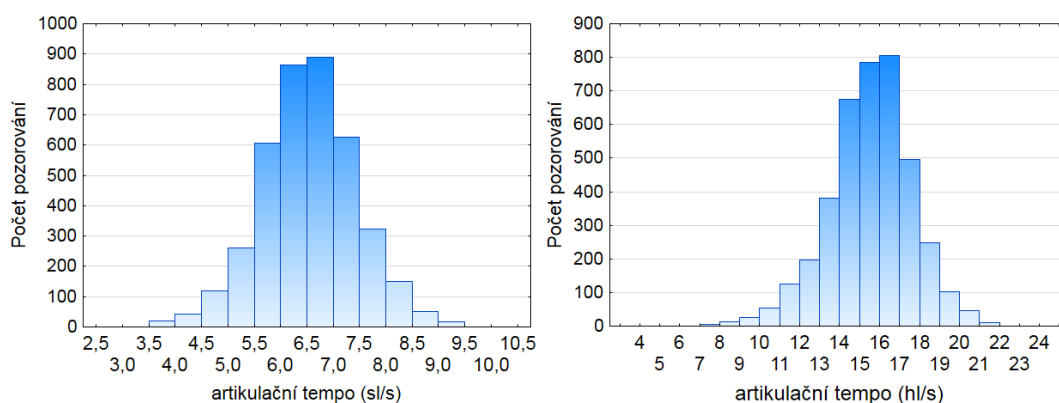
Měřeno bylo hláskové i slabičné tempo, to znamená hlásky a slabiky za sekundu. Tyto dva ukazatele spolu velmi silně korelují, Pearsonův korelační koeficient  $r$  je 0,98 ( $p < 0,01$ ), přesto bude prozkoumáno, zda některý z nich není vhodnější pro zachycení specifík mluvčích.

Celkové průměrné artikulační tempo v prozkoumaném materiálu je 6,54 slabik za sekundu (sl/s) a 15,55 hlásek za sekundu (hl/s). Nejnižší naměřené artikulační tempo je 3,53 sl/s a

<sup>7</sup> Analýza tempa a globálních temporálních ukazatelů (následující kapitola) je ve stručnější podobě prezentována ve Weingartová & Volín (2014).

7,38 hl/s (pro jednotlivou repliku) a 6,03 sl/s a 14,29 hl/s (pro jednotlivého mluvčího, konkrétně mluvčí REBA). Naopak nejvyšší naměřené tempo v jedné replice bylo 9,56 sl/s a 21,77 hl/s a průměrně má nejvyšší artikulační tempo mluvčí MORC – 7,38 sl/s a 17,51 hl/s. Stejného slabičného tempa dosahuje i mluvčí VASA, ta má ale o něco nižší hláskové tempo. Směrodatné odchylky artikulačního tempa jednotlivých mluvčích jsou poměrně nízké, pohybují se mezi 0,7 a 1,1 sl/s a mezi 1,5 a 2,6 hl/s.

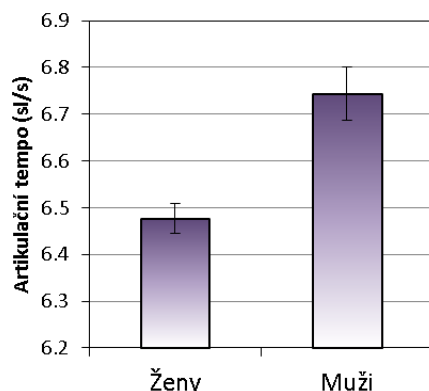
Histogramy na obrázku 8.85a a b zobrazují rozložení hodnot slabičného a hláskového tempa ve všech analyzovaných replikách v našem materiálu.



Obrázky 8.85a-b: Histogramy hodnot artikulačního tempa všech mluvčích měřeného v každé replice zvlášť ve slabikách za sekundu (vlevo) a hláskách za sekundu (vpravo).

Z histogramů je vidět, že hodnoty slabičného i hláskového tempa všech mluvčích jsou přibližně symetricky rozděleny kolem průměru.

Dále byl prozkoumán také vliv pohlaví mluvčího na artikulační tempo. Graf 8.86 zobrazuje výsledek jednofaktorového testu ANOVA s faktorem POHLAVÍ.



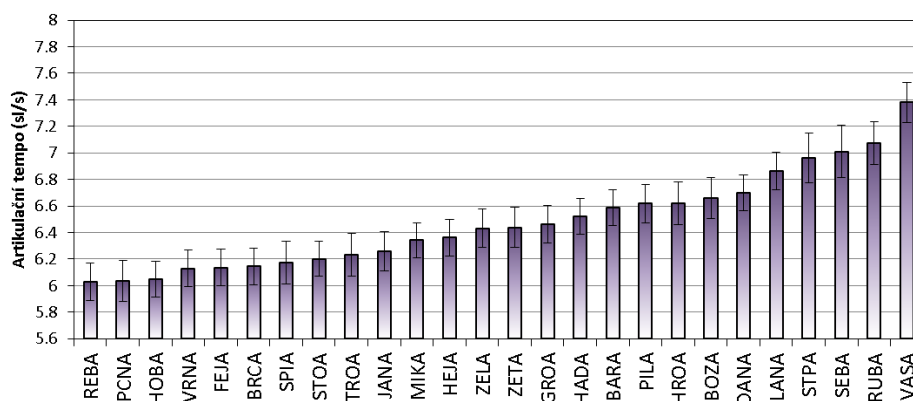
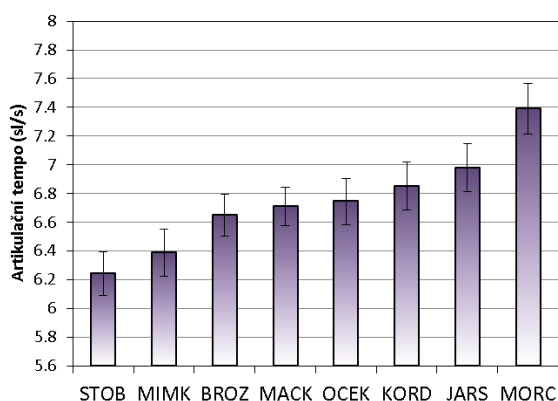
Obrázek 8.86: Průměrné artikulační tempo (ve slabikách za sekundu) podle pohlaví mluvčích.

Ve shodě s výsledky z literatury (viz např. Byrd, 1994; Künzel et al., 1995 nebo Jacewicz et al., 2010) mají i v našem materiálu muži průměrně vyšší artikulační tempo než ženy. Přestože

je mužů méně, rozdíl je statisticky významný:  $F(1, 3967) = 64,2$ ;  $p < 0,001$ . U hláskového tempa je rozdíl menší ( $F = 36,3$ ), přesto však stále signifikantní. Ženy mají průměrné artikulační tempo 6,48 sl/s a 15,44 hl/s, zatímco muži průměrně 6,74 sl/s a 15,9 hl/s.

Efekt mluvčího je také vysoce významný pro obě měření (slabičné i hláskové tempo), pro slabičné opět o něco více:  $F(33, 3935) = 23,4$ ;  $p < 0,001$  (hláskové tempo:  $F = 21,9$ ). Po provedení Tukeyho post-hoc testu bylo zjištěno, že ANOVA se závislou proměnnou slabičné tempo od sebe odliší 248 dvojic mluvčích z 561 možných, což odpovídá 44,2 % (tedy vyšší úspěšnost než trvání jednotlivých vokálů či konsonantů). Hláskové tempo má úspěšnost 41,4 % (232 dvojic).

Průměry slabičného tempa pro jednotlivé mluvčí jsou zobrazeny v grafech 8.87a (muži) a 8.87b (ženy).



Obrázky 8.87a-b: Průměrné hodnoty artikulačního tempa ve sl/s pro jednotlivé mluvčí, seřazené od nejnižších po nejvyšší. Mluvčí jsou rozděleni na muže (nahore) a ženy (dole).

Efekt mluvčího je významný v obou skupinách a pro obě měření. Muži, slabičné tempo:  $F(7, 927) = 19,3$ ;  $p < 0,001$ , hláskové tempo opět vykazuje nižší testové kritérium  $F = 17,8$ . Tukeyho post-hoc test pro slabičné tempo odlišil z celkového počtu 28 možných dvojic 15 (tedy 53,6 %), pro hláskové tempo 14 dvojic (50 %). Mluvčí MORC má signifikantně odlišné

slabičné tempo od všech ostatních mluvčích, dále jsou odlišní od většiny ostatních také mluvčí STOB a MIMK. Zbytek tvoří jakýsi poměrně homogenní průměr.

Žen je v našem materiálu mnohem více a rozdělení jejich průměrů je spojitější. Efekt mluvčího je významný pro slabičné tempo:  $F(25, 3008) = 22,7$ ;  $p < 0,001$  i pro hláskové tempo:  $F = 22,6$ . Slabičné tempo od sebe rozpoznalo 147 dvojic z 325 možných (45,2 %), hláskové tempo 143 dvojic (44 %).

Nejspecifičtější je mluvčí VASA, která se významně liší od všech ostatních mluvčích kromě RUBA a SEBA. Kromě těchto tří mluvčích se od nadpoloviční většiny ostatních žen signifikantně liší ještě LANA a STPA (rychlé mluvčí) a z pomalých mluvčích REBA a PCNA.

Kvůli prozkoumání rozdílů v artikulačním tempu v rámci jednoho mluvčího jsme jednotlivé repliky každého mluvčího náhodně rozdělili do dvou skupin a provedli t-testy pro nezávislá měření pro každého mluvčího zvlášť, tentokrát už pouze s proměnnou artikulační tempo ve slabikách za sekundu. Ukázalo se, že slabičné tempo je v rámci mluvčího dosti stabilní. Ze všech 34 mluvčích ukázal t-test významný rozdíl na hladině 0,05 pouze u mluvčích GROA, HOBA a STOA.

Prozkoumali jsme též, zda nehraje roli dvojice mluvčích, která spolu dialogy nahrávala – je možné, že mluvčí budou své řečové tempo navzájem přizpůsobovat. Rozdíl v rámci dvojice mezi slabičnými tempy obou mluvčích je průměrně 0,4 sl/s. Nejpodobnější je si dvojice ZELA a HEJA, které se liší pouze o 0,07 sl/s (zároveň se ale obě pohybují kolem průměru pro ženy). Největší rozdíl je mezi mluvčími VASA a GROA (0,91 sl/s), kdy GROA mluví průměrně rychle, zatímco VASA je nejrychlejší z mluvčích (vedle MORC). Korelace mezi průměrnými artikulačními tempy dvojic mluvčích je středně silná až silná,  $r = 0,7$ , a statisticky významná,  $p = 0,002$ . Seřadí-li se mluvčí do náhodně uspořádaných dvojic, je korelace průměrů pouze 0,1. Zdá se tedy, že mluvčí v dialogu se vzhledem ke svému artikulačnímu tempu navzájem částečně ovlivňují, akomodují.

### 9.3 Globální temporální ukazatele

Tyto ukazatele kvantifikují různými způsoby konsonantickou a vokalickou temporální variabilitu v rámci promluvy. V našem materiálu jsou jejich hodnoty spočítány na každé replice zvlášť (s vynecháním pauz a celých replik s dysfluencemi a hezitacemi). Stejně jako u artikulačního tempa ověřujeme jejich citlivost na faktory POHLAVÍ a MLUVČÍ.

V tabulce 8.23 jsou uvedeny výsledky jednofaktorových testů ANOVA pro každý naměřený globální temporální ukazatel.

Ukazatel	Faktor POHLAVÍ	Faktor MLUVČÍ
%V	$p < 0,001$ ; F < M	$p < 0,001$ ; 38,9 % (218)
$\Delta V$	$p = 0,025$ ; F > M	$p < 0,001$ ; 11,1 % (62)
$\Delta C$	nevýznamné	$p < 0,001$ ; 19,8 % (111)
VarcoV	$p < 0,001$ ; F > M	$p < 0,001$ ; 14,3 % (80)
VarcoC	$p < 0,001$ ; F < M	$p < 0,001$ ; 8,7 % (49)
rPVI-V	$p = 0,04$ ; F > M	$p < 0,001$ ; 2 % (11)
rPVI-C	nevýznamné	$p < 0,001$ ; 20,3 % (114)
nPVI-V	$p = 0,008$ ; F > M	$p < 0,001$ ; 2,1 % (12)
nPVI-C	$p = 0,02$ ; F < M	$p < 0,001$ ; 10,9 % (61)

Tabulka 8.23: Výsledky jednofaktorových testů ANOVA s faktory POHLAVÍ (druhý sloupec, F: ženy, M: muži) a MLUVČÍ (třetí sloupec). Ve třetím sloupci je uvedena úspěšnost testu jako procentuální podíl počtu dvojic, které Tukeyho post-hoc test označil za signifikantně odlišné (z 561 možných kombinací). V závorce následuje počet těchto dvojic.

Z devíti analyzovaných temporálních ukazatelů je sedm citlivých na pohlaví mluvčího, zdaleka nejvyšší významnost vykazuje VarcoC:  $F(1, 3967) = 42,3$ ;  $p < 0,001$ , vysoce významné jsou také %V:  $F(1, 3967) = 15,9$ ;  $p < 0,001$  a VarcoV:  $F(1, 3967) = 11,4$ ;  $p < 0,001$ . Naopak  $\Delta C$  a rPVI-C na tento faktor citlivé nejsou vůbec. Ve druhém sloupci je také uvedeno, zda mají ženy průměrně vyšší nebo nižší hodnoty než muži. Nižší %V u žen znamená, že vokalické intervaly zabírají průměrně menší část jejich promluv než u mužů. VarcoV mají ale vyšší, čili variabilita vokalických intervalů je naopak vyšší u žen. Variabilita konsonantických intervalů se ale podle výsledků VarcoC zdá být výrazně vyšší u mužů.

Rozdíly u indexů párové variability nejsou zcela snadno interpretovatelné, nicméně čím vyšší hodnota indexu, tím vyšší párová variabilita, tedy rozdíly v trvání po sobě následujících vokalických či konsonantických intervalů. Opět se zdá, že ženy vykazují vyšší variabilitu ve vokalických PVI a muži v konsonantickém nPVI-C.

Efekt mluvčího je sice významný pro všechny globální temporální ukazatele, nicméně z procentuálních úspěšností vidíme, že ne ve všech případech se jedná o ukazatel využitelný pro identifikaci. Nejúspěšnější je %V:  $F(33, 3935) = 23,4$ ;  $p < 0,001$ , které od sebe dokázalo navzájem odlišit 218 dvojic mluvčích (téměř 39 %). Nejdále od všech ostatních mluvčích přitom leží MIMK, jehož %V je zdaleka nejvyšší (50,5) a signifikantně se liší od všech

mluvčích. Od nadpoloviční většiny ostatních mluvčích se liší ještě BARA, LANA, MACK a OCEK (nízké hodnoty %V, kolem 39) a BRCA a BROZ (vyšší hodnoty %V, kolem 46).

Alespoň 100 dvojic mluvčích od sebe signifikantně rozlišil ještě ukazatel rPVI-C (úspěšnost 20,3 %),  $F(33, 3935) = 8,4$ ;  $p < 0,001$  a ukazatel  $\Delta C$  (úspěšnost 19,8 %),  $F(33, 3935) = 8,2$ ;  $p < 0,001$ . Tyto dva ukazatele zároveň nebyly citlivé na rozdíly mezi muži a ženami. Také je mezi nimi středně silná korelace,  $r = 0,68$ .

Pro ověření stability těchto tří slibných ukazatelů v rámci jednoho mluvčího jsme opět provedli sadu t-testů pro náhodně rozdělené repliky každého mluvčího do dvou skupin. %V ukázalo významný rozdíl mezi dvěma skupinami replik mluvčích JARS, MORC a ZETA,  $\Delta C$  pouze u mluvčí HROA a rPVI-C u mluvčí TROA. I tyto ukazatele jsou tedy forenzně využitelné, protože vykazují rozdíly mezi mluvčími a zároveň se chovají dosti stabilně v rámci replik jednoho mluvčího.

## 9.4 Diskuse k forezním analýzám

Z výše uvedeného vyplývá, že temporální charakteristiky mohou ukazovat na specifické rozdíly mezi mluvčími a vhodně tak doplňovat forezní analýzy v doméně spektrální. Je jasné, že čím více hlásek máme k dispozici, tím jednoznačnější jsou výsledky testů. Obecně se zdá, že v trvání vokálů se projevuje více individuální variability než v trvání konsonantů.

Artikulační tempo se rovněž ukázalo jako velice dobrý ukazatel identity mluvčího, protože vykazuje velkou variabilitu napříč subjekty a malou v rámci jednoho subjektu. Rovněž dle našich výsledků je slabičné tempo o něco lepším ukazatelem než tempo hláskové, které je pravděpodobně více ovlivněno hláskovou stavbou textu (jež byla pro všechny mluvčí stejná), zatímco na slabičném tempu se mohou více projevit individuální rytmické preference mluvčích.

V těchto analýzách jsme se rozhodli neměřit mluvní tempo jednotlivých mluvčích. Množství pauz v materiálu bylo nízké, takže jsme neočekávali výraznější rozdíly oproti tempu artikulačnímu. Rovněž by v takto krátkých dialozích s častým střídáním mluvčích nebyl prostor pro to, aby se projevilo individuální chování subjektů s ohledem na pauzy. Individualita pauz nicméně je rovněž předmětem výzkumu, viz např. pilotní studii pro češtinu Weingartové et al. (2014).

Co se týče globálních temporálních ukazatelů (neboli tzv. „rytmických ukazatelů“), zajímavým efektem, který se projevil v našem materiálu, byly rozdíly mezi muži a ženami. Ženy vykazovaly vyšší variabilitu vokalických intervalů, muži naopak vyšší variabilitu



konsonantických intervalů. Forezně nejúspěšnější ze zkoumaných ukazatelů bylo %V, rPVI-C a  $\Delta C$ . Tyto tři ukazatele jednak odhalují variabilitu mezi mluvčími, na druhou stranu jsou ale poměrně stabilní vzhledem k rozdílům v rámci mluvčího. Ovšem u rPVI-C a  $\Delta C$  je třeba obezřetnosti, jelikož oba ukazatele jsou nenormalizované, a tedy v sobě do určité míry přenášejí i rozdíly v artikulačním tempu.

Rozdíly v rámci mluvčích jsou pro forezní výzkum stejně důležitou oblastí jako rozdíly mezi mluvčími. Je potřeba vědět, do jaké míry jednotlivé parametry variiují v rámci promluv jednoho mluvčího. Tohoto tématu jsme se prozatím dotkli pouze okrajově, zasloužilo by si podrobnější pojednání.

V této práci rovněž nejsou zkoumány stylové změny – ty mohou vnést do materiálu ještě další variabilitu, i když z výsledků předchozího výzkumu se zdá, že přinejmenším artikulační tempo je vůči změnám stylu spíše rezistentní (Künzel, 1997; Trouvain, 2004; Mixdorff & Pfitzinger, 2005; pro češtinu Balkó, 2005 nebo Homolková, 2009). Nicméně i ve změně tempa v závislosti na mluvním stylu se dají nalézt individuální rozdíly (Balkó, 2005 nebo Jessen, 2007). Stylová variabilita temporálních charakteristik je další oblastí, která by měla být zevrubněji prozkoumána, také pro účely forezního využití, kde je často porovnávána spontánní řeč se čtenou (Svobodová & Voříšek, 2014).

## 10 Model temporálních charakteristik

Na základě získaných měření trvání hlásek a jeho modifikací v závislosti na různých faktorech byl vytvořen temporální model, který předpovídá trvání každé hlásky. Tento model je pravidlový a v principu podobný modelu Klattově (1976), liší se pouze tím, že neaplikuje minimální trvání („nestlačitelnou“ částí hlásky – viz kap. 3).

Vstupem do modelu je výchozí trvání jednotlivých hlásek uvedené v tabulce 9.1 níže, které je následně násobeno osmi různými koeficienty.

### 10.1 Výchozí hodnoty

Výchozí hodnoty trvání hlásek byly založeny na hodnotách z tabulek 8.5 a 8.18, nicméně tyto hodnoty bylo třeba pro některé hlásky opravit či pozměnit. Zejména jsme kontrolovali, zda vztahy mezi jednotlivými hláskami přibližně kopírují ostatní zdroje z literatury. Pro tuto korekci jsme vzali v úvahu tři zdroje, které měřily všechny (nebo téměř všechny) české hlásky, tedy Chlumského (1928), Borovičkovou a Maláče (1967), Psutku et al. (2006) – údaje pro čtenou řeč – a pro vokály hodnoty z Podlipského et al. (2009), jimiž jsme nahradili hodnoty Borovičkové a Maláče (1967). Porovnání jsme provedli vizuálně zobrazením hodnot trvání hlásek do spojnicových grafů. Pro tento účel byly hodnoty z Chlumského (1928) vynásobeny konstantou 0,7 a Borovičkové a Maláče (1967) 0,5, aby se pohybovaly ve stejných řádech a bylo možné je porovnávat.

Bylo zjištěno, že vztahům v porovnávaných zdrojích neodpovídají zejména sibilanty (kromě [d̥z] a [d̥ž], které měřili pouze Psutka et al., 2006), [d], [dʲ] a dlouhé vokály. Jejich hodnoty byly tedy upraveny. Vzali jsme vždy dvojici konsonantů (např [s]-[š]), spočítali poměr těchto dvou hodnot trvání pro každý ze tří starších zdrojů a následně je zprůměrovali. Hodnoty pro obě hlásky z této práce byly následně změněny o konstantu tak, aby jejich nový poměr odpovídal průměrnému poměru z literatury (vždy jedna hodnota byla snížena a druhá zvýšena). Dlouhé vokály byly upraveny podobným způsobem podle vztahů mezi krátkými vokály (ale jen z této práce, zde nebyla nutná korekce vzhledem ke starším zdrojům).

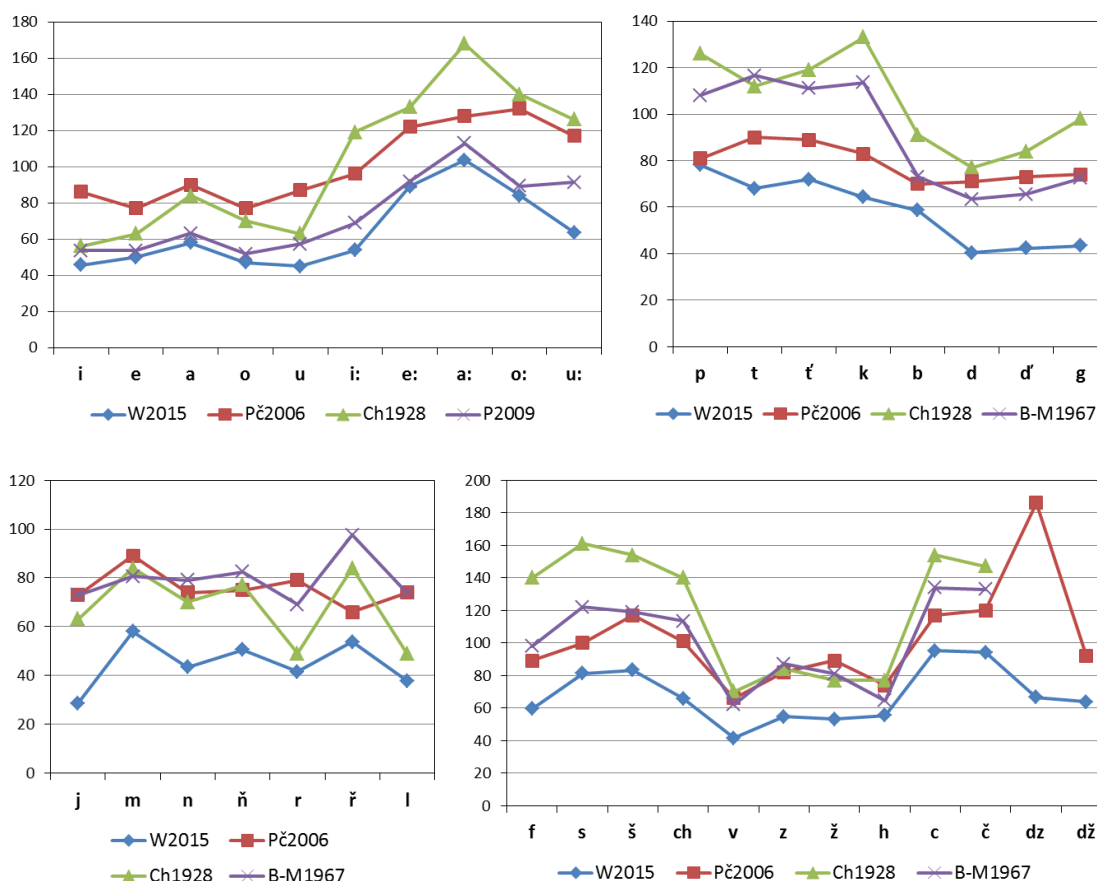
Korigované průměrné hodnoty jednotlivých českých hlásek jsou uvedeny v tabulce 9.1.

V tabulce nejsou uvedeny všechny hlásky, ale pouze ty, které byly měřeny ve starší literatuře. Jak je vidět, sonory nebylo třeba upravovat vůbec, jejich hodnoty odpovídaly hodnotám v citovaných zdrojích.

Vokály	Prům. trvání	Explozivy	Prům. trvání	Frikativy/ afrikáty	Prům. trvání	Sonory	Prům. trvání
i	45.6	p	78.1	f	59.6	m	58
e	49.9	t	68.1	s	<b>81.2</b>	n	43.4
a	58	ť	71.9	š	<b>83.3</b>	ň	50.5
o	47	k	64.4	ch	65.8	j	28.6
u	45	b	58.7	v	41.6	r	41.5
i:	<b>53.8</b>	d	<b>40.4</b>	z	<b>54.7</b>	ř	53.8
e:	<b>99.8</b>	d'	<b>42.4</b>	ž	<b>53.1</b>	l	37.7
a:	<b>103.6</b>	g	43.4	h	55.4		
o:	<b>84.0</b>			c	<b>95.2</b>		
u:	<b>63.7</b>			č	<b>94.2</b>		
				dz	66.7		
				dž	63.7		

Tabulka 9.1: Korigované průměrné hodnoty trvání jednotlivých hlásek. Tučně jsou označeny případy, které neodpovídají tabulkám 8.5 a 8.18. Korekce byla provedena na základě srovnání se studii Chlumského (1928), Borovičkové a Maláče (1967), Psutky et al. (2006) a Podlipského et al. (2009). Hodnoty jsou v milisekundách.

Grafy 9.1a-d zobrazují porovnání průměrných trvání všech hlásek vizuálně.



Obrázky 9.1a-d: Průměrné hodnoty trvání jednotlivých českých hlásek porovnávaných z pěti zdrojů – Chlumský (1928; Ch1928), Borovičková & Maláč (1967; B-M1967), Psutka et al. (2006; čtená řeč – Pč2006), Podlipský et al. (2009; P2009) a korigované hodnoty z této práce (W2015). Hodnoty z Chlumského jsou vynásobeny koeficientem 0,7 a z Borovičkové a Maláče koeficientem 0,5.

Po vytvoření a otestování několika prvních verzí modelu bylo nutné tyto výchozí hodnoty odladit, aby fungovaly na konkrétní materiál v našem korpusu. Výchozí hodnoty, které byly nakonec využity v poslední (deváté), nejpřesnější verzi temporálního modelu, jsou uvedeny v tabulce 9.2. Hodnoty z tabulky 9.1 by však měly být obecněji platné pro mluvenou češtinu.

Vokály	Prům. trvání	Explozivny	Prům. trvání	Frikativy/ afrikáty	Prům. trvání	Sonory	Prům. trvání
i	45.6	p	78.1	f	59.6	m	58
e	49.9	t	68.1	s	81.2	ŋ	92.7
a	58	ť	71.9	š	83.3	n	43.4
o	47	k	64.4	ch	65.8	ŋ	61.4
u	45	b	58.7	v	41.6	ň	50.5
i:	53.8	d	33.5	z	54.7	r	41.5
e:	99.8	d'	52.1	ž	53.1	r slab.	71.1
a:	103.6	g	43.4	h	55.4	ř	53.8
o:	84.0	?	63,8	c	87.9	l	37.7
u:	63.7			č	101.5	j	28.6
ou	80.3			dz	66.7		
				dž	63.7		

Tabulka 9.2: Výchozí hodnoty trvání jednotlivých hlásek (v milisekundách) využitých jako vstupní do temporálního modelu.

## 10.2 Koeficienty

Jednotlivé koeficienty byly vypočteny na základě hodnot uvedených v přechodných kapitolách této práce vydělených trváním výchozím. Vzorec pro výpočet daného koeficientu je tedy:

$$K_f = \frac{T_f}{T_v},$$

kde  $T_v$  je výchozí trvání hlásky z tabulky 9.2,  $T_f$  trvání hlásky pod vlivem faktoru  $f$  a  $K_f$  je výsledný koeficient faktoru  $f$ .

Pokud to bylo možné, využívali jsme hodnoty vypočítané ze seskupených dat (tedy stejně, jak je uvedeno vždy v příslušné části předchozí kapitoly). V některých případech ovšem dostatek dat nebyl k dispozici a koeficienty musely být extrapolovány (viz níže). Nastal-li případ, že daný faktor byl pro hlásku irelevantní, byl koeficient stanoven jako 1 (čili nijak nemění trvání hlásky). Ke stanovení koeficientu výše 1 při nedostatku dat jsme se uchýlovali jen výjimečně, vždy bylo primárním řešením zjistit extrapolaci koeficientu. Tabulka koeficientů a zdrojových hodnot je obsahem přílohy č. 5.

Při extrapolaci koeficientů byl důraz kladen na to, abychom daný faktor k extrapolaci nepoužívali. Čili například hodnota pro [ž] ve finálním slově prozodické fráze nebyla

dopočítávána z hodnoty mediálního [ž], ale z hodnot a vztahů pro ostatní hlásky ve finální pozici. Byly využity následující postupy, v uvedeném pořadí. K dalšímu postupu v pořadí bylo přistoupeno v případě, že nadřazený postup nebylo možné použít nebo vedl k neúměrně vysokým nebo naopak nízkým koeficientům.

Pro konsonanty:

- 1) Hodnota je dopočítána podle hodnoty pro znělostní protějšek a rozdílu mezi artikulačně nejbližšími hláskami.

*Příklad: Chybí hodnota pro [g] ve slabičné préture. Vezmeme tedy hodnotu pro znělostní protějšek [k] v préture a rozdíl mezi artikulačně nejbližšími explozivami v préture, což jsou [t] a [d]. Rozdíl mezi nimi činí 26,5 ms, za trvání [g] v préture tedy dosadíme trvání [k] minus 26,5 ms, což je 47,3 ms. Protože výchozí trvání [g] je 43,4 ms, koeficient pro trvání v préture bude  $47,3/43,4 = 1,09$ .*

- 2) Hodnota je dopočítána podle hodnoty pro znělostní protějšek a rozdílu mezi průměrnou hodnotou pro znělé a neznělé konsonanty.
- 3) Hodnota je dopočítána podle hodnoty pro znělostní protějšek a rozdílu mezi výchozími hodnotami pro tyto dvě hlásky.
- 4) Hodnota je dopočítána podle hodnoty artikulačně nejpříbuznější hlásky a rozdílu mezi výchozími hodnotami pro tyto dvě hlásky.

Dlouhé vokály byly extrapolovány podle krátkého protějšku a rozdílu ve výchozích hodnotách.

Krátké vokály bylo nutno extrapolovat jen ve dvou případech, kdy byly využity postupy obdobné pravidlům 1) a 3) – pouze nešlo o znělostní protějšek, ale o vokál sdílející stejnou charakteristiku horizontální polohy jazyka (přední či zadní). Extrapolované hodnoty spolu s pravidly použitými k jejich extrapolaci jsou k nalezení v příloze č. 5.

### 10.3 Modelované faktory

Rozhodli jsme se vůbec nemodelovat labiodentální nazálu [ŋ], jednak kvůli nedostatku dat, ale také vzhledem ke statusu hlásky v českém fonologickém systému.

Koeficient **pozice ve slabice** je aplikován pouze na konsonanty (vokály mají jedinou možnou hodnotu, a to *nukleus*). Extrapolován byl koeficient pro [g] v préture a [p], [b], [v] a [h] v kodě. Konsonanty, které se mohou vyskytovat pouze v kodě ([d̥z], [d̥ž] a [ŋ]), modelovány nebyly.

Koeficient **slovo v prozodické frázi** může nabývat tří hodnot, a to *iniciální*, *mediální* a *finální*. Slova tvořící samostatnou prozodickou frázi byla zahrnuta do kategorie *finální*. Pro vokály v iniciálních slovech nebyl v seskupených datech dostatek dokladů, rozhodli jsme se tedy koeficienty ponechat stejné jako ve slovech mediálních (vesměs se pohybovaly kolem 1). Bylo nutné dopočítat pouze mediální trvání [o:]. V případě konsonantů bylo dopočítáno iniciální trvání [b] a [dʰ], u frikativ a afrikát jsme iniciální opět ponechali stejné jako mediální, protože rozdíl mezi pozicemi nebyl statisticky významný (viz oddíly 8.3.3.1 a 8.3.4.1). U slabičného [r] a rázu jsme se rozhodli na pozicích s nedostatkem dat ponechat 1, protože citlivost těchto dvou hlásek na pozici v prozodické frázi se ukázala být nepatrná. Finální trvání bylo extrapolováno pro [g], [ž] a [x]. Zvláštním případem bylo [ŋ], [d͡z̥] a [d͡z̥]. Pro [ŋ] byl dostatek dokladů jen ve finálních slovech, vzali jsme tedy rozdíl mezi touto hodnotou a extrapolovanou hodnotou pro finální [g] a ostatní hodnoty jsme dopočítali podle hodnot [g] a tohoto rozdílu. Pro [d͡z̥] a [d͡z̥] byly dopočítány všechny tři koeficienty podle jejich znělostních protějšků.

Koeficient **finální slabika ve finálním slově** vyjadřuje o kolik se prodlouží hlásky v závěrečných slabikách finálních slov. Jeho hodnoty byly vztaženy k předchozímu faktoru tak, aby nedošlo ke dvojnásobnému násobení (a tím pádem i prodlužování) finálních hlásek. Téměř polovina hodnot tohoto koeficientu pro konsonanty musela být extrapolována, pro podrobnější postup odkazujeme na tabulku 5.3 v příloze. Pokud u nějakého koeficientu vyšla hodnota menší než 1, byla změněna na hodnotu 1, protože by jinak docházelo ke zkracování závěrečné slabiky.

Pouze pro konsonanty je relevantní koeficient **hláskové okolí**, který vyjadřuje změnu trvání podle toho, zda se hlásky nachází v konsonantickém shluku v rámci slabiky, mimo rámec slabiky, či je intervokalická. Nemodelujeme [r] a ráz, pro něž je tento faktor irelevantní a také [ŋ], kde máme nedostatek dat (navíc v češtině nemůže být intervokalické). Zde bylo extrapolováno mnoho hodnot, pro podrobnosti opět odkazujeme na přílohu č. 5 (tabulka 5.4).

Koeficient trvání **vokálů po rázu** byl spočten pouze pro krátké vokály, protože slova začínající na dlouhý vokál se v našem korpusu nevyskytují (jsou v češtině obecně vzácná). Také chyběl údaj pro trvání [e] po rázu, ten byl dopočítán podle trvání [a].

Dalším z faktorů byla **délka slova ve slabikách**. Rozhodli jsme se při počítání koeficientů pro délku slova i prozodické fráze použít vždy větší jednotku (tedy v případě slova slabiky, nikoliv fonémy a v případě prozodické fráze slova, nikoliv slabiky), protože by jinak faktor měl příliš mnoho úrovní a pro mnoho z nich bychom neměli dostatek dat. Modelovány jsou čtyři úrovně, jedno- až čtyřslabičná slova (tvoří 98,5 % všech případů). I tak nebyly v případě

mnoha konsonantů k dispozici hodnoty pro čtyřslabičná slova – tyto koeficienty pak nebyly extrapolovány, protože nebyl k dispozici dostatek hodnot pro jiné konsonanty, o které by se bylo možno opřít, a tedy jsme na dotčených místech byli nuceni ponechat 1. Pokud má slovo jiný počet slabik než 1 až 4, model hláskám automaticky přiřadí koeficient délky slova 1. (U neslabičných předložek mohl nastat i případ slova s počtem slabik 0.)

**Délka prozodické fráze ve slovech** funguje podobně, počítáme s úrovněmi 3 až 9, což pokrývá přibližně 85 % všech případů. Pokud má prozodická fráze jiný počet slov, model opět přiřazuje hláskám koeficient 1. Z důvodu nedostatku dat extrapolujeme vzhledem k tomuto faktoru hlásky [o:], [dž], [ŋ] a [r], u některých dalších hlásek jednotlivé hodnoty (viz příloha, tabulka 5.7).

Posledním modelovaným faktorem je **struktura slabiky**, která je vztažena pouze k trvání vokálů. Modelované úrovně jsou V, VC, CV, CVC, CCV, CCVC a CCCV (tyto slabiky tvoří přibližně 98 % materiálu). Má-li slabika jinou strukturu, je vokálu přiřazen koeficient 1. Při výpočtu koeficientů pro [ou] u posledních dvou faktorů byl zvolen následující postup – hodnota pro diftong byla vypočítána jako poměrná část součtu hodnot pro [o] a [u], poměr byl spočítán podle výchozích hodnot a činil 0,87.

Všechny koeficienty pro nedostatečně zastoupenou hlásku [ŋ] byly ponechány stejné jako pro [m] – tyto dvě hlásky se tedy v modelu liší pouze výchozím trváním.

Je třeba podotknout, že extrapolované koeficienty, ač jich je poměrně hodně, vznikly na místech, kde v materiálu nebyl dostatek dat. To tedy znamená, že při modelování korpusu budou tyto extrapolované koeficienty využity minimálně. Předpokládáme nicméně jejich užitečnost pro budoucí výzkum či pro aplikaci v praxi.

Modelová hodnota každé hlásky je tedy spočítána následovně:

$$T_m = T_v \times \prod_{f=1}^8 K_f,$$

kde  $T_m$  je modelové trvání hlásky,  $T_v$  je výchozí trvání hlásky a  $K_f$  je koeficient faktoru  $f$ .

## 10.4 Ladění modelu

Modelová trvání hlásek byla vypočítána na protříděných datech. Z korpusu jsme pro potřeby ladění i následného testování modelu vyřadili hezitace, dysfluence a finální slova před předělem hloubky 3.

Po výpočtu modelových hodnot bylo nutno post-hoc odladit některé z koeficientů pro jednotlivé hlásky. Byl zvolen následující postup: Ke každé reálné hlásce v korpusu byla spočítána odchylka  $O = T_r - T_m$ , kde  $T_r$  je reálné a  $T_m$  modelové trvání dané hlásky. Pak byla spočtena průměrná odchylka pro každou hlásku, a pokud se lišila od nuly více než 10 ms (tedy byla větší než 10 ms nebo menší než -10 ms), danou hlásku jsme prozkoumali a hledali příčinu této diskrepance.

Opraveny byly následující koeficienty (opravené hodnoty jsou i s komentářem také součástí přílohy č. 5):

Slovo v prozodické frázi – modelové trvání pro iniciální [h] se ukázalo být moc krátké. Iniciální [h] (jejichž koeficient byl původně stanoven stejný jako mediální) jsou v našem materiálu průměrně o 10 ms delší než mediální. Koeficient byl podle toho upraven.

Finální slabika ve finálním slově – přepočítán byl pro hlásky [a:], [ou], [t], [d] a [f]. V případě [a:] a [t] byl koeficient příliš nízký a předpovídal krátké trvání ve srovnání s reálným. Tento efekt byl pravděpodobně způsoben tím, že v původním výpočtu koeficientu nebyla zahrnuta individuální slova v prozodické frázi. U ostatních hlásek byl naopak koeficient příliš vysoký a modelová trvání byla delší než reálná. V obou případech byl koeficient přepočítán ze všech neseskupených dat.

Hláskové okolí – bylo třeba přepočítat koeficienty pro [p] ve všech třech pozicích, [t] v neintervokálních pozicích a [s] a [z] v pozici mimo shluk. Ve všech čtyřech případech předpovídal model příliš krátká trvání, koeficienty byly tedy přepočítány ze všech neseskupených dat, čímž se o něco zvýšily.

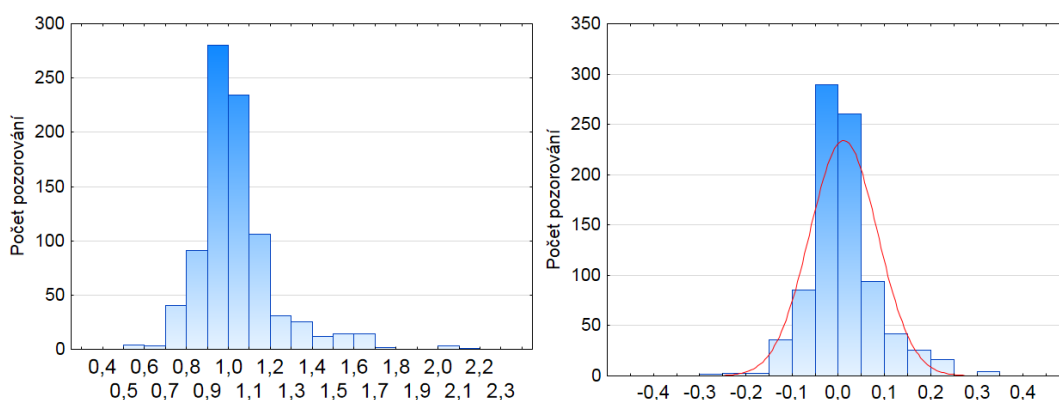
U koeficientu délky slova bylo třeba přepočítat některé z hodnot pro [p], [t], [g] a [f]. Kromě [g] byly všechny koeficienty snižovány, u [g] byl pro dvouslabičná slova naopak zvyšován.

U vokálu [u:] došlo k situaci, kdy hodnota pro tento vokál v šestislovných frázích byla identická pro hodnotu v CVC slabikách (výskyty [u:] pro tyto dva faktory byly totožné), a trvání vokálu bylo tedy násobeno dvakrát tímtéž číslem. Post-hoc jsme tedy koeficient pro CVC stanovili na 1.

Odladěný model má průměrnou odchylku od reálných hodnot 0,9 ms a průměr z absolutních hodnot odchylek je 16,54 ms. Obsahuje 860 koeficientů, které se nerovnjí jedné. Z toho téměř 60 % (514) spadá mezi 0,9 a 1,1 včetně. Celkově se hodnoty koeficientů pohybují mezi 0,5 a 2,1, což odpovídá zhruba polovičnímu a dvojnásobnému trvání. Lze tedy říci, že modifikace trvání (zkracování a prodlužování) je poměrově symetrická – dělení dvěma vs. násobení dvěma.



Graf 9.2a zobrazuje histogram hodnot koeficientů, graf 9.2b pak jejich dekadický logaritmus. Rozdělení je dosti špičaté, což odpovídá tomu, že většina koeficientů se skutečně pohybuje kolem 1. Vidíme, že rozdělení koeficientů je po zlogaritmování skutečně symetrické.



Obrázek 9.2a-b: Histogramy rozdělení hodnot koeficientů v temporálním modelu. Vlevo reálné hodnoty, vpravo zlogaritmované ( $\log_{10}$ ), červená čára znázorňuje křivku normálního rozdělení.

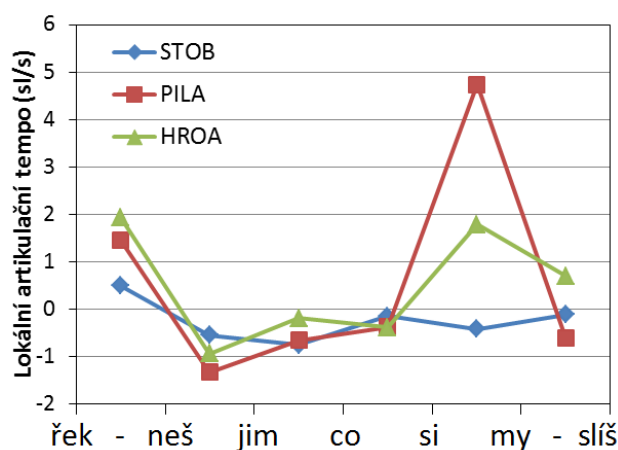
## 10.5 Modelové kontury LAR a individuální rezidua

Modelová trvání nyní použijeme k výpočtu modelových kontur lokálního artikulačního tempa (LAR; Volín, 2009), protože hodnoty vypočtené tímto temporálním modelem nejsou závislé na mluvčím, bude to znamenat, že daná prozodická fráze bude mít vždy jen jedinou modelovou konturu. Výpočet kontur LAR je popsán v sekci 7.6, zopakujme jen, že jde o převrácenou hodnotu vzdálenosti onsetů dvou slabičných jader. Pro porovnávání vybíráme z korpusu Minidialogy-H jen takové repliky, kde mluvčí stejně frázují, aby byla jejich reálná trvání přímo porovnatelná s modelovými a všichni byli vztáženi ke stejnému modelu.

Pro účely analýz specifík mluvčích byly vybrány repliky číslo 3, tzn. prostřední v každém dialogu. Tyto repliky vždy obsahují cílovou větu, lišící se ve čtveřicích dialogů pouze tvarem přísudku a modalitou (viz Metoda, oddíl 7.1). Kromě stejného frázování bylo ještě nutné zajistit, že každý mluvčí vyslovil stejný počet stejných hlásek (jinak by rozdíly v hodnotách LAR byly způsobeny elizemi, případně rozdíly v asimiliacích). Repliky s elidovanými hláskami a nestandardními asimilacemi byly tedy vyřazeny. Elize v préture první slabiky (např. [j] ve slově *zjistíš/zjistíte*, ve čtveřici H3) nemají na hodnoty LAR žádný vliv, tedy byly v materiálu ponechány. Také jsme ponechali repliky s desonorizovanými konsonanty, na základě úvahy, že fonologicky jde o stejnou hlásku, lišící se pouze v konkrétní realizaci, a tedy jde o specifikum mluvčího, které by se mohlo projevit v temporální doméně. Ovšem k většině desonorizací docházelo na začátku fráze, což je pro výpočet LAR irelevantní.

Analyzujeme celkem 645 vět. Je třeba rovněž poznamenat, že náš model nebere v úvahu rozdíly mezi otázkami a oznamovacími větami, tedy modelová trvání jsou pro obě modalities stejná.

Pro každou větu byla vypočítána jedna modelová kontura LAR a pro jednotlivé mluvčí pak odchylky v realizaci všech hodnot od hodnot modelových, tj. rezidua modelu. Rezidua, stejně jako hodnoty LAR, jsou uvedena ve slabikách za sekundu. Je nutné mít na paměti, že jedna hodnota LAR neodpovídá konkrétní slabice ani vokálu, ale lokálnímu artikulačnímu tempu mezi dvěma sousedícími vokály. Slabičné konsonanty se ve vybraných replikách nevyskytují.



Obrázek 9.3: Příklady kontur reziduí LAR z věty H1a\_3, *Řekneš jim, co si myslíš*, mluvčích HROA, PILA a STOB. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Příklad tří kontur reziduí uvádí obrázek 9.3. Jeden bod v grafu tedy odpovídá rozdílu reálné a modelové hodnoty LAR. Je-li bod umístěn blízko nule, znamená to, že trvání reálných hlásek použitých pro výpočet LAR je podobné trvání modelovému. Liší-li se směrem nahoru, mluvčí hlásky zkracuje a tempo se zvyšuje, naopak hodnoty záporné ukazují na prodlužování hlásek a zpomalování lokálního artikulačního tempa. Mluvčí STOB v grafu 9.3 má jen malé odchylky od modelového trvání, na rozdíl od mluvčích PILA a HROA, které na začátku a na konci zkoumané repliky zrychlují, a na druhé pozici naopak zpomalují. Postup výpočtu uvedeme na příkladu hodnoty 4,7 sl/s pro mluvčí PILA na páté pozici. Tato hodnota je vypočítána tak, že od převrácené hodnoty vzdálenosti onsetů obou krátkých [i] ve slabikách *si my* (vzdálenost činí 71 ms,  $1/0,071 = 14,1$ ), byla odečtena tatáž hodnota modelová (107 ms, tedy  $1/0,107 = 9,4$ ).

Pokud se všichni mluvčí liší od modelového trvání jedním směrem, jde o systematickou odchylku od modelu, která je způsobena faktorem, jenž model nevzal v úvahu. Forezně nejzajímavější jsou nicméně případy, kde se mluvčí podstatně liší od sebe navzájem. Pro

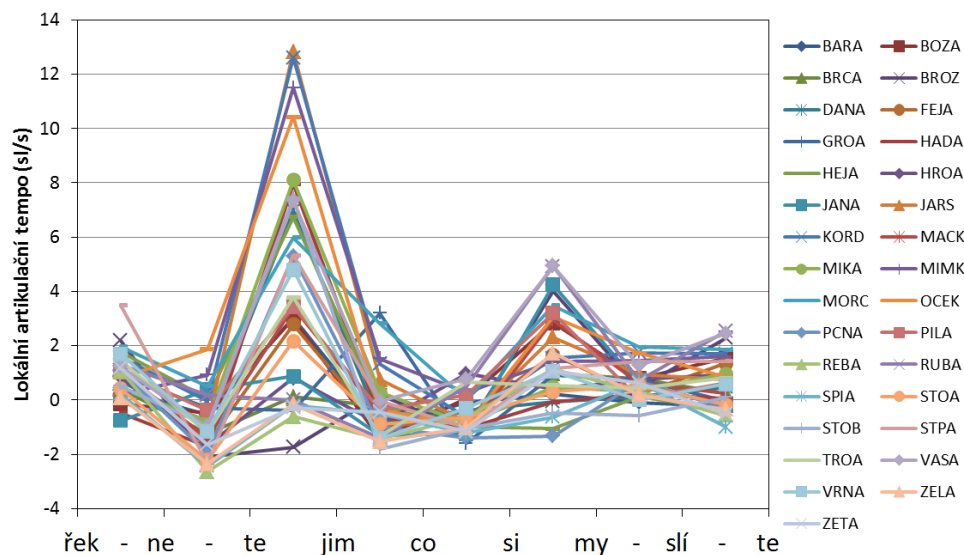
každou pozici LAR tedy počítáme ukazatele variability – variační rozpětí (tj. rozdíl maximální a minimální hodnoty) a percentilové rozpětí (10.–90. percentil a 20.–80. percentil). Variační rozpětí je citlivé na extrémní hodnoty, které jsou zajímavé pro určení pozice s největší variabilitou, percentilová rozpětí pak odhalí i zdánlivě variabilní případy, kde ale extrémní hodnotu tvoří například jen jeden mluvčí a všichni ostatní jsou shluknutí u sebe. Grafy všech zkoumaných replik a mluvčích jsou obsahem Přílohy 6.

Repliky H1a\_3, H1b\_3, H1c\_3 a H1d\_3 mají tvar *Řekneš jim, co si myslíš*, respektive *Řeknete jim, co si myslíte*. V kratších replikách (tedy H1a\_3 a H1b\_3) je variabilita mezi mluvčími obecně menší. Absolutní odchylky se pohybují kolem 0,9 sl/s pro kratší větu a 1,5 sl/s pro delší větu. Není to způsobeno pouze jednou pozicí v kontuře, obecně jsou ve větách c a d rozdíly na všech pozicích větší. U krátkých vět nacházíme největší rozdíly mezi mluvčími na pozicích *řek-neš* a *si-my*. Ve druhém případě se objevují tři mluvčí, kteří na této pozici extrémně zrychlují (více než dvojnásobně oproti ostatním). Nicméně i variabilita mezi ostatními je stále vysoká. O něco málo větší průměrnou variabilitu vykazuje přitom oznamovací věta. Otázka (tedy replika H1b\_3) vykazuje menší variabilitu zejména na čtvrté pozici, tedy *jim-co*, což může být způsobeno menší pravděpodobností výraznějšího prozodického předělu. Je možné, že v oznamovací větě někteří mluvčí častěji na této pozici zdůrazňují taktový předěl, který se pak odráží ve variabilitě lokálního tempa.

V delších replikách (tj. H1c\_3 a H1d\_3; *Řeknete jim, co si myslíte*) je na první pohled zřetelná masivní variabilita na pozici *te-jim*. U některých se lokální tempo drží průměru, u části výrazně stoupá (jak mluvčí hlásku [j] redukuje) až na 16 sl/s. Zásadní je, že nejde o pár osamocemných extrémů, ale o kontinuum, v jehož rámci se mluvčí rozprostřeli. Věta H1d\_3 je vidět na obrázku 9.4.

Druhý největší rozdíl se pak nachází mezi slabikami *si-my*, třetí pak u *řek-ne*. Ty jsou řádově podobně velké jako u kratších replik.

Naopak nejstabilnější jsou v prvních dvou větách hodnoty odpovídající rozdílům mezi slabikami *neš-jim* a *my-slíš*. Ve druhých dvou větách je to pak rozdíl *my-slí* a *řek-ne*. Zajímavé je, že v řetězci *neš-jim* se také vyskytuje [j], ale zřejmě v sousedství [š] je jeho variabilita kompenzována stabilitou sykavky.



Obrázek 9.4: Graf kontur reziduí lokálního artikulačního tempa pro větu H1d\_3, *Řeknete jim, co si myslíte?* Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

V replice H1a\_3 vybočují především mluvčí BOZA a PILA, které na pozici *si-my* extrémně zrychlují. U mluvčí BOZA je to způsobeno redukcí vokálu [ɪ] v *si*, která podle reziduí LAR není kompenzována následujícím [m] ani předcházejícím [s], které mají trvání téměř stejné jako model. Mluvčí PILA tohoto výsledku dosahuje redukcí [m], které je spíše aproximantou. Idiosynkraticky se chová také mluvčí RUBA, jejíž bod největšího zrychlení se nachází na pozici *co-si* vysoko nad ostatními díky redukovanému [s], které je téměř prodloužením předchozího [c].

V rámci repliky H1b\_3 se jinak než ostatní chová především mluvčí VASA, jež podstatně zrychluje na celé části *co si myslíš*. Naopak téměř všechny body pod modelovým trváním má mluvčí HADA, zejména se od zbytku liší opět v části *co si myslíš*, kde realizuje výrazný stoupavý melodém (v kontrastu s mluvčí VASA, která stoupá jen nepatrně) a slovo *myslíš* je slyšitelně prominentní.

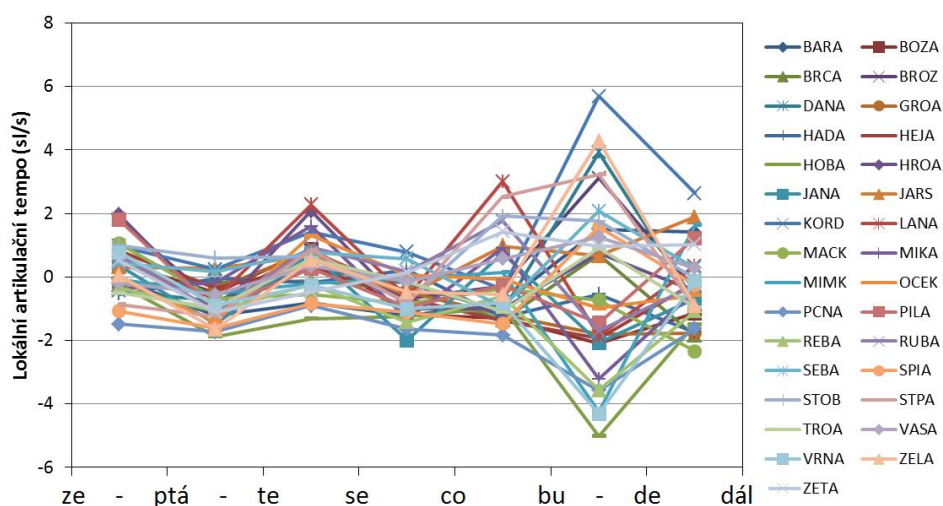
Na pozici *te-jim* (ze slov *řeknete jim* v replice H1c\_3) zdaleka nejvíce zrychlují mluvčí STOB a SEBA, výrazně na tomto místě zpomaluje mluvčí BROZ, což je doprovázeno poměrně výrazným stoupaním F0 a intenzity v tomto místě a navozuje tak percepční dojem drobného prozodického předělu mezi *řeknete jim* a *co si myslíte*.

Na stejném místě ve větě H1d\_3 je signifikantně rychlejších více mluvčích (viz obr. 9.4), zejména JARS, KORD, MIMK a OCEK. Kontrastují zejména s mluvčím BROZ, který je téměř 2 sl/s pod modelovým trváním. U čtyř zmíněných nejrychlejších mluvčích by se dala celá část *řeknete jim, co si* považovat za jeden šestislabičný takt, zatímco mluvčí BROZ

realizuje zřetelný taktový předěl – slabiky *te jim* jsou prominentnější (v kontrastu s okolními), s vyšší F0 a intenzitou. Mluvčí JARS, KORD, MIMK a OCEK naopak realizují tyto slabiky bez melodického pohybu a s podobnou intenzitou jako následující *co si*.

Věty v dialozích H2 znějí *Zeptáš se, co bude dál* (H2a\_3 a H2b\_3) a *Zeptáte se, co bude dál* (H2c\_3 a H2d\_3). Zde nelze pozorovat žádný velký rozdíl v průměrné variabilitě mezi prvními a druhými dvěma větami (nejvyšší průměrnou variabilitu má H2d\_3). Nejzásadnější rozdíly se ve všech čtyřech případech nacházejí uvnitř a kolem slova *bude*, kde se mluvčí navzájem liší na obě strany od nuly – někteří zpomalují, jiní zrychlují, což je způsobeno jednak různým zacházením s hláskou [d], jednak také celkovým zrychlením nebo zpomalením na posledních třech slovech (*co bude dál*).

Všechny čtyři zkoumané věty mají první tři (respektive čtyři) pozice LAR velmi stabilní, variabilita se soustředí až na poslední tři slova. Nejstabilnější jsou přitom první dvě pozice (tj. *ze-ptáš* a *ptáš-se*, respektive *ze-ptá* a *ptá-te*). Větu H2c\_3 ukazuje obrázek 9.5.



Obrázek 9.5: Graf kontur reziduí lokálního artikulačního tempa pro větu H2c\_3, *Zeptáte se, co bude dál*. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

V této čtveřici mají oznamovací věty nižší průměrnou variabilitu, nicméně rozdíl není příliš velký. Jiné systematické rozdíly mezi otázkou a oznamovací větou nebyly nalezeny.

V replice H2a\_3 nacházíme tři mluvčí, výrazně se lišící od ostatních. Mluvčí VASA nejvýrazněji zrychluje na slabikách *bude*, přičemž [d] realizuje jako alveolární švih. Výsledkem je percepční dojem pouze jediného vrcholu na slove *bude*, které díky tomu budí dojem jediné slabiky. Mluvčí BOZA na tomtéž místě podstatně zpomaluje a vyslovuje velmi pečlivě obě slabiky. Vyvolává percepční dojem hyperkorektnosti, nebo až lehké

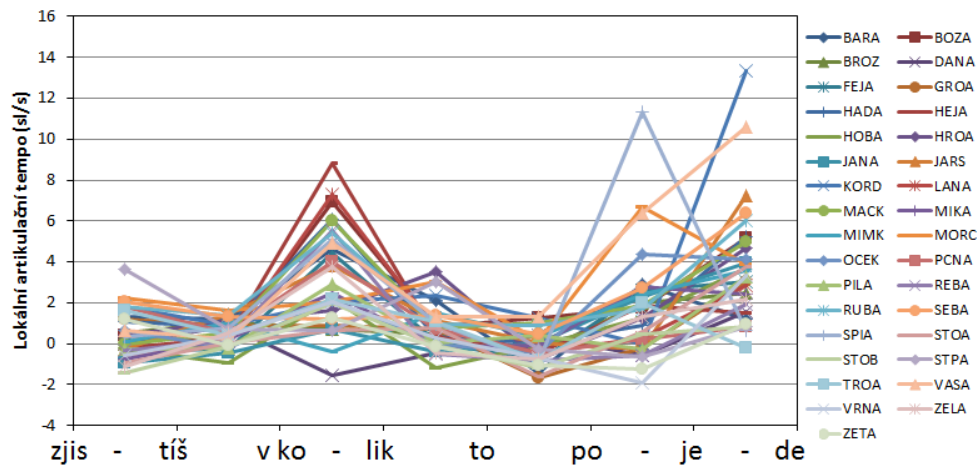
dysfluence. Třetí zvláštním případem je mluvčí JARS, který zrychluje na pozici *co-bu*, což je zapříčiněno redukcí [o]. Percepční dojem z rytmu mluvčích JARS a VASA je ovšem velmi podobný.

Mluvčí VASA zrychluje na stejném místě také v replice H2b\_3. Zpomalují zde naopak mluvčí HEJA a HOBA. Pro tyto případy platí totéž, co bylo napsáno v předchozím odstavci o mluvčích VASA a BOZA, obzvláště obě realizace mluvčí VASA jsou téměř identické, a to i přes to, že obě věty se vyskytují v různém kontextu a při nahrávání byly odděleny řadou jiných minidialogů.

Totéž lze říci i o replice H2c\_3, v níž na tomtéž místě nejvíce zrychluje mluvčí KORD a nejvíce zpomaluje mluvčí HOBA (viz obrázek 9.5). Mluvčí HOBA se podobně chová i v replice čtvrté, tedy H2d\_3, v níž opět nejvíce zrychluje mluvčí VASA. Také realizace mluvčího JARS vykazuje stejnou odlišnost jako v případě repliky H2a\_3.

Čtveřice replik H3a\_3, H3b\_3, H3c\_3 a H3d\_3 obsahuje věty *Zjistíš, v kolik to pojede* a *Zjistíte, v kolik to pojede*. Ve všech čtyřech větách se nacházejí dvě pozice s největší variabilitou, a to mezi slabikami *ko-lik* a *je-de*. Na obou pozicích dochází u části mluvčích ke zrychlování (rozdíl oproti modelu je i více než 6 sl/s), zbylá část se drží kolem modelového trvání, a v malé části případů dochází ke zpomalování. Stejně jako v případě předchozích vět nejde o pár extrémů, ale o kontinuum, ve kterém se mluvčí nacházejí na různých místech. Zdá se přitom, že mluvčí mohou využít možnost zrychlit v jednom případě, ale také v obou najednou nebo v žádném. Třetí nejvyšší variabilita (i když již ne tak výrazná) se objevuje mezi předposledními slabikami *po-je*. Stabilní body jsou konzistentně na pozicích *tíš-v ko*, *te-v ko* a *to-po*, u replik a a c také *zjis-tí(š)*. Průměrně mají obě oznamovací věty poněkud menší variabilitu než otázky.

Na pozici *ko-lik* v replikách H3a\_3 a H3b\_3 nejvíce zrychluje mluvčí HEJA, jejíž [o] je velmi krátké a redukované na [ə], následující [l] a [ɪ] jsou taktéž dost krátké. Obě repliky znějí co do rytmu velmi podobně a dochází zde ke slyšitelné taktové izochronii – obě realizace dvouslabičného taktu *zjistíš* jsou jen o pár milisekund jinak dlouhé než obě realizace taktů *v kolik to* (kolem 350 ms). Takt *pojede* je pak v obou případech delší, zhruba 430 ms.

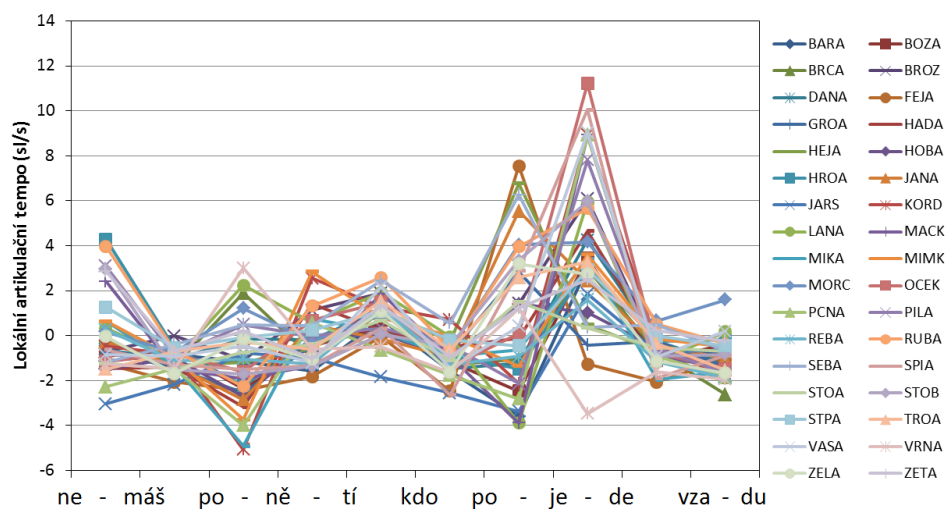


Obrázek 9.6: Graf kontur reziduí lokálního artikulačního tempa pro větu H3b\_3, *Zjistíš, v kolik to pojedede?* Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Mluvčí MIMK na slově *ko-lik* zpomaluje ve všech čtyřech replikách. Především v první čtveřici vytváří stoupnutím F0 a intenzity výraznou prominenci na slabice *ko*, kterou také protahuje. Mluvčí MORC a MIKA, kteří oba zrychlují ve slově *pojedede* (repliky H3a\_3), silně redukují [j] i [d] a celé slovo tak budí dojem dvouslabičného. Zajímavě se také chová mluvčí BROZ, který v replice H3a\_3 nezrychluje na *ko-lik*, nýbrž na následující pozici *lik-to*. Oproti ostatním mluvčím neredukuje [o] v první slabice, ale [i] a následující [k] je s nevypuštěným závěrem. V replice H3b\_3 se idiosynkraticky chová mluvčí SPIA, která zrychluje na pozici *po-je* (viz obr. 9.6). [j] je téměř elidováno a obě slabiky tvoří jeden intenzitní vrchol. Dojem dvou slabik je zachován díky výraznému pohybu F0 směrem vzhůru. Na poslední pozici repliky H3b\_3 spolu kontrastují mluvčí KORD a TROA. KORD vyslovuje [d] jako krátký alveolární švih a výrazně zrychluje na posledních dvou slabikách. Naopak mluvčí TROA mnohem více zpomaluje. Rozdíl mezi nimi spočívá také v tom, že TROA stoupá melodii již na předposlední slabice *je*, zatímco KORD až na poslední slabice *de*. V této replice je také zajímavá mluvčí DANA, která zpomaluje na pozici *ko-lik* – velmi pečlivě vyslovuje [l] a neredukuje ani jeden z vokálů.

V případě replik H3c\_3 a H3d\_3 na konci výrazně nejvíce zrychluje mluvčí MORC. V otázce můžeme pozorovat stejný melodický kontrast, kdy MORC (a také druhý nejrychlejší KORD) stoupá až na poslední slabice, zatímco mluvčí, kteří mají nízkou hodnotu LAR, již na předposlední, případně ještě o slabiku dříve (např. REBA). Na pozici *ko-lik* je v obou replikách nejrychlejší mluvčí LANA (která je rovněž jednou z nejrychlejších v replikách H3a\_3 a H3b\_3).

Čtveřice replik H4a\_3, H4b\_3, H4c\_3 a H4d\_3 s větami *Nemáš poněti, kdo pojede vzadu* a *Nemáte poněti, kdo pojede vzadu*, je nejdelší z celého korpusu (11, respektive 12 slabik). Přesto zde (především ve větách a a b) velká část mluvčích stále ještě realizovala celou repliku jako jedinou prozodickou frázi. Pozice s největší variabilitou kontury LAR se nacházejí (stejně jako u H3) ve slově *pojede* a o něco menší mezi slabikami *po-ně* ve slově *poněti*. Na slově *pojede* mluvčí převážně zrychlují, v extrémních případech o 9–11 sl/s oproti modelu. Ve slově *poněti* dochází u velké většiny mluvčích ke zpomalení, což může signalizovat důraz, který se na tomto slově často objevuje. Nejstabilnější body kontur nejsou opět zcela konzistentní, nejčastěji jde o pozici *de-vza* a u prvních dvou vět také *máš-po*. Na obr. 9.7, na příkladu repliky H4a\_3, jsou zejména body největší variability zřetelně vidět.



Obrázek 9.7: Graf kontur reziduí lokálního artikulačního tempa pro větu H4a\_3, *Nemáš poněti, kdo pojede vzadu*. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Repliky H4a a b mají obecně nejvyšší průměrnou variabilitu odchylek ze všech. H4b a c vykazují o něco nižší, to je ale způsobeno tím, že je zde k dispozici menší počet mluvčích (17, respektive 8); jak již bylo zmíněno výše, mluvčí častěji dělili repliku na dvě prozodické fráze. Z obr. 9.7 je také vidět, že část mluvčích zrychluje pouze na slabikách *po-je*, část na *je-de* a třetí skupina na obou. Extrémem, který zrychluje na prvních dvou slabikách slova *pojede*, jsou v replice H4a\_3 mluvčí FEJA, HEJA a SEBA. Na druhých dvou slabikách pak nejvíce zrychlují OCEK a SPIA. V kontrastu k nim stojí mluvčí MACK, LANA a VRNA, kteří zde naopak výrazně zpomalují. Ve slově *poněti* nejvíce zpomalují KORD a MIKA a naopak zrychluje zde VRNA. Lokální zpomalení v tomto slově je důsledkem větného důrazu.

V replice H4b\_3 jsou velmi zajímavá rezidua mluvčího MORC, který extrémně zrychluje na *po-ně* i *po-je* zároveň. Celá replika díky tomu zní překotně. Odchyluje se také mluvčí STPA,



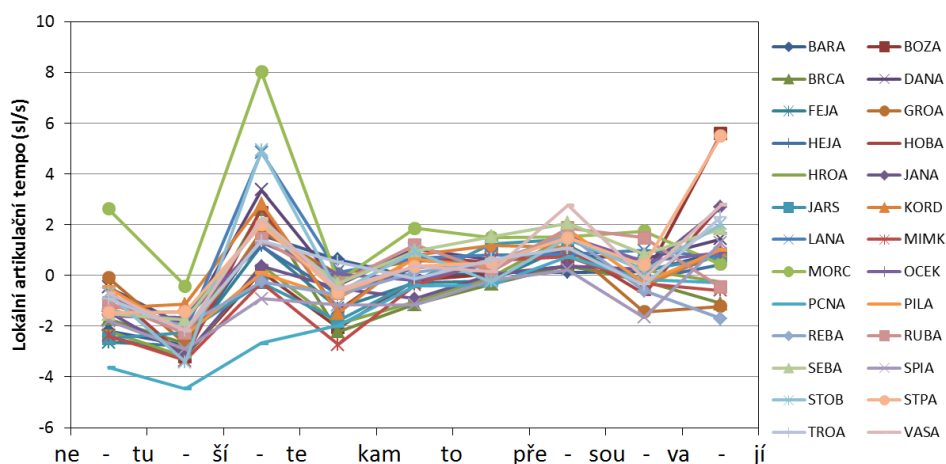
vrchol jejíchž reziduí je mezi slabikami *ně-tí*. Je to způsobeno redukcí [e], které se stává součástí [ň], a slovo *poněti* pak budí dojem dvouslabičnosti. Na pozici *je-de* je nerychlejší mluvčí SPIA, která podstatně redukuje [j] i následující [e] a opět výsledný dojem je dvouslabičný.

V replice H4c\_3 se bod největší variability jednoznačně nachází ve slově *pojede*, naopak ve slově *poněti* téměř všichni mluvčí zpomalují, což opět vytváří lehký, nicméně slyšitelný dojem zdůraznění. Nejodlišnější svým zrychlováním jsou mluvčí JARS, VASA a HADA, kteří silně redukují [j] i [d]. V případě repliky H4d\_3, kde máme k dispozici jen sedm mluvčích, je odlišná mluvčí RUBA, která výrazně zpomaluje slovo *poněti* a naopak zrychluje *pojede*. To vytváří dojem lehkého prozodického předělu za slovem *poněti* a druhá část věty působí překotně. Mohlo by jít o tendenci k izochronii, protože obě části věty jsou díky tomu velmi podobně dlouhé (690 a 670 ms).

Co se týče rozdílu mezi oznamovacími větami a otázkami, průměrná variabilita je podobná, ale v oznamovacích větách je častější výrazné zpomalení na slově *poněti* – zdůraznění tohoto slova se totiž v otázce může vyskytnout jen těžko.

Repliky ze čtveřice H5, *Netušíš, kam to přesouvají* a *Netušíte, kam to přesouvají*, mají poměrně nízkou průměrnou variabilitu, je srovnatelná s H2 a jen o něco vyšší než u H1. Obě otázky jsou průměrně o něco méně variabilní než oznamovací věty. Mluvčí ve všech čtyřech větách vykazují ve slově *netušíš* a *netušíte* hodnoty LAR nižší než modelové (často až o 3 sl/s), což je důsledkem větného důrazu na tomto slově, který realizuje většina mluvčích.

Ve větách H5a\_3 a H5b\_3 jsou pozicí s největší variabilitou poslední dvě slabiky, *va-jí*. U otázky (H5b\_3) je velká variabilita také ve slově *netušíš*, protože stejně jako v případě H4 je zde větný důraz méně pravděpodobný, část mluvčích zde tedy zrychluje. Nejstabilnějším bodem H5a\_3 je pozice *pře-sou*, u H5b\_3 je o něco stabilnější ještě předcházející pozice *to-pře*. U delších replik, H5c\_3 a H5d\_3 se pozice s největší variabilitou přesunula na *ší-te* (ze slova *netušíte*) a je řádově vyšší než v obou předchozích větách. Především ve větě c je ovšem poslední pozice stále ještě dost variabilní (viz obr. 9.8).



Obrázek 9.8: Graf kontur reziduí lokálního artikulačního tempa pro větu H5c\_3, *Netušíte, kam to přesouvají.* Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

V H5a\_3 je zajímavá mluvčí HADA, jejíž kontura začíná nízko (silný větný důraz na slově *netušíš*) a stoupá až do konce, kde dosahuje jedné z nejvyšších odchylek. Zrychlení je jasně slyšitelné. Dále zde nacházíme dvě lokální maxima mluvčích RUBA a VASA, na slabikách *kam-to*, respektive *pře-sou*. Tyto individuální změny rytmu jsou také dobře slyšitelné. U mluvčí SPIA, která je jednou z nejpomalejších na konci, dochází k podstatnému frázovému zpomalování, nejde o důraz na slově *přesouvají*.

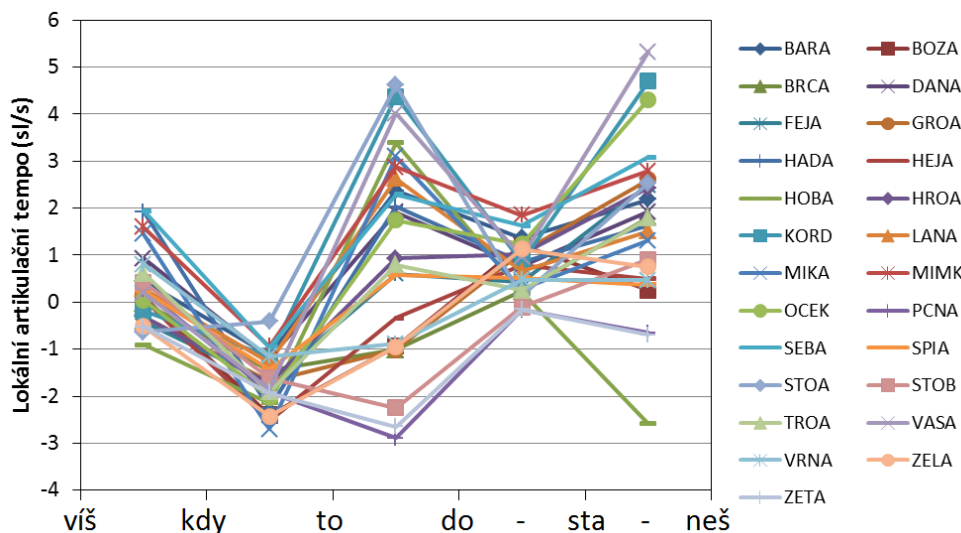
Věta H5b\_3 je obecně poměrně málo variabilní, mluvčí se kromě začátku a konce drží ve shluku. Mluvčí MIMK a BROZ extrémně zrychlují na prvních dvou slabikách, ovšem i přes to nejsou obě repliky poslechově ničím výjimečné, všechny hlásky jsou zřetelně vysloveny. Naopak mluvčí SPIA, která na začátku i na konci výrazně zpomaluje, působí otráveně či rezignovaně. Mluvčí MORC na rozdíl od ostatních zrychluje také v poslední slabice slova *netušíš*, jeho [i:] je přibližně stejně dlouhé jako okolní krátké vokály. Na posledních třech slabikách zrychlují mluvčí RUBA, SEBA a STPA, ovšem každá jiným způsobem, RUBA má vysokou odchylku na pozici *sou-va*, na následující už je hodnotou blízko průměru, STPA se chová přesně naopak a SEBA má výrazně vysoké obě hodnoty. RUBA redukuje [ou̯] na [u], STPA zřetelně zkracuje poslední dvě slabiky a u mluvčí SEBA je slyšitelné celkové zrychlování směrem ke konci věty. BRCA, jež má nejpomalejší poslední hodnotu, zde realizuje výrazný stoupavý melodém, který končí velmi vysoko.

Mluvčí MORC je nejodchylnějším případem ve větě H5c\_5, zřetelně zrychluje na prvních třech pozicích (viz obr. 9.8). Obě [t] v prvním slově jsou redukována na alveolární švih a celá jeho promluva zní překotně. Oproti tomu je na začátku nejpomalejší mluvčí PCNA, která na prvním slově realizuje výrazný větný důraz, který je kromě trvání signalizován také

melodickým pohybem a velkým intenzitním kontrastem vůči zbytku věty. Zrychlení na konci, které je nejvýraznější u mluvčích BOZA a STPA, je u nich doprovázeno úbytkem intenzity a výrazným klesnutím F0 až do třepené fonace. Oproti tomu mluvčí GROA a REBA, které zde nejvíce zpomalují, přestože také klesají melodií, nepoužily třepenou fonaci. Rychlý, až překotný konec promluvy mluvčích BOZA a STPA působí dojmem přísnosti, netrpělivosti – má negativní valenci ve srovnání s pozvolným zpomalením druhých dvou mluvčích.

Většina kontur v replice H5d\_3 se velmi podobá předchozím, ale je zde skupina tří mluvčích (BARA, BOZA a MORC), jejichž hodnoty se drží výše než ostatní nejen v bodu *ší-te*, kde je největší variabilita, ale i v okolních. Poslechový dojem je ten, že tyto tři mluvčí nerealizovali taktový předěl mezi *netušíte* a *kam to* a tedy vyslovili tato tři slova jako jediný přízvukový takt. Na konci repliky nejvíce zrychluje mluvčí VASA a zpomaluje mluvčí BRCA. Obě sice realizují stoupavý melodém, BRCA na poslední slabice ale stoupne ještě o něco více.

Věty ze čtveřice H6, tedy H6a\_3, H6b\_3, H6c\_3 a H6d\_3, v podobě *Víš, kdy to dostaneš a Víte, kdy to dostanete*, jsou i přes svoji nevelkou délku velmi variabilní. Obzvláště obě delší věty (c a d) dosahují vysoké průměrné variability srovnatelné s H4. Obě otázky jsou o něco málo variabilnější, to ale může souviset i s tím, že pro ně máme o něco více mluvčích. Kontury LAR jsou si ve všech čtyřech replikách nápadně podobné. Body nejvyšší variability se nacházejí na slabikách *to-do* a *sta-ne(š)*, přičemž nejstabilnější bod, kde se mluvčí od sebe navzájem liší nejméně, je umístěn mezi nimi, na slabikách *do-sta*. Směr těchto změn se ale v různých větách liší. V H6a\_3 jsou v obou bodech odchylky oběma směry, mluvčí zde tedy zrychlují i zpomalují. Využívají přitom všech možností, tzn. zrychlování na jednom z bodů, zpomalování na druhém, případně obráceně, zrychlují či zpomalují na obou, případně se drží modelových hodnot. Tyto odlišnosti jsou jednak způsobeny zacházením s oběma intervokalickými alveolárami, [d] a [n], často redukovanými na alveolární švih. Také by mohly souviset s rozdílnou realizací melodie na slově *dostaneš*, obzvláště mluvčí, kteří zpomalují, vyslovují na prvních dvou slabikách zřetelný melodický vrchol. V replice H6b\_3 (tj. v otázce) kromě dvou mluvčích na konci všichni zrychlují, zatímco na *to-do* jich zhruba třetina zpomaluje. Opět, nejrychlejší mluvčí (KORD, STOA a VASA) silně zkracují obě alveoláry. U nejpomalejších mluvčích opět zpomalení souvisí buďto s realizací větného důrazu (PCNA) a/nebo výrazné melodie (HOPA, ZETA). Replika H6b\_3 je zobrazena v grafu 9.9.



Obrázek 9.9: Graf kontur reziduí lokálního artikulačního tempa pro větu H6b\_3, *Víš, kdy to dostaneš?* Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

V replice H6c\_3 kromě stejného vzorce jako výše (ovšem tentokrát s odchylkami nahoru, tj. všichni mluvčí až na jednoho v bodech s největší variabilitou zrychlují) vykazuje velké odchylky mluvčí MORC, který výrazně zkracuje shluk [gd] a následující [t] ve slovech *kdy to*, což způsobuje, že jeho odchylka je na těchto pozicích zdaleka největší. Mluvčí JARS, OCEK a HROA zrychlují nejvíce na *to-do* a *sta-ne*, každý sice trochu jinak, ale výsledný dojem je velmi podobný. BOZA a RUBA, které mají nejnižší hodnoty na těchto dvou pozicích, vyslovují [d] jako aproximantu, což prodlužuje jeho trvání, a zároveň redukuje [o] na [ə]. Jejich slabika *do* tedy zní nedbale, přestože je trváním delší než u ostatních mluvčích.

Kontury reziduí v replice H6d\_3 jsou velmi podobné replice předchozí, jediným rozdílem je větší množství mluvčích, kteří zpomalují na *to-do*, nejvíce VRNA a RUBA, jejichž realizace jsou od sebe nicméně dosti odlišné. Mluvčí RUBA vyslovuje nedbale, hranice mezi jejími vokály a konsonanty se stírají, obě [t] v první části věty vyslovuje jako znělé dentální frikativy a její [d] na bodu nejvyšší variability je rovněž dentální. Mluvčí VRNA naopak vyslovuje nesmírně pečlivě a realizuje na slabice *do* velmi výraznou větnou prominenci s nízkým F0 (v porovnání s okolními slabikami) a vysokou intenzitou.

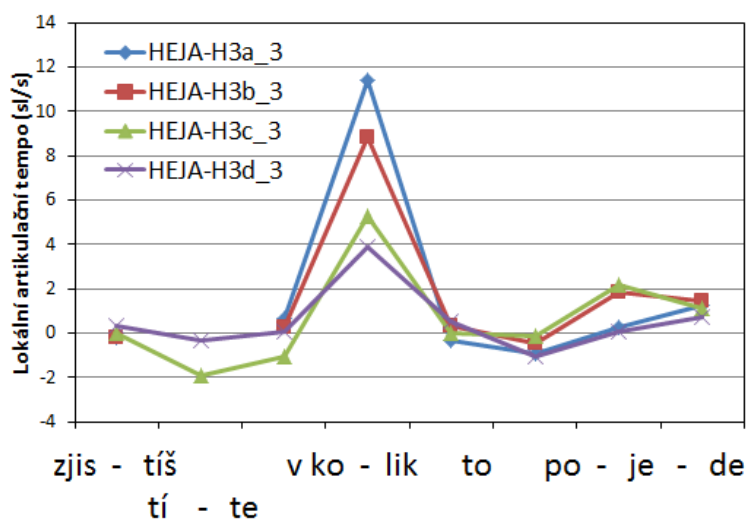
Pokles hodnot pod modelové na rozhraní slov *kdy to*, který je zřetelný ve všech čtyřech větách, je způsoben nadprůměrně dlouhým [t]. Mohlo by to být způsobeno tendencí ke slabičné izochronii – tak, aby slabika *to* byla stejně dlouhá jako sousední *kdy*, které má ovšem v onsetu dva konsonanty. Druhým možným vysvětlením by mohla být tendence nemít ve dvou sousedících slabikách dva zkrácené konsonanty, tedy zároveň [t] i [d], protože by slabičné vrcholy všech tří sousedících vokálů (tedy *kdy to do*) pak byly příliš blízko u sebe.

Průměrná absolutní odchylka jednotlivých mluvčích od modelových hodnot LAR se pohybuje mezi 0,58 a 1,38 sl/s. Nejvíce podobná modelu je mluvčí STOA, nejméně pak mluvčí JARS. Kromě mluvčího BROZ jsou všichni muži v polovině, která je odlišnější od modelu, což odpovídá i našim neformálním zkušenostem s jejich promluvami.

Pokusili jsme se také odhadnout stabilitu kontur LAR v rámci jednoho mluvčího – a to spočítáním Pearsonova korelačního koeficientu pro hodnoty ve stejných dvojicích vět (otázka vs. oznamovací věta). Korelační koeficienty, které kvantifikují konzistentnost v produkci kontur lokálního artikulačního tempa jednotlivých mluvčích, se pohybují mezi 0,3 (mluvčí STPA) a 0,84 (mluvčí STOB), průměr je 0,57. Šest mluvčích má vyšší korelaci než 0,7 a celkem 28 mluvčích vyšší než 0,5. Obecně ale neplatí vztah, že by mluvčí s největší odchylkou byli zároveň nejméně konzistentní, ani naopak.

Nízká korelace může být kromě nekonzistentnosti mluvčího ovšem také způsobena tím, že mluvčí různě frázuje, a tedy nebyly do této analýzy zahrnuty všechny jeho promluvy. Nižší korelace pak může být pouhým matematickým důsledkem menšího množství korelovaných případů.

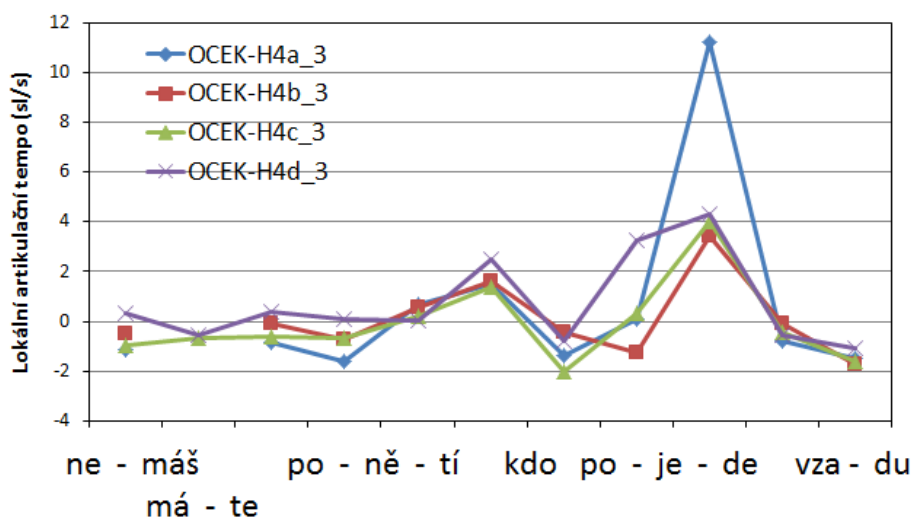
Na šest vnitřně nejkonzistentnějších mluvčích jsme se podívali blíže, a to v dialozích, kde jsou k dispozici data pro všechny čtyři repliky. Jejich kontury reziduí jsou si v jednotlivých dialozích velmi podobné, což znamená, že mluvčí realizují změny lokálního tempa konzistentně.



Obrázek 9.10: Graf kontur reziduí lokálního artikulačního tempa pro mluvčí HEJA a čtveřici replik H3, *Zjistíš, v kolik to pojedete* a *Zjistíte, v kolik to pojedete*. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Mluvčí HEJA lokálně zrychluje v replikách H1c\_3 a H1d\_3 na pozici *te-jim*, jinak se pohybuje kolem modelového trvání. Hodnoty LAR ve všech čtyřech replikách H1 ve druhé části věty (*co si myslíš/te*) jsou téměř totožné. Ve čtveřici H2 se také chová velmi konzistentně, oproti modelu vždy zpomaluje ve vedlejší větě *co bude dál* (vyslovuje velmi zřetelně) – ani jednou nerealizuje lokální zrychlení jako většina ostatních mluvčích. Jediná odlišnost je ve větě d, kdy HEJA realizuje důraz na slově *dál*, čemuž přizpůsobuje lokální načasování a v předchozích slabikách oproti ostatním replikám lehce zrychluje (její hodnoty se tak blíží modelovým). Také ve čtveřici H3 (obrázek 9.10) jsou si křivky jejich reziduí velmi podobné. Nastává zde zrychlení ve slově *kolik*; jinde se hodnoty téměř neliší od modelových, a tedy mluvčí nerealizuje tak výrazné zrychlení na slově *pojede*, které se objevuje u mnoha jiných mluvčích. Ve čtveřici dialogů H5 nacházíme lehké odlišnosti mezi realizací otázek a oznamovacích vět, nicméně opět se hodnoty mluvčí HEJA pohybují kolem modelových. Nerealizuje tak výrazné zrychlení na pozici *ší-te* jako jiní mluvčí (pouze o 1 sl/s rychlejší než model). Také nezrychluje na posledních dvou slabikách.

U mluvčí HOBA jsme porovnali věty z dialogů H2 a H3. V první čtveřici mluvčí výrazně zpomaluje na pozici *bu-de*, mnohem více než mluvčí HEJA. Rozdíly na ostatních pozicích se pravděpodobně dají připsat různému artikulačnímu tempu v promluvěch – křivky reziduí jsou si velmi podobné, ale položené výše či níže. Ve čtveřici H3 stejně jako HEJA zrychluje na pozici *ko-lik*, ale celkově méně výrazně. Naopak více než ona zrychluje na posledních slabikách *je-de*.



Obrázek 9.11: Graf kontur reziduí lokálního artikulačního tempa pro mluvčího OCEK a čtveřici replik H4, *Nemáš poněti, kdo pojede vzadu* a *Nemáte poněti, kdo pojede vzadu*. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Mluvčí OCEK se v dialozích H2 a H5 od modelových hodnot téměř neodchyluje. Zrychluje ve čtveřici H3 na slovech *kolik* a *pojede* (kolem 4 sl/s), na pozici *je-de* také ve čtveřici H4. Zde se chová také velmi konzistentně na pozicích *tí-kdo-po*, kde nejprve zrychlí na 1-2 sl/s více než model, a následně podobně zpomalí, viz obrázek 9.11. Všechny čtyři promluvy jsou si v tomto chování velice podobné. Ve čtveřici H6 mají jeho rezidua dvě lokální maxima, a to na pozicích *to-do* a *sta-ne(š)*.

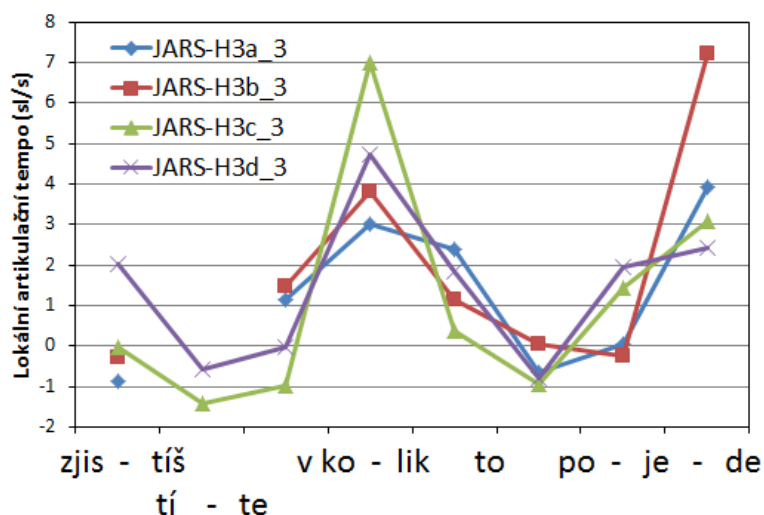
Mluvčí SEBA má ve čtveřici H2 podobné všechny kontury kromě repliky a, ve které zpomaluje na slově *bude*, zatímco v ostatních třech replikách zde oproti modelu lehce zrychluje. V dialozích H3 výrazně zrychluje na slově *pojede*, a to více v otázkách než v oznamovacích větách. Konzistentní je také ve zpomalení na předchozí pozici *to-po*. Podobně jako OCEK má i SEBA dvě lokální maxima v konturách reziduí dialogů H6, *to-do* a *sta-ne(š)*, nicméně na rozdíl od něj je v jejich realizaci méně konzistentní.

Kontury mluvího STOB jsou ve čtveřici H2 téměř identické s modelovými, pouze drobně klesají na posledních dvou slabikách. Ve všech čtyřech replikách H3 velmi konzistentně zrychluje na slově *kolik* a také na *pojede*, v otázkách jsou kontury výše na poslední, v oznamovacích větách pak na předposlední slabice. V dialozích H3 ve všech případech zpomaluje na pozici *tu-ší(š)* a následně zrychluje ve větách c a d na *ší-te*. Ve čtveřici H6 vykazuje na rozdíl od předchozích dvou mluvích jen jedno lokální maximum, a to *sta-ne(š)*, kromě věty a, ve které drobně zpomaluje. Velmi konzistentní je na pozici *kdy-to*, kde jsou hodnoty všech čtyř reziduí téměř identické.

Mluvčí VASA vykazuje velice konzistentní chování v dialozích H2, H3, H5 i H6 (v ostatních nejsou k dispozici všechny repliky). V H2 (kromě c) výrazně zrychluje na *bude*, v H3 na *kolik* a *pojede*, v H5 na *ší-te* a *va-jí* a v H6 na *to-do* a *sta-ne(š)*. Výrazněji pod modelovou konturu se dostává pouze na pozici *tu-ší(š)* (H5) a *kdy-to* (H6).

Prozkoumali jsme také dva mluví s nejmenší vnitřní konzistencí podle korelace, STPA a JARS. Jak už bylo zmíněno, nejsou u nich k dispozici všechny repliky od všech dialogů, u STPA bylo možné zkoumat pouze čtveřici H3. I zde se však dá najít konzistentní chování. Ve slově *pojede* tato mluví oproti modelu téměř nezrychluje, kromě věty c, kde nepatrně protahuje první [e] na úkor předchozího [oj]. Rovněž zrychluje na *ko-lik*, kromě věty b, kde vyslovuje zřetelněji [l] než ve větách ostatních. Zbylá rezidua jsou vzájemně poměrně konzistentní.

Mluvčí JARS využívá v dialozích H2 dva temporální vzorce. Ve větách a a d výrazně zrychluje na *co-bu* a naopak zpomaluje na následujícím *bu-de*, zatímco ve větách b a c jsou obě pozice lehce nad modelovými hodnotami. V dialozích H3 je naopak poměrně konzistentní, ve všech čtyřech případech zrychluje na *ko-lik* a *je-de* (viz obrázek 9.12). Ve čtveřici H5 vykazují jeho kontury velkou variabilitu na slovech *netušíš* a *netušíte*, což souvisí s realizací (nebo naopak absencí) větného důrazu na tomto slově.



Obrázek 9.12: Graf kontur reziduí lokálního artikulačního tempa pro mluvčího JARS a čtveřici replik H3, *Zjistíš, v kolik to pojedě* a *Zjistíte, v kolik to pojedě*. Reziduum je vypočítáno jako rozdíl modelové a reálné hodnoty LAR. Nula zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu.

Lze tedy uzavřít, že temporální vzorce zkoumaných mluvčích jsou ve stejných či podobných větách velmi konzistentní, pokud jsou dodrženy srovnatelné podmínky – zejména dělení na prozodické fráze či větný důraz. Naopak modalita věty ve většině případů roli nehraje. Ve forenzních podmínkách může být tento výsledek užitečný, pokud pracujeme s takto porovnatelnými promluvami (například čtený přepis telefonátu s jeho nahrávkou).

## 10.6 Diskuse k temporálnímu modelu

Za pomoci deskriptivních analýz z první části práce byl vytvořen model temporálních charakteristik hlásek na korpusu Minidialogy-H. Z modelových hodnot trvání hlásek byly pro každou repliku vypočítány modelové kontury lokálního artikulačního tempa (podle Volína, 2009, viz oddíl 7.6) a jednotlivé reálné kontury mluvčích s nimi byly porovnány. Odchytky v reziduích reálných kontur ukazují na dva jevy, které nebyly při vytváření modelu vzaty v úvahu, a to větný důraz a předěl mezi přízvukovými takty. Obojí vytváří lokální zpomalení,



kteřé model není schopn zachytit. Bylo by řádoucí tyto parametry v budoucnu také zakomponovat.

V kapitole 4 byly uvedeny dva příklady temporálních modelů, jednak pravidlový model Carlsona (1991), který je založen na modelu Klattově (1976), a statistický model Lazaridise et al. (2010). Carlsonův model předpovídal hodnoty trvání hlásek se směrodatnou odchylkou 20 ms. Model popsany v této práci je o něco úspěšnější, průměrná odchylka rozdílů modelových hodnot od reálných činí 16,5 ms.<sup>8</sup> Stejně tak lze úspěšnost našeho modelu porovnat i s modernějším statistickým modelem Lazaridise et al. (2010), který uvádí korelaci modelových a skutečných hodnot mezi 0,6 a 0,8. Naš temporální model dosahuje srovnatelné korelace 0,66 (Pearsonův korelační koeficient). Tyto výsledky jsou o to cennější, že oba citované modely byly vytvořeny pouze na základě jediného mluvčího, zatímco zde je variabilita mnohem větší, modelujeme 34 mluvčích.

Naším cílem bylo také postihnout variabilitu v lokálním artikulačním tempu mezi mluvčími. Ukazuje se, že místa v promluvě, kde dochází k individuální variabilitě, nejsou náhodná. Mluvčí si neupravuje načasování segmentů libovolně, místa, na kterých se mluvčí od sebe navzájem odchyľují, nejsou náhodná a vyskytují se systematicky. V řádne z replik se nestalo, že by stejná pozice LAR, případně stejné slabiky, byly jednou místem s velkou variabilitou a jindy místem stabilním. V těch částech čtveřic vět, které mají stejný text, jsou místa s největší variabilitou (a obvykle také místa s největší stabilitou) stejná. Pokud se v některých dialozích liší, tak v závislosti na změně textu – například v H1 se pozice s největší variabilitou přidáním jedné slabiky přesunula ze slabik *řek-neš* na slabiky *te-jim* (*z řeknete jim*), podobně i v H5.

Také další výsledky vedou k závěru, že pozice největší individuální variability je zejména závislá na textu, tedy na segmentálním obsahu promluvy. Slovo *pojede*, které se vyskytuje ve dvou analyzovaných dialozích, má vždy nejvyšší variabilitu z celé věty, přestože v dialozích H3 tvoří finální takt celé fráze a v H4 je součástí mediálního taktu (*kdo pojede*).

Důležitější se přitom zdají být konsonanty než vokály. Shrneme-li si pozice s největší variabilitou, jde o řetězce hlásek [ed] (9×), [oj] (7×), [od], [ud], [ul], [an] (4×), [aj], [ob], [im], [oň] (3×), [ej], [i:t] (2×) a [em], [ekn] a [uš] (1×). Kromě jedné výjimky jde vždy o intervokalický konsonant (hodnota LAR je počítána od onsetu prvního k onsetu druhého vokálu, tedy za konsonantem následuje další vokál) a kromě dvou výjimek o sonoru nebo znělou explozivu (bilabiální či alveolární). Ze znělých exploziv je přitom zdaleka nejčastější [d] a ze sonor [j] – zacházení s těmito hláskami umožňuje velký prostor pro individuální

---

<sup>8</sup> Směrodatná odchylka je o necelé dvě desetiny ms vyšší, nicméně koncepčně pokládáme směrodatnou odchylku odchylek za problematickou.

variabilitu. Co se týče [j], mohli bychom namítat, že hranice této hlásky se špatně značkují, a tedy může být variabilita také výsledkem neshody v anotaci. Proti tomu ovšem hovoří za první výsledek z oddílu 8.3.6.2, na obrázku 8.81, kde je zřetelně vidět, že intervokalická [j] jsou velmi stabilní (na rozdíl například od intervokalických [r]), za druhé výsledek z oddílu 9.1, tabulka 8.22, ze které je vidět, že samotné trvání [j] není specifické pro mluvčí (rozlišilo pouze 2,3 % dvojic mluvčích) – jde o to, jak se v okolí této hlásky bortí temporální struktura výpovědi. Stejně můžeme uvažovat i v případě ostatních míst největší variability – samotné trvání znělých exploziv ani sonor k rozpoznání mluvčího nestačí (úspěšnost maximálně 19 %), jde o komplexnější zacházení s temporálními charakteristikami v okolí těchto hlásek, co prozradí více o individualitě mluvčího.

Vokály jsou přitom mezi body variability a body stability poměrně rovnoměrně rozprostřené, jedině dlouhý vokál [a:] není nikdy v bodu největší variability (a v šesti výskytech z osmi dokonce v bodu nejvyšší stability), zatímco krátké vokály a [i:] ano, i když [i:] pouze ve dvou výskytech z 16. V budoucnu by bylo užitečné ověřit, zda jsou dlouhé vokály obecně temporálně stabilnější a méně náchylné k individuálním rozdílům než krátké – [i:] je pravděpodobně temporálně tvárnější díky oslabenému kontrastu délky.

Z ostatních prozodických faktorů se zdá mít na umístění bodu největší variability vliv ještě délka taktu – tyto pozice se vyskytují téměř vždy v delších taktech – tří a čtyřslabičných. Výjimkou je pouze čtveřice H1, kde místo s největší variabilitou připadá na přelom taktů *co si* a *myslíš*, ovšem je možné, že mluvčí ve skutečnosti realizují tyto dva dvouslabičné takty jako jeden čtyřslabičný. Tato tendence by také vysvětlovala, proč ve slově *vzadu* (v replikách H4), kde se rovněž nachází intervokalické [d], není variabilita zdaleka tak vysoká jako v sousedním slově *pojede*. Stejně bychom mohli i vysvětlit, proč nenalézáme variabilitu ve slově *nemáš* kolem intervokalického [m] (také z replik H4a a b). Nízká variabilita mezi mluvčími ve slově *nemáte* (repliky H4c a d) je ovšem spíše zaviněna nedostatkem mluvčích v těchto dvou replikách.

V analyzovaných větách se nacházejí ještě další intervokalické konsonanty, [p], [t], [tʰ], [s], [š], [c] a [v]. Žádné z nich však nepředstavují bod největší variability a některé z nich naopak tvoří pozici s variabilitou nejmenší. Je otázka, zda se intervokalické znělé frikativy (především [z]) budou také chovat idiosynkraticky, to bude třeba ověřit v dalším výzkumu.

Pozice největší stability se zdají být rovněž podmíněny segmentálně. Nacházejí se převážně na souhláskových shlučích, a to často s neznělou frikativou: [isl] v *myslíš*, [a:šs] v *zeptáš se*, [istʰ] ve *zjistíš*, [efk] ve *zjistíte v kolik*, [a:šp] v *nemáš poněti*, [ost] v *dostaneš*. Ostatní případy

obsahují nejčastěji neznělý obstruent ([es] z *přesouvají*, [ec] ze *se co*, [a:t] ze *zeptáte, nemáte* a [opř] v *to přesouvají*).

Body stability jsou často (i když ne vždy) těsně vedle bodů největší variability, oba mohou být i součástí jediného slova (např. *přesouvají* z H5 či *dostaneš* z H6). Z hlediska lokálního artikulačního tempa by tedy bylo příliš zjednodušující brát slova jako celek – jak vidíme, i v rámci slova dochází k zajímavým temporálním odlišnostem.

Z výsledků také nelze vyvodit, že by zásadní roli hrála délka repliky (tedy prozodické fráze). Průměrná absolutní variabilita (spočítaná jako průměr absolutních hodnot všech odchylek) repliky koreluje s délkou prozodické fráze jen slabě (Pearsonův korelační koeficient  $r = 0.4$  pro délku prozodické fráze vyjádřenou ve slabikách i fonémech).

Poměrně dispreferovanou pozicí pro bod největší variability je přízvučná slabika (tedy její vokál, případně koda). Jen čtvrtina bodů největší variability (ale zato polovina bodů největší stability) se nachází na přízvučném vokálu. Z celkového počtu přízvučných vokálů je to 20,5 %. Prozkoumána byla i pozice kolem taktového předělu, ta se ovšem nezdá být nijak zvlášť preferována pro variabilitu ani stabilitu mezi mluvčími.

Závěry z předchozí studie (Volín & Weingartová, 2012) ohledně kategorie slova se nepotvrdily – nebyl nalezen žádný rozdíl mezi autosémantickými a synsémantickými slovy vzhledem k umístění bodů největší variability a stability. Variabilitu i stabilitu najdeme v obou kategoriích slov a také na přechodech mezi nimi (oběma směry).

Forenzně zajímavější je přitom variabilita směrem vzhůru, tj. k vyšším hodnotám lokálního artikulačního tempa. Odchyly směrem dolů jsou, jak bylo zmíněno výše, způsobeny často zdůrazněním, které je vědomě ovládáno mluvčím. Zatímco segmentální redukce, zkracování, a tím vyvolané změny v rytmicizaci (např. skladba fráze vzhledem k přízvukovým taktům), spíše nebudou přímo vědomě ovládány.

Z našich výsledků je zřetelná provázanost jednak segmentální a prozodické vrstvy řeči, jednak také temporální domény, F0 a intenzity. V prvním případě jde o to, že příčinná souvislost funguje oběma směry. Záměrem mluvčího není vyslovit např. [d] jako alveolární švih, což následně způsobí jeho zkrácení a lokální změnu rytmu, ale naopak nějak realizovat rytmus a načasování slabik v promluvě, k čemuž jsou mu k dispozici určité segmentální prostředky v rámci fonologického systému jazyka. Zároveň je variabilita temporálních charakteristik vázána na segmentální rovinu, protože mluvčí může hlásky variovat jen do té míry, aby nedošlo k jejich záměně a následnému nepochopení.

Ve druhém případě jde o fakt, že změny lokálního artikulačního tempa jsou často doprovázeny realizací výrazné melodie, či intenzitního vrcholu, případně kombinací těchto

jevů. Nelze říci, že by některý z těchto parametrů byl příčinou a jiný důsledkem, všechny tři jsou provázány a příčinou jejich změn je plánování promluvy na vyšší úrovni. Uvědomme si, že v tomto případě nejde o slovní přízvuk (který v našich datech nehraje žádnou větší roli), ale o prominenci na úrovni prozodické fráze, která se často projevuje kontrastivním průběhem melodie (zvýšení nebo snížení vůči okolním slabikám), intenzity a zpomalením lokálního artikulačního tempa.

Zajímavým jevem, na který jsme v průběhu analýz narazili, je fakt, že zpomalení lokálního artikulačního tempa nemusí znamenat pečlivější artikulaci. Například změna explozivny na frikativu, ke které dochází při nedbalé řeči, způsobí lokální zpomalení, stejně tak i některé změny místa artikulace.

Jeden z mluvčích, MACK vykazuje řečovou vadu, konkrétně rotacismus. Nicméně se neukázalo, že by tato vada měla vliv na jeho temporální charakteristiky. [r] mluvčího MACK jsou průměrně dlouhá a statisticky významně se trváním liší pouze od jediného mluvčího (BROZ), jehož [r] jsou velmi krátká. Také v analýzách lokálních artikulačních kontur mluvčí MACK nijak zvlášť nevyčnívá, jeho průměrná absolutní odchylka od modelu je sice vyšší, 1,02 sl/s, ale stále ještě existuje 9 mluvčích, jejichž průměrná odchylka je větší.

Zmiňme též otázku slabičné a taktové izochronie. Tato problematika nebyla v této práci systematicky zkoumána, nicméně některá pozorování při analýzách kontur LAR nás vedou k domněnce, že mluvčí mohou obě formy izochronie používat jako prostředky k modifikaci rytmu svých promluv. Přísná slabičná izochronie se s našimi daty neslučuje – pokud by byla mezi onsetsy vokálů stále stejná vzdálenost, znamenalo by to, že hodnoty LAR by se pohybovaly stále v podobné výši. K tomu však nedochází. Průměrné směrodatné odchylky hodnot LAR jednoho mluvčího v jedné replice (nenormalizovaných k modelu) se pohybují kolem 3 sl/s, což znamená, že jsou v rámci repliky dost variabilní.

Rozhodnutí nevyužít v modelu jako parametr modalitu věty se ukázalo být efektivní. Otázky a oznamovací věty se v konturách LAR neliší. Předběžné analýzy naznačují, že se od sebe liší pouze trváním posledního vokálu (který je výrazně delší u otázek), pokud replika končí na otevřenou slabiku.

## 11 Závěr

Tato disertační práce se skládá ze dvou hlavních částí. První (kapitoly 8 a 9) je deskriptivní popis temporálních charakteristik českých hlásek. Trvání hlásek je podrobně popsáno a prozkoumáno v závislosti na frázovém zpomalování, pozici ve vyšší jednotce (slabice, slově, prozodické frázi), délce vyšší jednotky (slova a prozodické fráze), hláskovém okolí, případně struktuře slabiky. Propojuje se tak segmentální a suprasegmentální rovina řeči, jelikož je bráno v úvahu trvání individuálních hlásek a jeho změna v závislosti na prozodických faktorech. Tyto výsledky jednak obohacují lingvistický popis češtiny, ale také nabízejí srovnání pro budoucí výzkum v rozličných – nejen fonetických – oborech.

Na základě naměřených údajů je pak vytvořen model temporálních charakteristik češtiny (kapitola 10), který se skládá z výchozího trvání každé hlásky a sady pravidel, která toto trvání modifikuje. Tento model byl pak využit ke zkoumání kontur lokálního artikulačního tempa a hledání individuálních odchylek jednotlivých mluvčích v reziduích modelu.

Co se týče závěrů této práce užitečných pro identifikaci mluvčího: v první části bylo ukázáno, že artikulační tempo a některé rytmické ukazatele (zejména %V,  $\Delta C$  a rPVI-C) mohou být velmi užitečné pro rozpoznávání mluvčího, protože vykazují variabilitu mezi mluvčími a stabilitu v rámci jednoho mluvčího (viz odd. 9.2 a 9.3). Kromě toho lze individuální specifika nalézt i v hodnotách trvání krátkých vokálů, či v poměru dlouhého a krátkého [I] (viz 9.1).

Z analýz kontur lokálního artikulačního tempa vyplývá, že mluvčí se od sebe liší i v tom, jak zacházejí s artikulačním tempem v průběhu jedné promluvy, a to dokonce i v rámci jediného slova. Tyto odlišnosti přitom nejsou náhodné – mluvčí nevariují na libovolných místech v promluvě. Je to dané primárně segmentálním obsahem promluvy (tj. identitou hlásek), sekundárně také délkou přízvukového taktu. Naopak se zdá, že kategorie slova (autosémantické vs. synsémantické) roli nehraje. K největší individuální variabilitě přitom dochází v okolí intervokalických znělých exploziv (především [d]) a intervokalických sonor (zejména [j]).

Výhodnější přitom je sledovat variabilitu směrem nahoru, tj. k vyšším artikulačním tempům – snížení hodnot je často důsledkem větného důrazu, zatímco zkracování hlásek může spíše být idiosynkratické.

Vytvořený temporální model najde uplatnění i jinde než ve zkoumání specifík mluvčích. Užitečným může být například při syntéze nebo rozpoznávání řeči, také může sloužit jako srovnávací měřítko při různých dalších výzkumech prozodické struktury, cizineckého přízvuku, řečových vad, a podobně – díky němu je možné si vytvořit představu o tom, co je v češtině obvyklé a co je již vybočující.

Přínosem práce je rovněž ručně označovaný řečový korpus, který obsahuje 4046 replik od 34 mluvčích s celkovým trváním větším než dvě hodiny. Každá replika je anotována na vrstvách prozodických frází, slov, slabik, typů hlásek, hlásek a fonémů.

Omezení této práce spočívají především ve faktu, že kontexty v korpusu Minidialogy nevyužívají všech kombinatorických a distribučních možností češtiny – ne všechny hlásky se jsou dostatečně zastoupeny v různých pozicích a okolích. To je problémem zejména pro dlouhé vokály. Nicméně pro forenzní účely bylo na druhou stranu mnohem výhodnější mít mluvčí, kteří vyslovují téže promluvy, než textovou bohatost. Vždy je třeba nalézt kompromis, který je nejvhodnější vzhledem k výzkumné otázce.

Na důkladnější prozkoumání také čekají rozdíly v rámci jednoho mluvčího – jak stabilní jsou temporální kontury v promluvách jediné osoby. Tento směr výzkumu je svým způsobem ortogonální k problematice této práce a nebylo možno mu zde věnovat náležitou pozornost.

Do modelu je také možné přidávat další pravidla, která budou modelovat dosud nezachycené faktory – zejména prozodické předěly hloubky 2 (tedy přízvukové takty) a 3. Rovněž by bylo užitečné přidat v budoucnu pravidlo kompenzující vliv artikulačního tempa na jednotlivé hlásky.

Přestože množství řečového materiálu, které jsme měli k dispozici, je na poměry fonetických studií velmi vysoké, nelze naměřená trvání hlásek nazvat referenčními pro všechny mluvčí češtiny. Nicméně je možné považovat hodnoty předložené v této práci za referenční pro mladé dospělé mluvčí z Čech. Pro vytvoření skutečně referenční databáze bychom museli mít k dispozici širší vzorek mluvčích vzhledem k pohlaví, věku, vzdělání, případně regionu, odkud mluvčí pochází. K tomuto účelu by v budoucnu mohl posloužit korpus VASST, který je v současné době budován na Fonetickém ústavu (Volín & Weingartová, v přípravě). Zde bude možné také ověřit závěry této práce i v rámci jiných mluvních stylů (především spontánní řeči).

## Reference

- Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, 66, 46–63.
- Asu, E. L. & Nolan, F. (2006). Estonian and English rhythm: a twodimensional quantification based on syllables and feet. In: *Proceedings of Speech Prosody 2006*, Dresden, Germany.
- Balkó, I. (2005). K výzkumu tempa řeči a tempa artikulace v různých řečových úlohách. *Bohemistyka*, 3, 185–198.
- Barbosa, P. A. (2002): Explaining Cross-Linguistic Rhythmic Variability via a Coupled-Oscillator Model of Rhythm Production. *Proceedings of Speech Prosody 2002*, 163–166, Aix-en-Provence.
- Barry, W., Andreeva, B. & Koreman, J. (2009). Do Rhythm Measures Reflect Perceived Rhythm? *Phonetica*, 66/1-2, 78–94.
- Bartoň, T., Cvrček, V., Čermák, F., Jelínek, T. & Petkevič V. (2009): *Statistiky češtiny*. Praha: Nakladatelství Lidové noviny.
- Beckman, M. E. & Elam, G. A. (1997): *Guidelines for ToBI labelling*. Verze 3.0, březen 1997. Publikováno online: [http://www.ling.ohio-state.edu/~tobi/ame\\_tobi/labelling\\_guide\\_v3.pdf](http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf).
- Beňuš, Š. (2009): Are we 'in sync': Turn-taking in collaborative dialogues. *Proceedings of the 10th INTERSPEECH*: 2167–2170.
- Boersma, P. & Weenink, D. (2014): *Praat: doing phonetics by computer* [počítačový program], verze 5.3.71. Získáno z <http://www.praat.org/>.
- Borovičková, B. & Maláč, M. (1967): *The Spectral Analysis of Czech Sound Combinations*. Praha: Academia.
- Boucher, V. J. (2002): Timing relations in speech and the identification of voice-onset times: A stable perceptual boundary for voicing categories across speaking rates. *Perception & Psychophysics*, 64/1, 121–130.
- Bouchrika, I., Goffredo, M., Carter, J. & Nixon, M. (2011). On using gait in forensic biometrics. *Journal of Forensic Sciences*, 56, 882–889.
- Bourland, H., Hermansky, H. & Morgan, N. (1996): Towards increasing speech recognition error rates. *Speech Communication*, 18/3, 205–231.
- Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, 15, 39–54.
- Caisse, M. (1988): *Modelling English vowel durations*. Nepublikovaná disertační práce. University of California, Berkeley.
- Carlson, R. (1991): Duration models in use. *Proceedings of the XIIth ICPhS*, 278–281, Aix-en-Provence.
- Carlson, R., Granström, B., & Klatt, D. H. (1979): Some notes on the perception of temporal patterns in speech. In Lindblom, B., & Öhman, S. (eds.), *Frontiers of Speech Communication Research*, 223–243. London: Academic Press.
- Cummins, F. (2002): Speech Rhythm and Rhythmic Taxonomy. *Proceedings of Speech Prosody 2002*, 121–126, Aix-en-Provence.
- Cummins, F. (2003). Practice and performance in speech produced synchronously. *Journal of Phonetics*, 31, 139–148.

- Cummins, F. (2009): Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics*, 37/1, 16–28.
- Cummins, F. & Port, R. F. (1998): Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26/2, 145–171.
- Dankovičová, J. (2001): *The Linguistic Basis of Articulation Rate Variation in Czech*. Forum Phonicum, 71. Hector: Frankfurt am Main.
- de Jong, Kenneth J. (2004): Stress, Lexical Focus, and Segmental Focus in English: Patterns of Variation in Vowel Duration. *Journal of Phonetics*, 32, 493–516.
- Dellwo, V. (2006). Rhythm and Speech Rate: A Variation Coefficient for C. In: P. Karnowski & I. Szigeti (eds.), *Language and Language Processing*, 231–241. Frankfurt am Main: Peter Lang.
- Dellwo, V. & Koreman, J. (2008). How speaker idiosyncratic is measurable speech rhythm? *Proceedings of IAFPA 2008*. Lausanne: IAFPA.
- Dellwo, V., Ramyeard, S. & Dankovičová, J. (2009): The influence of voice disguise on temporal characteristics of speech. *Proceedings of IAFPA 2009*. Cambridge: IAFPA.
- Dellwo, V., Kolly, M. & Leemann, A. (2012). Speaker identification based on temporal information: A forensic phonetic study of speech rhythm and timing in the Zurich variety of Swiss German. *Proceedings of IAFPA 2012*. Santander: IAFPA.
- Donovan, A. & Darwin C. J. (1979): The perceived rhythm of speech. *Proc. 9th ICPhS, vol. II*, 268–274, Copenhagen.
- Eriksson, A. & Wretling, P. (1997). How flexible is the human voice? A case study of mimicry. *Proceedings of Eurospeech*, 97, 1043–1046.
- Ghitza, O. & Greenberg, S. (2009): On the possible role of brain rhythms in speech perception: Intelligibility of time compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66, 113–126.
- Gold, E. & French, P. (2011) An international investigation of forensic speaker comparison practices. *Proc. 17th ICPhS, Hong Kong*.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In: N. Warner, & C. Gussenhoven (eds.), *Papers in laboratory phonology 7*, 515–546. Berlin: Mouton de Gruyter.
- Hála, B. (1962): *Uvedení do fonetiky češtiny na obecně fonetickém základě*. Praha: Československá akademie věd.
- Hitchcock, L. & Greenberg, S. (2001): Vowel height is intimately associated with stress accent in spontaneous American English discourse. *Proceedings of EUROSPEECH-2001*, 79–82.
- Hollien, H. (2002). *Forensic Voice Identification*. London: Academic Press.
- Homolková, V. (2009): *Temporální vlastnosti českých frikativ*. Diplomová práce. Praha: Fonetický ústav FF UK.
- Huggins, A. F. (1979): Some effects on intelligibility of inappropriate temporal relations within speech units. *Proc. 9th ICPhS, vol. II*, 283–289, Copenhagen.



- Chlumský, J. (1911): *Pokus o měření českých zvuků a slabik v řeči souvislé*. Praha: Česká akademie.
- Chlumský, J. (1928): *Česká kvantita, melodie a přízvuk*. Praha: Česká akademie věd a umění.
- Chung, Y. & Arvaniti, A. (2012): Speech cycling in Korean. *Proceedings of Perspectives on Rhythm and Timing*, 10, Glasgow.
- Churaňová, E. (2013): *Vliv lokálních a globálních temporálních změn na úspěšnost rozpoznávání souvislé řeči v úlohách s velkým slovníkem*. Nepublikovaná seminární práce. Praha: Fonetický ústav.
- Churaňová, E., Šturm, P. & Weingartová, L. (v recenzním řízení): *Changes in segmental timing in slow and fast metronome-synchronized speech*. Nepublikovaný rukopis.
- Jacewicz, E., Fox, R. A. & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America*, 128, 839–850.
- Janota, P. & Palková, Z. (1974): Auditory evaluation of stress under the influence of context. *AUC Philologica 2/1974, Phonetica Pragensia*, 4, 29–59.
- Jessen, M. (2007). Forensic reference data on articulation rate in German. *Science & Justice*, 47, 50–67.
- Jun, S.-A. (2005): Prosodic Typology. In S.-A. Jun (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*, 430–458. Oxford University Press.
- Kaiser, L. (1964): Phonetic similarity apart from linguistic affinity. *Zeitschrift für Phonetik 17*: 243–249.
- Kim, E. Y. & McAuley, J. D. (2012). Dynamic attending and the perceived duration of auditory oddballs. *Proceedings of Perspectives on Rhythm and Timing*, 26, Glasgow.
- Klatt, D. H. (1976): Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208–1221.
- Knight, S. & Cross, I. (2012): Rhythms of persuasion: The perception of periodicity in oratory. *Proceedings of Perspectives on Rhythm and Timing*, 27, Glasgow.
- Kohler, K. J. (2009a): Rhythm in Speech and Language: A New Research Paradigm. *Phonetica*, 66, 29–45.
- Kohler, K. J. (2009b): Whither speech rhythm research? *Phonetica*, 66, 5–14.
- Kučera, H. & Monroe, G. K. (1968): *A comparative quantitative phonology of Russian, Czech and German*. New York: American Elsevier Publishing Company.
- Künzel, H. J. (1997). Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics*, 4, 48–83.
- Künzel, H. J., Masthoff, H. R. & Köster, J.-P. (1995). The relation between speech tempo, loudness, and fundamental frequency: an important issue in forensic speaker recognition. *Science & Justice*, 35, 291–295.
- Large, E. W., Almonte, F. & Velasco, M. (2010). A canonical model for gradient frequency neural networks. *Physica D*, 239, 905–911.
- Lazaridis, A., Ganchev, T., Kostoulas, T., Mporas, I. & Fakotakis, N. (2010): Phone duration modeling: Overview of techniques and performance optimization via feature selection in the context of emotional speech. *International Journal of Speech Technology 13/3*, 175–188.

- Leemann, A., Kolly, M.-J. & Dellwo, V. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic Science International*, 238, 59–67.
- Lehiste, I. (1970): *Suprasegmentals*. Cambridge, MA: The MIT Press.
- Lehiste, I. (1973): Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America*, 54, 1228–1234.
- Lehiste, I. (1977): Isochrony reconsidered. *Journal of Phonetics*, 5, 253–263.
- Lehiste, I. (1979): Temporal relations within speech units. *Proc. 9th ICPhS*, vol. II, 241–244.
- Lindblom, B. (1967): Vowel duration and a model of lip-mandible coordination. *STL-QPRS* 4, 1–29.
- Loukina, A., Kochanski, G., Rosner, B., Keane, E. & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustical Society of America*, 129/5, 3258–3270.
- Low, E. L., Grabe, E. & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language & Speech*, 43/4, 377–401.
- Lowit, A. (2012). Application of rhythm metrics in disordered speech: What should we measure and what do they really tell us? *Proceedings of Perspectives on Rhythm and Timing*, 36, Glasgow.
- Ludvíková, M. (1987): Čísla o hláskách. In: M. Těšitelová et al. (eds.), *O češtině v číslech*, 91–108. Praha: Academia.
- Machač, P. (2008): Desonorizace českých intervokálních frikativ. In: J. Volín & J. Janoušková (eds.), *AUC Philologica 2/2007, Phonetica Pragensia XI*, 105–116. Praha: Karolinum.
- Machač, P. (2009): *Temporální a spektrální struktura českých explozív*. Disertační práce. Praha: Fonetický ústav.
- Machač P. & Skarnitzl R. (2009): *Fonetická segmentace hlásek*. Praha: Nakladatelství Epoque.
- Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *Journal of the Acoustical Society of America*, 125, 3962–3973.
- Mixdorff, H. & Pfitzinger, H. R. (2005). Analysing fundamental frequency contours and local speech rate in map task dialogs. *Speech Communication*, 46, 310–325.
- Niebuhr, O. (2010). On the phonetics of intensifying emphasis in German. *Phonetica*, 67, 170–198.
- O'Dell, M. & Nieminen, T. (2009): Coupled Oscillator Model for Speech Timing: Overview and Examples. *Nordic Prosody, Proceedings of the Xth Conference*, 179–190, Helsinki.
- Ondrušková, L. (2011): *Zvukové vlastnosti jednoslabičných slov v semispontánním dialogu a hlasitém čtení*. Diplomová práce. Praha: Fonetický ústav.
- O'Shaughnessy, D. (1995): Timing Patterns in Fluent and Disfluent Spontaneous Speech. *Proceedings IEEE Conference on Acoustics, Speech and Signal Processing, vol. I*, 600–603.
- Palková, Z. (1994): *Fonetika a fonologie češtiny*. Praha: Karolinum.

- Petr, J. (ed.) et al. (1986). *Mluvnice češtiny I: Fonetika, fonologie, morfonologie a morfemika, tvoření slov*. Praha: Academia.
- Pickett, E. R., Blumstein, S. E. & Burton, M. W. (1999): Effects of Speaking Rate on the Singleton/Geminate Consonant Contrast in Italian. *Phonetica*, 56, 135–157.
- Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Podlipský, V. J. (2009): *Reevaluating perceptual cues: Native and non-native perception of Czech vowel quantity*. Disertační práce. Olomouc: Katedra anglistiky a amerikanistiky.
- Podlipský V. J., Skarnitzl R. & Volín J. (2009): High Front Vowels in Czech: a Contrast in Quantity or Quality? *Proceedings of Interspeech 2009*, 132–135. Brighton: ISCA.
- Podlipský, V. J., Chládková, K. & Šimáčková, Š. (2012): Length or height? Front-back symmetry in the Czech vowel system. *Proceedings of LabPhon 13*, Stuttgart.
- Pollák P., Volín J. & Skarnitzl R. (2007): HMM-Based Phonetic Segmentation in Praat Environment. *Proceedings of the XIIth International Conference "Speech and computer – SPECOM 2007"*, 537–541, Moscow.
- Port, R. F., Al-Ani, S. & Maeda, S. (1980). Temporal compensation and universal phonetics. *Phonetica* 37, 235–252.
- Pytkönen, J. & Kurimo, M. (2004): Duration Modeling Techniques for Continuous Speech Recognition. *Proceedings of Interspeech 2004*, 385–388.
- Quené, H. & Port, R. (2005): Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica*, 62, 1–13.
- Ramus, F. (2002). Acoustic correlates of linguistic rhythm: Perspectives. *Proceedings of Speech Prosody 2002*, 115–120, Aix-en-Provence.
- Ramus, F., Nespore, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73/3, 265–292.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review*, 12, 969–992.
- Russell, M. J. & Cook, A. E. (1987): Experimental evaluation of duration modeling techniques for automatic speech recognition. In *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2376–2379.
- Skarnitzl, R. (2010): Prague Phonetic Corpus: status report. *AUC Philologica 1/2009, Phonetica Pragensia, XII*, 65–67.
- Skarnitzl, R. (2011): *Znělostní kontrast nejen v češtině*. Praha: Nakladatelství Epoque.
- Solé, M.-J. & Ohala, J. J. (2010): What is and what is not under the control of the speaker: intrinsic vowel duration. In: C. Fougerson, B. Kühnert, M. D'Imperio & N. Vallée (eds.): *Papers in Laboratory Phonology 10*, 607–655. Berlin: de Gruyter.
- Studenovský, D. (2012): *Akustické vlastnosti českých diftongů*. Disertační práce. Praha: Fonetický ústav.
- Svobodová, M. & Voříšek, L. (2014): Identifikace mluvčích z pohledu autentické kriminalistické praxe v České republice. In: R. Skarnitzl (ed.), *Fonetická identifikace mluvčího*. Praha: Vydavatelství FF UK.
- Šimek, J. (2010): *Explozivní v češtině: temporální vlastnosti a variabilita při realizaci*. Diplomová práce. Praha: Fonetický ústav.

- Šimko, J., Cummins, F. (2010). Embodied task dynamics. *Psychological Review*, 117/4, 1229–1246.
- Tajima, K. (1998): *Speech Rhythm in English and Japanese: Experiments in Speech Cycling*. Disertační práce. Bloomington: Indiana University.
- Trouvain, J. (2004). *Tempo Variation in Speech Production. Implications for Speech Synthesis*. Disertační práce. Saarbrücken: Universität Saarland.
- van Donzel, M. E. & Koopmans-van Beinum, F. J. (1996). Pausing strategies in discourse in Dutch. *Proceedings of ICSLP 1996*, 1029–1032.
- Vernerová, T. (2006): *Trvání slabikotvorných likvid v češtině*. Diplomová práce. Praha: Fonetický ústav.
- Volín, J. (2007): *Statistické metody ve fonetickém výzkumu*. Praha: Nakladatelství Epona.
- Volín, J. (2009): Metric warping in Czech newsreading. In: R. Vích (ed.), *Speech Processing – 19th Czech-German Workshop*, 52–55, Praha.
- Volín, J. (2010): On the significance of the temporal structuring of speech. In: M. Malá & P. Šaldová (eds.) ...for thy speech bewrayeth thee (A Festschrift for Libuše Dušková), 289–305. Praha: UK FF.
- Volín, J. (2012): Jak se v Čechách „rázuje“. *Naše řeč*, 95/1, 51–54.
- Volín, J., Churaňová, E. & Šturm, P. (2014): P-centre Position in Natural Two-Syllable Czech Words. In: N. Campbell, D. Gibbon & D. Hirst (eds.), *Proceedings of the 7th International Conference on Speech Prosody*, 920–924, Dublin.
- Volín, J. & Skarnitzl, R. (2007): Temporal downtrends in Czech read speech. *Proceedings of Interspeech 2007*, 442–445, Antwerpen.
- Volín, J. & Weingartová, L. (2012): Idiosyncrasies in local articulation rate trajectories in Czech. *Proceedings of Perspectives on Rhythm and Timing*, 67. Glasgow.
- Volín, J. & Weingartová, L. (2014): Acoustic correlates of word stress as a cue to accent strength. *Research in Language*, 12/2, 175–183.
- Volín, J. & Weingartová, L. (v přípravě): *Současný stav zkoumání zvukové stránky mluvních stylů*. Nepublikovaný rukopis.
- Weingartová, L. (2013): Rhythm metrics for speaker identification in Czech. *AUC Philologica 1/2014, Phonetica Pragensia XIII*, 33–42.
- Weingartová, L., Churaňová, E. & Šturm, P. (2014a): Transitions, pauses and overlaps: Temporal characteristics of turn-taking in Czech. In: N. Campbell, D. Gibbon & D. Hirst (eds.), *Proceedings of the 7th International Conference on Speech Prosody*, 502–506. Dublin: TCD.
- Weingartová, L., Poesová, K. & Volín, J. (2014b): Prominence contrasts in Czech English as a predictor of learner's proficiency. In: N. Campbell, D. Gibbon & D. Hirst (eds.), *Proceedings of the 7th International Conference on Speech Prosody*, 236–240. Dublin: TCD.
- Weingartová, L. & Volín, J. (2014): Temporální charakteristiky. In: R. Skarnitzl (ed.), *Fonetická identifikace mluvčího*. Praha: Vydavatelství FF UK.
- Wells, J. C. et al. (2003): Czech SAMPA Home Page. In: SAMPA - computer readable phonetic alphabet. University College London, 2003. Publikováno online: <http://www.phon.ucl.ac.uk/home/sampa/czech-uni.htm>.

- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., and Mattys, S. (2010) How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, 127, 1559-1569.
- Whalen, D. H. & Levitt, A. G. (1995): The universality of intrinsic f<sub>0</sub> of vowels. *Journal of Phonetics*, 23, 349–366.
- White, L. & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35, 501–522.
- Wretling, P. & Eriksson, A. (1998): Is articulatory timing speaker specific? – evidence from imitated voices. *Proc. FONETIK 98*, 48–52.
- Zheng, X. & Pierrehumbert, J. B. (2010): The effects of prosodic prominence and serial position on duration perception. *Journal of the Acoustical Society of America*, 128, 851–859

## Přílohy

### Příloha 1: Text korpusu Minidialogy-H

#### H1a

Já ti řeknu, co uděláš. Nejdřív najdeš Hučku a Atamana.

Dobře. A až je najdu?

Řekneš jim, co si myslíš.

To asi budou dost nadávat, co?

To bych se nebál.

#### H1b

Začni opatrně. A o další spolupráci bych se nezmiňoval.

To by mě ani nenapadlo.

Řekneš jim, co si myslíš?

Nevím, snad nebudu muset.

Dobře, nezapomeň, hlavní heslo: opatrnost.

#### H1c

Nejlepší bude, když zatroubíte a počkáte, až vylezou.

Hmm, no a potom?

Řeknete jim, co si myslíte.

A vy na nás počkáte?

Jasně, ani se neheme z místa.

#### H1d

Možná budou problémy. Promysleli jste si to pořádně?

Jo, krok za krokem.

Řeknete jim, co si myslíte?

Určitě, hnedka zkraje.

No, připravte se, že budou prskat.

#### H2a

Takže se posadiš a budeš se usmívat. Žádnou paniku.

Jasně. A co mám dělat, až přijde ten jejich šéf?

Zeptáš se, co bude dál.

To je všechno? Nemám mu říct, že už to víme?

Ne, nic nevíš. Ani slovo.

#### H2b

Už se někdo ozval?

Ne. Všichni to vědí, ale vesele se předstírá, že nic.

To je teda situace. Zeptáš se, co bude dál?

Budu muset. Jinak budeme předstírat a předstírat, až už bude pozdě.

Hmm, to ti nezavídím.

#### H2c

Když budou chtít napřed vidět peníze, řekněte jim, že jsou na cestě.

Jasně. A čím máme začít, až nás vezmou dovnitř?

Zeptáte se, co bude dál.

Myslíš, že v tom jedou všichni?

Určitě. Je to jejich priorit.

#### H2d

Máte nějakou představu, jak to bude probíhat? Zhruba. Od nich přijdou taky tři.

To by mohlo zaskřípat. Zeptáte se, co bude dál?

Zeptáme, ale ne hned zkraje. Musíme pomalu.

#### H3a

Až začnou o tom transportu, nastražíš uši.

Vždyť mě znáš. Jen mě trochu nasměruj. O co jde?

Zjistíš, v kolik to pojede.

No, to je snad samo sebou.

Jo, ale potřebujeme to přesně. Hodně přesně.

#### H3b

Tak, ještě kartáček na zuby a je to. Snad mám všechno.

Dobře, chceš ještě nějak pomoci?

Zjistíš, v kolik to pojede?

No, na ty informace jsem volal, ale pořád bylo obsazeno.

Na internetu to není?

#### H3c

Trochu se bojím, že je tím rozčílíte.

Aha, my asi potřebujeme, aby byli v klidu. O co půjde?

Zjistíte, v kolik to pojede.

No jo, ale to je to poslední, co by nám chtěli prozradit.

Právě proto žádný provokace. Klid.

#### H3d

Přestaňte se mi tu motat pod nohy. Nevím, co dřív.

Ale my jsme přišli pomáhat. A na to nádraží taky dojdeme.

Zjistíte, v kolik to pojede?

Dobře, něco po poledni, jo?

Jo. Určitě před druhou.

**H4a**

Moc se neošívěj, ale dělej, že to je pro tebe novinka.

Proč? Hrozí něco?

Nemáš poněťí, kdo pojede vzadu.

Jasně, chápu.

Hlavně buď bez obav, už jsme zvládli horší věci.

**H4b**

Hodilo by se trochu víc informací.

Já už nemám čas na to myslet.

Nemáš poněťí, kdo pojede vzadu?

Nevím. Doufám, že ne ten idiot Kukla.

To by byl kolosální malér. Jen ne Kukla.

**H4c**

Začíná mě mrazit v zádech. V jste v pohodě?

Jasně. Ty myslíš, že by se to mohlo zvrtnout?

Nemáte poněťí, kdo pojede vzadu.

No, to nemáme. Snad ne posily.

Když tak radši nic nezkoušejte. Ještě bude spousta šancí.

**H4d**

Kam bych to měl soustředit? Máte nákresy?

Kousek za půlku. A nebo prostředek.

Nemáte poněťí, kdo pojede vzadu?

Krmivo, pomocná síla a sanitka.

Jo, takže ne moc za půlku.

**H5a**

A pamatuj: o místě určení ani slovo.

Takže se nemám zapojovat?

Nedávej na sobě nic znát. Netušíš, kam to přesouvají.

Jasně. Mám je k tomu nějak nasměřovat?

Ne. Až se o tom začnou bavit, tvař se překvapeně.

**H5b**

Kdo ví, kam až se s tím povlečeme.

A to se to ještě může zkomplikovat kvůli ostraze.

Netušíš, kam to přesouvají?

Ne. Ale dělají s tím zbytečně tajnosti.

Jako by to bylo bůhví co.

**H5c**

Takže všechno to teď závisí na vás.

No jo, ale jak se k tomu máme stavět.

Nesmíte ani mrknout. Netušíte, kam to přesouvají.

Jo, to se lehce řekne. Nám se klepou ruce už teď, že jo, Frede?

Klid – copak jste začátečníci?

**H5d**

To je teda náklad'áků. To snad nemá konce.

Hmm. Na tohle si čas a prostředky najdou.

Netušíte, kam to přesouvají?

Někam za Rudnou. Ale tam se teď nikdo nedostane.

A ze vzduchu nic vidět není?

**H6a**

Tentokrát nemají šanci. Máš to v kapse.

Proč myslíš, že jsem ve výhodě?

Víš, kdy to dostaneš.

No, to mi pár minut získá. Ale jinak nevíme nic.

Já ti říkám, že je to v suchu.

**H6b**

Měl bys mít všechno po ruce. Bude zmatek.

No jo. Už jsem si to kontroloval nejmíň pětkrát.

Víš, kdy to dostaneš?

Hned jak Batulka zavře bránu.

Teda tam bych se nechtěl přimotat.

**H6c**

Máte jeden trumf, a na tom se dá vydělat.

Jak to?

Víte, kdy to dostanete.

No jo, to je pravda.

A právě s tím nikdo nepočítá.

**H6d**

Jen abyste měli dost peněz, až to přijde.

Bez obav. S tím se počítá.

Víte, kdy to dostanete?

No, někdy po druhý hodině. Až zmizí první směna.

Musíte hrát hodně opatrně. Nepřehánět sázky.

## Příloha 2: Fonémová transkripce korpusu Minidialogy-H

Pro přepis je využita česká transkripce SAMPA (viz Wells et al. 2003). Pomlčky označují hranice mezi značkami fonémů.

H1a\_1 j-a:-c-i-P\~e-k-n-u-t\_s-o-u-J\~e-l-a:-S-n-e-j-d-P\~i:-f-n-a-j-d-e-Z-h-u-t\_S-k-u-a-a-t-a-m-a-n-a  
H1a\_2 d-o-b-P\~e-a-a-S-j-e-n-a-j-d-u  
H1a\_3 P\~e-k-n-e-S-j-i-m-t\_s-o-s-i-m-i-s-l-i:-S  
H1a\_4 t-o-a-s-i-b-u-d-o\_u-d-o-s-t-n-a-d-a:-v-a-t-t\_s-o  
H1a\_5 t-o-b-i-x-s-e-n-e-b-a:-l  
H1b\_1 z-a-t\_S-J-i-o-p-a-t-r-J-e-a-o-d-a-l-S-i:-s-p-o-l-u-p-r-a:-t\_s-i-b-i-x-s-e-n-e-z-m-i-J-o-v-a-l  
H1b\_2 t-o-b-i-m-J-e-a-J-i-n-e-n-a-p-a-d-l-o  
H1b\_3 P\~e-k-n-e-S-j-i-m-t\_s-o-s-i-m-i-s-l-i:-S  
H1b\_4 n-e-v-i:-m-s-n-a-t-n-e-b-u-d-u-m-u-s-e-t  
H1b\_5 d-o-b-P\~e-n-e-z-a-p-o-m-e-J-h-l-a-v-J-i:-h-e-s-l-o-o-p-a-t-r-n-o-s-t  
H1c\_1 n-e-j-l-e-p-S-i:-b-u-d-e-g-d-i-Z-z-a-t-r-o\_u-b-i:-t-e-a-p-o-t\_S-k-a:-t-e-a-S-v-i-l-e-z-o\_u  
H1c\_2 hm-n-o-a-p-o-t-o-m  
H1c\_3 P\~e-k-n-e-t-e-j-i-m-t\_s-o-s-i-m-i-s-l-i:-t-e  
H1c\_4 a-v-i-n-a-n-a:-s-p-o-t\_S-k-a:-t-e  
H1c\_5 j-a-s-J-e-a-J-i-s-e-n-e-h-n-e-m-e-z-m-i:-s-t-a  
H1d\_1 m-o-Z-n-a:-b-u-d-o\_u-p-r-o-b-l-e:-m-i-p-r-o-m-i-s-l-e-l-i-j-s-t-e-s-i-t-o-p-o-P\~a:-d-J-e  
H1d\_2 j-o-k-r-o-g-z-a-k-r-o-k-e-m  
H1d\_3 P\~e-k-n-e-t-e-j-i-m-t\_s-o-s-i-m-i-s-l-i:-t-e  
H1d\_4 u-r-t\_S-i-c-e-h-n-e-t-k-a-s-k-r-a-j-e  
H1d\_5 n-o-p-Q\~i-p-r-a-f-t-e-s-e-Z-e-b-u-d-o\_u-p-r-s-k-a-t  
H2a\_1 t-a-g-Z-e-s-e-p-o-s-a-J\~i:-S-a-b-u-d-e-S-s-e-u-s-m-i:-v-a-t-Z-a:-d-n-o\_u-p-a-n-i-k-u  
H2a\_2 j-a-s-J-e-a-t\_s-o-m-a:-m-J\~e-l-a-t-a-S-p-Q\~i-j-d-e-t-e-n-j-e-j-i-x-S-e:-f  
H2a\_3 z-e-p-t-a:-S-s-e-t\_s-o-b-u-d-e-d-a:-l  
H2a\_4 t-o-j-e-f-S-e-x-n-o-n-e-m-a:-m-m-u-P\~i:-t\_s-t-Z-e-u-S-t-o-v-i:-m-e  
H2a\_5 n-e-J-i-t\_s-n-e-v-i:-S-a-J-i-s-l-o-v-o  
H2b\_1 u-S-s-e-J-e-g-d-o-o-z-v-a-l  
H2b\_2 n-e-f-S-i-x-J-i-t-o-v-j-e-J\~i:-a-l-e-v-e-s-e-l-e-s-e-p-Q\~e-t-s-c-i:-r-a:-Z-e-J-i-t\_s  
H2b\_3 t-o-j-e-t-e-d-a-s-i-t-u-a-t\_s-e-z-e-p-t-a:-S-s-e-t\_s-o-b-u-d-e-d-a:-l  
H2b\_4 b-u-d-u-m-u-s-e-t-j-i-n-a-g-b-u-d-e-m-e-p-Q\~e-t-s-c-i:-r-a-t-a-p-Q\~e-t-s-c-i:-r-a-t-a-S-u-Z-b-u-  
d-e-p-o-z-J\~e  
H2b\_5 hm-t-o-c-i-n-e-z-a:-v-i-J\~i:-m  
H2c\_1 g-d-i-Z-b-u-d-o\_u-x-c-i:-t-n-a-p-Q\~e-t-v-i-J\~e-t-p-e-J-i:-z-e-P\~e-k-J-e-t-e-j-i-m-Z-e-j-s-o\_u-n-  
a-t\_s-e-s-c-e  
H2c\_2 j-a-s-J-e-a-t\_S-i:-m-m-a:-m-e-z-a-t\_S-i:-t-a-S-n-a:-s-v-e-z-m-o\_u-d-o-v-J-i-t-Q\~  
H2c\_3 z-e-p-t-a:-t-e-s-e-t\_s-o-b-u-d-e-d-a:-l  
H2c\_4 m-i-s-l-i:-S-Z-e-f-t-o-m-j-e-d-o\_u-f-S-i-x-J-i  
H2c\_5 u-r-t\_S-i-c-e-j-e-t-o-j-e-j-i-x-p-r-i-j-o-r-i-t-a  
H2d\_1 m-a:-t-e-J-e-j-a-k-o\_u-p-Q\~e-t-s-t-a-v-u-j-a-k-t-o-b-u-d-e-p-r-o-b-i:-h-a-t  
H2d\_2 z-h-r-u-b-a-o-d-J-i-x-p-Q\~i-j-d-o\_u-t-a-k-i-t-Q\~i  
H2d\_3 t-o-b-i-m-o-h-l-o-z-a-s-k-Q\~i:-p-a-t-z-e-p-t-a:-t-e-s-e-t\_s-o-b-u-d-e-d-a:-l  
H2d\_4 z-e-p-t-a:-m-e-a-l-e-n-e-h-n-e-t-s-k-r-a-j-e-m-u-s-i:-m-e-p-o-m-a-l-u  
H3a\_1 a-Z-z-a-t\_S-n-o\_u-o-t-o-m-t-r-a-n-s-p-o-r-t-u-n-a-s-t-r-a-Z-i:-S-u-S-i  
H3a\_2 v-Z-d-i-c-m-J-e-z-n-a:-S-j-e-n-m-J-e-t-r-o-x-u-n-a-s-m-J-e-r-u-j-o-t\_s-o-j-d-e  
H3a\_3 z-j-i-s-c-i:-S-f-k-o-l-i-k-t-o-p-o-j-e-d-e  
H3a\_4 n-o-t-o-j-e-s-n-a-t-s-a-m-o-s-e-b-o\_u  
H3a\_5 j-o-a-l-e-p-o-t-Q\~e-b-u-j-e-m-e-t-o-p-Q\~e-s-J-e-h-o-d-J-e-p-Q\~e-s-J-e  
H3b\_1 t-a-k-j-e-S-c-e-k-a-r-t-a:-t\_S-e-k-n-a-z-u-b-i-a-j-e-t-o-s-n-a-t-m-a:-m-f-S-e-x-n-o  
H3b\_2 d-o-b-P\~e-x-t\_s-e-S-j-e-S-c-e-J-e-j-a-k-p-o-m-o-t\_s-t  
H3b\_3 z-j-i-s-c-i:-S-f-k-o-l-i-k-t-o-p-o-j-e-d-e  
H3b\_4 n-o-n-a-t-i-i-n-f-o-r-m-a-t\_s-e-j-s-e-m-v-o-l-a-l-a-l-e-p-o-P\~a:-d-b-i-l-o-o-p-s-a-z-e-n-o



H3b\_5 n-a-i-n-t-e-r-n-e-t-u-t-o-n-e-J-i:  
 H3c\_1 t-r-o-x-u-s-e-b-o-j-i:-m-Z-e-j-e-c-i:-m-r-o-s-t\_S-i:-l-i:-t-e  
 H3c\_2 a-h-a-m-i-a-s-i-p-o-t-Q\~e-b-u-j-e-m-e-a-b-i-b-i-l-i-f-k-l-i-d-u-o-t\_s-o-p-u:-j-d-e  
 H3c\_3 z-j-i-s-c-i:-t-e-f-k-o-l-i-k-t-o-p-o-j-e-d-e  
 H3c\_4 n-o-j-o-a-l-e-t-o-j-e-t-o-p-o-s-l-e-d-J-i:-t\_s-o-b-i-n-a:-m-x-c-e-l-i-p-r-o-z-r-a-J\~i-t  
 H3c\_5 p-r-a:-v-j-e-p-r-o-t-o-Z-a:-d-n-i:-p-r-o-v-o-k-a-t\_s-e-k-l-i-t  
 H3d\_1 p-Q\~e-s-t-a-J-t-e-s-e-m-i-t-u-m-o-t-a-t-p-o-d-n-o-h-i-n-e-v-i-m-t\_s-o-d-P\~i:-f  
 H3d\_2 a-l-e-m-i-j-s-m-e-p-Q\~i-S-l-i-p-o-m-a:-h-a-t-a-n-a-t-o-n-a:-d-r-a-Z-i:-t-a-k-i-d-o-j-d-e-m-e  
 H3d\_3 z-j-i-s-c-i:-t-e-f-k-o-l-i-k-t-o-p-o-j-e-d-e  
 H3d\_4 d-o-b-P\~e-J-e-t\_s-o-p-o-p-o-l-e-d-J-i-j-o  
 H3d\_5 j-o-u-r-t\_S-i-c-e-p-Q\~e-d-d-r-u-h-o\_u  
 H4a\_1 m-o-t\_s-s-e-n-e-o-S-i:-v-e-j-a-l-e-J\~e-l-e-j-Z-e-t-o-j-e-p-r-o-t-e-b-e-n-o-v-i-N-k-a  
 H4a\_2 p-r-o-t\_S-h-r-o-z-i:-J-e-t\_s-o  
 H4a\_3 n-e-m-a:-S-p-o-J-e-c-i:-g-d-o-p-o-j-e-d-e-v-z-a-d-u  
 H4a\_4 j-a-s-J-e-x-a:-p-u  
 H4a\_5 h-l-a-v-J-e-b-u-J\~b-e-s-o-b-a-f-u-S-j-s-m-e-z-v-l-a:-d-l-i-h-o-r-S-i:-v-j-e-t\_s-i  
 H4b\_1 h-o-J\~i-l-o-b-i-s-e-t-r-o-x-u-v-i:-t\_s-i-n-f-o-r-m-a-t\_s-i:  
 H4b\_2 j-a:-u-S-n-e-m-a:-m-t\_S-a-s-n-a-t-o-m-i-s-l-e-t  
 H4b\_3 n-e-m-a:-S-p-o-J-e-c-i:-g-d-o-p-o-j-e-d-e-v-z-a-d-u  
 H4b\_4 n-e-v-i:-m-d\_o\_u-f-a:-m-Z-e-n-e-t-e-n-i-d-i-j-o-t-k-u-k-l-a  
 H4b\_5 t-o-b-i-b-i-l-k-o-l-o-s-a:-l-J-i:-m-a-l-e:-r-j-e-n-n-e-k-u-k-l-a  
 H4c\_1 z-a-t\_S-i:-n-a:-m-J-e-m-r-a-z-i-t-v-z-a:-d-e-x-v-i-j-s-t-e-f-p-o-h-o-J\~e  
 H4c\_2 j-a-s-J-e-t-i-m-i-s-l-i:-S-Z-e-b-i-s-e-t-o-m-o-h-l-o-z-v-r-t-n-o\_u-t  
 H4c\_3 n-e-m-a:-t-e-p-o-J-e-c-i:-g-d-o-p-o-j-e-d-e-v-z-a-d-u  
 H4c\_4 n-o-t-o-n-e-m-a:-m-e-s-n-a-t-n-e-p-o-s-i-l-i  
 H4c\_5 g-d-i-S-t-a-k-r-a-t-S-i-J-i-t\_s-n-e-s-k-o\_u-S-e-j-t-e-j-e-S-c-e-b-u-d-e-s-p-o\_u-s-t-a-S-a-n-t\_s-i:  
 H4d\_1 k-a-m-b-i-x-t-o-m-J-e-l-s-o\_u-s-t-Q\~e-J\~i-t-m-a:-t-e-n-a:-k-r-e-s-i  
 H4d\_2 k-o\_u-s-e-g-z-a-p-u:-l-k-u-a-n-e-b-o-p-r-o-s-t-Q\~e-d-e-k  
 H4d\_3 n-e-m-a:-t-e-p-o-J-e-c-i:-g-d-o-p-o-j-e-d-e-v-z-a-d-u  
 H4d\_4 k-r-m-i-v-o-p-o-m-o-t\_s-n-a:-s-i:-l-a-a-s-a-n-i-t-k-a  
 H4d\_5 j-o-t-a-g-Z-e-n-e-m-o-d\_z-z-a-p-u:-l-k-u  
 H5a\_1 a-p-a-m-a-t-u-j-o-m-i:-s-c-e-u-r-t\_S-e-J-i:-a-J-i-s-l-o-v-o  
 H5a\_2 t-a-g-Z-e-s-e-n-e-m-a:-m-z-a-p-o-j-o-v-a-t  
 H5a\_3 n-e-d-a:-v-e-j-n-a-s-o-b-j-e-J-i-d\_z-z-n-a:-t-n-e-t-u-S-i:-S-k-a-m-t-o-p-Q\~e-s-o\_u-v-a-j-i:  
 H5a\_4 j-a-s-J-e-m-a:-m-j-e-k-t-o-m-u-J-e-j-a-k-n-a-s-m-J-e-r-o-v-a-t  
 H5a\_5 n-e-a-S-s-e-o-t-o-m-z-a-t\_S-n-o\_u-b-a-v-i-t-t-v-a-Q\~s-e-p-Q\~e-k-v-a-p-e-J-e  
 H5b\_1 g-d-o-v-i:-k-a-m-a-S-s-e-s-c-i:-m-p-o-v-l-e-t\_S-e-m-e  
 H5b\_2 a-t-o-s-e-t-o-j-e-S-c-e-m-u:-Z-e-s-k-o-m-p-l-i-k-o-v-a-t-k-v-u:-l-i-o-s-t-r-a-z-e  
 H5b\_3 n-e-t-u-S-i:-S-k-a-m-t-o-p-Q\~e-s-o\_u-v-a-j-i:  
 H5b\_4 n-e-a-l-e-J\~e-l-a-j-i:-s-c-i:-m-z-b-i-t-e-t\_S-J-e-t-a-j-n-o-s-c-i  
 H5b\_5 j-a-k-o-b-i-t-o-b-i-l-o-b-u:-x-v-i:-t\_s-o  
 H5c\_1 t-a-g-Z-e-f-S-e-x-n-o-t-o-t-e-J\~z-a:-v-i-s-i:-n-a-v-a:s  
 H5c\_2 n-o-j-o-a-l-e-j-a-k-s-e-k-t-o-m-u-m-a:-m-e-s-t-a-v-j-e-t  
 H5c\_3 n-e-s-m-i:-t-e-a-J-i-m-r-k-n-o\_u-t-n-e-t-u-S-i:-t-e-k-a-m-t-o-p-Q\~e-s-o\_u-v-a-j-i:  
 H5c\_4 j-o-t-o-s-e-l-e-x-t\_s-e-P\~e-k-n-e-n-a:-m-s-e-k-l-e-p-o\_u-r-u-t\_s-e-u-S-t-e-c-Z-e-j-o-f-r-e-d-e  
 H5c\_5 k-l-i-t-t\_s-o-p-a-k-j-s-t-e-z-a-t\_S-a:-t-e-t\_S-J-i:-t\_s-i  
 H5d\_1 t-o-j-e-t-e-d-a-n-a:-k-l-a-J\~a:-k-u:-t-o-s-n-a-t-n-e-m-a:-k-o-n-t\_s-e  
 H5d\_2 h-m-n-a-t-o-h-l-e-s-i-t\_S-a-s-a-p-r-o-s-t-Q\~e-t-k-i-n-a-j-d-o\_u  
 H5d\_3 n-e-t-u-S-i:-t-e-k-a-m-t-o-p-Q\~e-s-o\_u-v-a-j-i:  
 H5d\_4 J-e-k-a-m-z-a-r-u-d-n-o\_u-a-l-e-t-a-m-s-e-t-e-c-J-i-g-d-o-n-e-d-o-s-t-a-n-e  
 H5d\_5 a-z-e-v-z-d-u-x-u-J-i-t\_s-v-i-J\~e-t-n-e-J-i:  
 H6a\_1 t-e-n-t-o-k-r-a:-t-n-e-m-a-j-i:-S-a-n-t\_s-i-m-a:-S-t-o-f-k-a-p-s-e  
 H6a\_2 p-r-o-t\_S-m-i-s-l-i:-S-Z-e-j-s-e-m-v-e-v-i:-h-o-J\~e  
 H6a\_3 v-i:-Z-g-d-i-t-o-d-o-s-t-a-n-e-S (Pozn.: Jednoslabičná hlavní věta, nepředpokládáme prozodický předěl.)

H6a\_4 n-o-t-o-m-i-p-a:-r-m-i-n-u-t-z-i:-s-k-a:-a-l-e-j-i-n-a-k-n-e-v-i:-m-e-J-i-t\_s  
H6a\_5 j-a:-c-i-P\i:-k-a:-m-Z-e-j-e-t-o-f-s-u-x-u  
H6b\_1 m-J-e-l-b-i-s-m-i:-t-f-S-e-x-n-o-p-o-r-u-t\_s-e-b-u-d-e-z-m-a-t-e-k  
H6b\_2 n-o-j-o-u-S-j-s-e-m-s-i-t-o-k-o-n-t-r-o-l-o-v-a-l-n-e-j-m-i:-J-p-j-e-t-k-r-a:t  
H6b\_3 v-i:-Z-g-d-i-t-o-d-o-s-t-a-n-e-S  
H6b\_4 h-n-e-t-j-a-g-b-a-t-u-l-k-a-z-a-v-P\e-b-r-a:-n-u  
H6b\_5 t-e-d-a-t-a-m-b-i-x-s-e-n-e-x-c-e-l-p-Q\i-m-o-t-a-t  
H6c\_1 m-a:-t-e-j-e-d-e-n-t-r-u-m-f-a-n-a-t-o-m-s-e-d-a:-v-i-J\e-l-a-t  
H6c\_2 j-a-k-t-o  
H6c\_3 v-i:-t-e-g-d-i-t-o-d-o-s-t-a-n-e-t-e  
H6c\_4 n-o-j-o-t-o-j-e-p-r-a-v-d-a  
H6c\_5 a-p-r-a:-v-j-e-s-c-i:-m-J-i-g-d-o-n-e-p-o-t\_S-i:-t-a:  
H6d\_1 j-e-n-a-b-i-s-t-e-m-J-e-l-i-d-o-s-t-p-e-J-e-s-a-S-t-o-p-Q\i-j-d-e  
H6d\_2 b-e-s-o-b-a-f-s-c-i:-m-s-e-p-o-t\_S-i:-t-a:  
H6d\_3 v-i:-t-e-g-d-i-t-o-d-o-s-t-a-n-e-t-e  
H6d\_4 n-o-J-e-g-d-i-p-o-d-r-u-h-i:-h-o-J\i-J-e-a-Z-z-m-i-z-i:-p-r-v-J-i:-s-m-J-e-n-a  
H6d\_5 m-u-s-i:-t-e-h-r-a:-d-h-o-d-J-e-o-p-a-t-r-J-e-n-e-p-Q\e-h-a:-J-e-t-s-a:-s-k-i

### Příloha 3: Sylabifikace textu korpusu Minidialogy-H

Pro přepis je využita česká transkripce SAMPA (viz Wells et al. 2003). Slabičné hranice jsou naznačeny pomlčkami.

H1a\_1 ja-ci-P\ek-nu-t\_so-?u-J\e-la:S-nej-dP\i:f-naj-deS-h\ut\_S-ku-?a-?a-ta-ma-na  
H1a\_2 do-bP\e-?a-?aS-je-naj-du  
H1a\_3 P\ek-neS-jim-t\_so-si-mis-li:S  
H1a\_4 to-?a-si-bu-do\_u-dost-na-da:-vat-t\_so  
H1a\_5 to-bix-se-ne-ba:l  
H1b\_1 zat\_S-Ji-?o-pa-tr-Je-?a-?o-dal-Si:-spo-lu-pra:-t\_si-bix-se-ne-zmi-Jo-val  
H1b\_2 to-bi-mJe-a-Ji-ne-na-pad-lo  
H1b\_3 P\ek-neS-jim-t\_so-si-mis-li:S  
H1b\_4 ne-vim-snat-ne-bu-du-mu-set  
H1b\_5 do-bP\e-ne-za-po-meJ-h\lav-Ji:-h\e-slo-?o-pa-tr-nost  
H1c\_1 nej-lep-Si:-bu-de-gdiZ-za-tro\_u-bi:-te-a-pot\_S-ka:-te-aS-vi-le-zo\_u  
H1c\_2 h\m-no-a-po-tom  
H1c\_3 P\ek-ne-te-jim-t\_so-si-mis-li:-te  
H1c\_4 ?a-vi-na-na:s-pot\_S-ka:-te  
H1c\_5 jas-Je-?a-Ji-se-ne-h\ne-me-zmi:-sta  
H1d\_1 moZ-na:-bu-do\_u-pro-ble:-mi-pro-mis-le-li-jste-si-to-po-P\ a:d-Je  
H1d\_2 jo-krog-za-kro-kem  
H1d\_3 P\ek-ne-te-jim-t\_so-si-mis-li:-te  
H1d\_4 ?ur-t\_Si-ce-h\net-ka-skra-je  
H1d\_5 no-pQ\i-praf-te-se-Ze-bu-do\_u-prs-kat  
H2a\_1 tag-Ze-se-po-sa-J\i:S-?a-bu-deS-se-?u-smi:-vat-Za:-dno\_u-pa-ni-ku  
H2a\_2 jas-Je-?a-t\_so-ma:m-J\e-lat-?aS-pQ\i:j-de-ten-je-jix-Se:f  
H2a\_3 ze-pta:S-se-t\_so-bu-de-da:l  
H2a\_4 to-je-fSex-no-ne-ma:m-mu-P\i:t\_st-Ze-uS-to-vi:-me  
H2a\_5 ne-Jit\_s-ne-vi:S-?a-Ji-slo-vo  
H2b\_1 uS-se-Je-gdo-?o-zval  
H2b\_2 ne-fSix-Ji-to-vje-J\i:-a-le-ve-se-le-se-pQ\et-sci:-ra:-Ze-Jit\_s  
H2b\_3 to-je-te-da-si-tu-a-t\_se-ze-pta:S-se-t\_so-bu-de-da:l  
H2b\_4 bu-du-mu-set-ji-nag-bu-de-me-pQ\et-sci:-rat-?a-pQ\et-sci:-rat-?aS-uZ-bu-de-poz-J\ e  
H2b\_5 h\m-to-ci-ne-za:-vi-J\i:m  
H2c\_1 gdiZ-bu-do\_u-xci:t-na-pQ\et-vi-J\et-pe-Ji:-ze-P\ek-Je-te-jim-Ze-jso\_u-na-t\_ses-ce  
H2c\_2 jas-Je-?a-t\_Si:m-ma:-me-za-t\_Si:t-aS-na:s-vez-mo\_u-do-vJitQ\  
H2c\_3 ze-pta:-te-se-t\_so-bu-de-da:l  
H2c\_4 mis-li:Z-Ze-ftom-je-do\_u-fSix-Ji  
H2c\_5 ?ur-t\_Si-ce-je-to-je-jix-pri-jo-ri-ta  
H2d\_1 ma:-te-Je-ja-ko\_u-pQ\et-sta-vu-jak-to-bu-de-pro-bi:-h\at  
H2d\_2 zh\ru-ba-?od-Jix-pQ\i:j-do\_u-ta-ki-tQ\  
H2d\_3 to-bi-mo-h\lo-zas-kP\i:-pat-ze-pta:-te-se-t\_so-bu-de-da:l  
H2d\_4 ze-pta:-me-?a-le-ne-h\net-skra-je-mu-si:-me-po-ma-lu  
H3a\_1 ?aZ-zat\_S-no\_u-?o-tom-tran-spor-tu-na-stra-Zi:S-?u-Si  
H3a\_2 vZdic-mJe-zna:S-jen-mJe-tro-xu-na-smJe-ruj-?o-t\_so-jde  
H3a\_3 zjis-ci:S-fko-lik-to-po-je-de  
H3a\_4 no-to-je-snat-sa-mo-se-bo\_u  
H3a\_5 jo-?a-le-po-tQ\ e-bu-je-me-to-pQ\es-Je-h\od-Je-pQ\es-t  
H3b\_1 tak-je-Sce-kar-ta:-t\_Sek-na-zu-bi-?a-je-to-snat-ma:m-fSex-no  
H3b\_2 do-bP\e-xt\_seS-jeS-ce-Je-jak-po-mot\_st  
H3b\_3 zjis-ci:S-fko-lik-to-po-je-de  
H3b\_4 no-na-ti-in-for-ma-t\_se-jsem-vo-lal-a-le-po-P\ a:d-bi-lo-?op-sa-ze-no  
H3b\_5 na-?in-ter-ne-tu-to-ne-Ji:  
H3c\_1 tro-xu-se-bo-ji:m-Ze-je-ci:m-ros-t\_Si:-li:-te

H3c\_2 ?a-h\|a-mi-a-si-po-tQ\|e-bu-je-me-a-bi-bi-li-fkli-du-?o-t\_so-pu:j-de  
 H3c\_3 zjis-ci:-te-fko-lik-to-po-je-de  
 H3c\_4 no-jo:-a-le-to-je-to-po-sled-Ji:-t\_so-bi-na:m-xce-li-pro-zra-J\|it  
 H3c\_5 pra:-vje-pro-to-Za:-dni:-pro-vo-ka-t\_se-klit  
 H3d\_1 pQ\|e-staJ-te-se-mi-tu-mo-tat-pod-no-h\|i-ne-vim-t\_so-dP\|i:f  
 H3d\_2 ?a-le-mi-jsme-pQ\|i-Sli-po-ma:-h\|at-?a-na-to-na:-dra-Zi:-ta-ki-doj-de-me  
 H3d\_3 zjis-ci:-te-fko-lik-to-po-je-de  
 H3d\_4 do-bP\|e-Je-t\_so-po-po-leJ\|Ji-jo  
 H3d\_5 jo-?ur-t\_Si-ce-pQ\|ed-dru-h\|o\_u  
 H4a\_1 mot\_s-se-ne-?o-Si:-vej-?a-le-J\|e-lej-Ze-to-je-pro-te-be-no-viN-ka  
 H4a\_2 prot\_S-h\|ro-zi:-Je-t\_so  
 H4a\_3 ne-ma:S-po-Je-ci:-gdo-po-je-de-vza-du  
 H4a\_4 jas-Je-xa:-pu  
 H4a\_5 h\|lav-Je-buJ\|bes-?o-baf-?uS-jsme-zvla:-dli-h\|or-Si:-vje-t\_si  
 H4b\_1 h\|o-J\|i-lo-bi-se-tro-xu-vi:t\_s-?in-for-ma-t\_si:  
 H4b\_2 ja:-uS-ne-ma:m-t\_Sas-na-to-mis-let  
 H4b\_3 ne-ma:S-po-Je-ci:-gdo-po-je-de-vza-du  
 H4b\_4 ne-vi:m-do\_u-fa:m-Ze-ne-ten-?i-di-jot-ku-kla  
 H4b\_5 to-bi-bil-ko-lo-sa:l-Ji:-ma-le:r-jen-ne-ku-kla  
 H4c\_1 za-t\_Si:-na:-mJe-mra-zit-vza:-dex-vi-jste-fpo-h\|o-J\|e  
 H4c\_2 jas-Je-ti-mis-li:Z-Ze-bi-se-to-mo-h\|lo-zvrt-no\_ut  
 H4c\_3 ne-ma:-te-po-Je-ci:-gdo-po-je-de-vza-du  
 H4c\_4 no-to-ne-ma:-me-snat-ne-po-si-li  
 H4c\_5 gdiS-tak-rat-Si-Jit\_s-ne-sko\_u-Sej-te-jeS-ce-bu-de-spo\_u-sta-San-t\_si:  
 H4d\_1 kam-bix-to-mJel-so\_u-stQ\|e-J\|it-ma:-te-na:-kre-si  
 H4d\_2 ko\_u-sek-za-pu:l-ku-a-ne-bo-pro-stQ\|e-dek  
 H4d\_3 ne-ma:-te-po-Je-ci:-gdo-po-je-de-vza-du  
 H4d\_4 kr-mi-vo-po-mot\_s-na:-si:-la-?a-sa-nit-ka  
 H4d\_5 jo-tag-Ze-ne-mod\_z-za-pu:l-ku  
 H5a\_1 ?a-pa-ma-tuj-?o-mi:-sce-?ur-t\_Se-Ji:-?a-Ji-slo-vo  
 H5a\_2 tag-Ze-se-ne-ma:m-za-po-jo-vat  
 H5a\_3 ne-da:-vej-na-so-bje-Jid\_z-zna:t-ne-tu-Si:S-kam-to-pQ\|e-so\_u-va-ji:  
 H5a\_4 jas-Je-ma:m-je-kto-mu-Je-jak-na-smJe-ro-vat  
 H5a\_5 ne-?aS-se-?o-tom-zat\_S-no\_u-ba-vit-tvaQ\|se-pQ\|e-kva-pe-Je  
 H5b\_1 gdo-vi:-kam-?aS-se-sci:m-po-vle-t\_Se-me  
 H5b\_2 ?a-to-se-to-jeS-ce-mu:-Ze-skom-pli-ko-vat-kvu:-li-?o-stra-ze  
 H5b\_3 ne-tu-Si:S-kam-to-pQ\|e-so\_u-va-ji:  
 H5b\_4 ne-a-le-J\|e-la-ji:-scim-zbi-tet\_S-Je-taj-no-sci  
 H5b\_5 ja-ko-bi-to-bi-lo-bu:x-vi:-t\_so  
 H5c\_1 tag-Ze-fSex-no-to-teJ\|za:-vi-si:-na-va:s  
 H5c\_2 no-jo-a-le-jak-se-kto-mu-ma:-me-sta-vjet  
 H5c\_3 ne-smi:-te-?a-Ji-mrk-no\_ut-ne-tu-Si:-te-kam-to-pQ\|e-so\_u-va-ji:  
 H5c\_4 jo-to-se-le-xt\_se-P\|ek-ne-na:m-se-kle-po\_u-ru-t\_se-uS-teJ\|Ze-jo-fre-de  
 H5c\_5 klit-t\_so-pak-jste-za-t\_Sa:-te-t\_SJi:-t\_si  
 H5d\_1 to-je-te-da-na:-kla-J\|a:-ku:-to-snat-ne-ma:-kon-t\_se  
 H5d\_2 h\|m-na-to-h\|le-si-t\_Sas-a-pro-stQ\|et-ki-naj-do\_u  
 H5d\_3 ne-tu-Si:-te-kam-to-pP\|e-so\_u-va-ji:  
 H5d\_4 Je-kam-za-ru-dno\_u-?a-le-tam-se-tec-Jig-do-ne-do-sta-ne  
 H5d\_5 ?a-ze-vzdu-xu-Jit\_s-vi-J\|et-ne-Ji:  
 H6a\_1 ten-to-kra:t-ne-ma-ji:-San-t\_si-ma:S-to-fka-pse  
 H6a\_2 prot\_S-mis-li:Z-Ze-jsem-ve-vi:-h\|o-J\|e  
 H6a\_3 vi:Z-gdi-to-do-sta-neS  
 H6a\_4 no-to-mi-pa:r-mi-nut-zi:-ska:-?a-le-ji-nak-ne-vi:-me-Jit\_s  
 H6a\_5 ja:-ci-P\|i:-ka:m-Ze-je-to-fsu-xu  
 H6b\_1 mJel-bis-mi:t-fSex-no-po-ru-t\_se-bu-de-zma-tek

H6b\_2 no-jo:-?uS-jsem-si-to-kon-tro-lo-val-nej-mi:J-pjet-kra:t  
H6b\_3 vi:Z-gdi-to-do-sta-neS  
H6b\_4 h\net-jak-ba-tul-ka-zav-P\|e-bra:-nu  
H6b\_5 te-da-tam-bix-se-nex-cel-pP\|i-mo-tat  
H6c\_1 ma:-te-je-den-trumf-?a-na-tom-se-da:-vi-J\|e-lat  
H6c\_2 jak-to  
H6c\_3 vi:-te-gdi-to-do-sta-ne-te  
H6c\_4 no-jo-to-je-prav-da  
H6c\_5 pra:-vje-sci:m-Jig-do-ne-po-t\_Si:-ta:  
H6d\_1 jen-a-bi-ste-mJe-li-dost-pe-Jes-aS-to-pP\|i:j-de  
H6d\_2 bes-?o-baf-scim-se-po-t\_Si:-ta:  
H6d\_3 vi:-te-gdi-to-do-sta-ne-te  
H6d\_4 no:-Je-gdi-po-dru-h\|i:-h\|o-J\|i-Je-aS-zmi-zi:-pr-vJi:-smJe-na  
H6d\_5 mu-si:-te-h\|ra:t-h\|od-Je-o-pa-tr-Je-ne-pP\|e-h\|a:-Jet-sa:-ski

## Příloha 4: Tabulky deskriptivních statistik českých hlásek

**4.1 Tabulka hodnot trvání českých vokálů.** V šedě podbarvených řádcích jsou hodnoty pro všechny výskyty dané hlásky v korpusu Minidialogy-H. V nepodbarvených řádcích jsou tyto výskyty rozděleny podle pozice vokálu ve slově. Hodnoty ve třetím až devátém sloupci jsou uvedeny v milisekundách.

Hlásky / Pozice ve slově	Počet	Průměr	Medián	Min.	Max.	10. perc.	90. perc.	Sm.odch.
<b>a</b>	7777	<b>63.03</b>	59.79	9.91	295.89	37.35	90.14	24.68
iniciální	287	<b>62.61</b>	61.14	24.09	131.69	43.38	85.04	16.67
mediální	5699	<b>62.46</b>	59.52	11.63	295.89	38.26	88.28	22.88
finální	1686	<b>65.16</b>	60.40	9.91	284.21	32.59	99.88	31.18
individuální	105	<b>61.26</b>	60.29	30.36	104.94	44.55	84.21	15.83
<b>a:</b>	2694	<b>112.49</b>	105.16	20.61	360.00	69.93	165.86	39.59
mediální	2285	<b>112.99</b>	104.86	25.10	360.00	71.28	166.92	39.67
finální	408	<b>109.60</b>	107.01	20.61	256.59	62.69	162.64	39.11
<b>e</b>	13876	<b>59.68</b>	53.83	7.52	250.13	34.20	91.20	26.48
iniciální	35	<b>45.99</b>	46.57	20.11	70.31	34.04	62.15	12.24
mediální	6882	<b>55.47</b>	53.03	8.95	232.00	34.29	79.07	19.01
finální	6863	<b>63.94</b>	54.72	7.52	250.13	34.09	108.03	31.81
individuální	96	<b>61.10</b>	56.88	24.81	126.81	36.87	92.47	21.45
<b>e:</b>	227	<b>126.29</b>	115.66	44.53	271.60	83.56	181.03	39.67
mediální	102	<b>129.46</b>	125.99	44.53	245.86	93.45	170.28	32.64
finální	125	<b>123.70</b>	107.07	66.37	271.60	81.32	190.03	44.55
<b>i</b>	5478	<b>52.94</b>	48.00	6.98	230.73	29.98	80.61	24.21
iniciální	61	<b>45.55</b>	45.36	26.33	65.51	30.37	59.49	10.18
mediální	3164	<b>51.11</b>	46.55	8.24	229.07	29.11	77.10	22.83
finální	2250	<b>55.71</b>	49.87	6.98	230.73	31.15	86.31	26.02
<b>i:</b>	3273	<b>66.70</b>	61.22	7.68	336.00	36.56	101.93	29.71
mediální	2214	<b>64.94</b>	61.28	7.68	204.87	36.60	96.52	25.76
finální	1059	<b>70.39</b>	61.10	13.34	336.00	36.25	117.13	36.35
<b>o</b>	8374	<b>55.38</b>	50.30	12.89	303.76	32.00	79.95	25.98
iniciální	100	<b>57.96</b>	57.16	21.76	109.36	37.49	78.45	16.43
mediální	4375	<b>53.56</b>	51.33	12.89	244.53	32.72	75.20	19.29
finální	3854	<b>57.35</b>	48.85	15.19	303.76	31.51	92.54	32.00
individuální	45	<b>56.51</b>	50.67	31.58	167.50	37.65	78.94	22.71
<b>u</b>	3017	<b>55.33</b>	49.30	8.82	228.96	30.44	86.24	26.33
iniciální	167	<b>54.84</b>	54.53	8.82	120.73	35.16	74.67	16.25
mediální	1902	<b>48.24</b>	45.86	11.38	154.50	29.13	70.44	17.83
finální	948	<b>69.64</b>	58.94	13.12	228.96	34.00	121.59	34.88
<b>u:</b>	235	<b>70.05</b>	65.54	22.23	197.37	38.48	107.52	29.74
mediální	201	<b>62.38</b>	60.96	22.23	153.41	37.66	87.52	20.75
finální	34	<b>115.38</b>	112.41	40.91	197.37	84.68	158.02	34.45
<b>ou</b>	1020	<b>101.92</b>	91.71	32.46	342.18	63.51	159.10	39.18
mediální	407	<b>98.74</b>	92.46	45.08	241.48	66.53	134.76	30.61
finální	613	<b>104.03</b>	91.54	32.46	342.18	61.59	171.53	43.85

**4.2 Tabulka hodnot trvání českých konsonantů.** V šedě podbarvených řádcích jsou hodnoty pro všechny výskyty dané hlásky v korpusu Minidialogy-H. V nepodbarvených řádcích jsou tyto výskyty rozděleny podle pozice konsonantu ve slabice. Hodnoty ve třetím až devátém sloupci jsou uvedeny v milisekundách.

Hláška / Pozice ve slabice	Počet	Průměr	Medián	Min	Max	Sm. odch.	10. perc.	90. perc.
<b>b</b> <b>prétura</b>	2095	<b>59.41</b>	59.48	8.74	150.84	16.74	37.63	80.00
<b>c</b>	1427	<b>106.37</b>	98.42	29.39	263.74	36.44	67.33	156.72
<b>prétura</b>	1133	<b>102.26</b>	96.54	29.39	248.52	30.73	66.01	146.92
<b>koda</b>	294	<b>122.17</b>	107.50	30.17	263.74	49.97	71.84	202.02
<b>č</b>	843	<b>107.15</b>	105.05	43.58	229.39	24.09	79.08	137.46
<b>prétura</b>	544	<b>107.60</b>	106.98	43.58	181.85	23.73	78.42	137.74
<b>koda</b>	299	<b>106.20</b>	103.32	48.62	229.39	24.55	80.21	132.82
<b>d</b>	3614	<b>38.98</b>	36.94	8.82	149.93	14.74	22.69	56.82
<b>prétura</b>	3305	<b>38.49</b>	36.52	8.82	149.93	14.73	22.37	56.05
<b>koda</b>	309	<b>44.21</b>	42.54	16.99	105.13	13.75	28.08	61.50
<b>d'</b>	898	<b>57.08</b>	55.99	12.97	160.14	18.87	33.36	81.15
<b>prétura</b>	612	<b>59.71</b>	58.35	12.97	160.14	18.68	36.43	82.23
<b>koda</b>	286	<b>51.44</b>	49.46	15.09	128.94	18.03	30.36	72.75
<b>dz koda</b>	99	<b>64.73</b>	63.18	27.07	138.65	21.24	39.06	91.66
<b>dž koda</b>	23	<b>63.68</b>	52.89	33.41	127.46	23.21	37.77	98.21
<b>f</b>	874	<b>69.34</b>	61.48	7.71	233.17	36.16	31.06	123.44
<b>prétura</b>	642	<b>58.74</b>	56.14	7.71	150.84	25.26	28.87	94.22
<b>koda</b>	232	<b>98.68</b>	95.38	24.00	233.17	44.67	43.98	162.09
<b>g</b>	781	<b>46.82</b>	45.81	7.07	132.43	16.09	28.65	66.22
<b>prétura</b>	496	<b>48.72</b>	48.19	7.07	132.43	16.45	29.72	66.76
<b>koda</b>	285	<b>43.51</b>	42.50	7.86	89.43	14.90	25.97	65.29
<b>h</b>	940	<b>58.32</b>	55.69	11.31	170.35	21.49	33.59	87.31
<b>prétura</b>	937	<b>58.29</b>	55.59	11.31	170.35	21.52	33.53	87.35
<b>koda</b>	3	<b>65.22</b>	65.19	62.21	68.25	3.02	62.21	68.25
<b>j</b>	3077	<b>39.02</b>	34.53	7.02	179.33	20.59	17.96	64.89
<b>prétura</b>	2556	<b>36.43</b>	32.04	7.02	157.95	19.16	17.05	61.31
<b>koda</b>	521	<b>51.72</b>	47.67	10.08	179.33	22.55	28.53	79.80
<b>k</b>	2781	<b>77.36</b>	72.92	20.43	361.06	27.78	48.00	113.15
<b>prétura</b>	1981	<b>79.98</b>	76.45	20.43	273.99	24.47	53.54	113.04
<b>koda</b>	800	<b>70.89</b>	62.45	21.59	361.06	33.79	40.05	114.21
<b>l</b>	3074	<b>46.14</b>	42.23	7.08	192.00	19.89	25.41	71.07
<b>prétura</b>	2461	<b>41.94</b>	40.00	7.08	146.96	15.10	24.62	61.62
<b>koda</b>	613	<b>63.02</b>	58.41	20.28	192.00	26.72	32.38	99.71
<b>m</b>	4846	<b>58.01</b>	56.94	10.67	264.00	20.95	33.56	80.19
<b>prétura</b>	3309	<b>56.11</b>	56.96	10.67	161.49	17.92	32.08	78.19
<b>koda</b>	1537	<b>62.10</b>	56.93	17.02	264.00	25.85	37.70	91.04
<b>m̩ koda</b>	27	<b>83.06</b>	83.47	32.00	130.06	23.79	56.20	117.31
<b>n</b>	4780	<b>45.13</b>	42.10	7.10	327.55	20.40	24.14	67.63
<b>prétura</b>	4272	<b>43.88</b>	41.16	7.10	327.55	20.15	24.00	64.00
<b>koda</b>	508	<b>55.71</b>	53.67	16.26	115.81	19.44	31.84	81.15

<b>ň</b>	2504	<b>53.32</b>	50.48	4.04	176.00	21.46	28.98	80.72
<b>prétura</b>	2403	<b>52.88</b>	50.14	4.04	176.00	21.18	28.88	80.14
<b>koda</b>	101	<b>63.91</b>	59.34	25.59	168.71	25.29	37.85	92.84
<b>ŋ koda</b>	34	<b>86.86</b>	88.56	50.16	114.41	16.38	68.27	111.52
<b>p</b>	3474	<b>74.85</b>	73.41	19.91	187.08	20.45	48.59	100.80
<b>prétura</b>	3406	<b>75.08</b>	73.72	19.91	187.08	20.43	48.80	101.04
<b>koda</b>	68	<b>63.68</b>	61.01	38.30	114.44	18.08	41.57	89.64
<b>r neslabičné</b>	2241	<b>45.90</b>	44.51	15.58	222.98	13.58	32.00	60.54
<b>prétura</b>	1833	<b>45.28</b>	44.53	15.82	110.07	10.97	32.20	59.27
<b>koda</b>	408	<b>48.73</b>	44.47	15.58	222.98	21.53	29.82	69.60
<b>r slabičné</b>	273	<b>71.15</b>	71.86	23.85	117.79	14.63	51.59	88.42
<b>ráz</b>	1292	<b>63.85</b>	57.73	9.64	209.36	31.74	28.91	106.59
<b>ř</b>	1629	<b>64.09</b>	62.37	16.47	237.33	18.97	41.95	87.12
<b>prétura</b>	1561	<b>63.18</b>	62.05	16.47	132.24	17.21	41.49	85.61
<b>koda</b>	68	<b>84.85</b>	76.13	40.94	237.33	37.26	47.75	125.69
<b>s</b>	5420	<b>89.39</b>	86.17	17.44	263.11	31.50	52.01	129.99
<b>prétura</b>	4501	<b>86.52</b>	84.68	17.44	255.42	28.16	51.05	123.50
<b>koda</b>	919	<b>103.47</b>	96.16	28.60	263.11	41.61	55.12	160.84
<b>š</b>	2146	<b>85.42</b>	79.13	13.95	246.64	38.15	43.22	141.37
<b>prétura</b>	809	<b>89.90</b>	88.00	13.95	199.25	30.92	52.46	133.41
<b>koda</b>	1337	<b>82.71</b>	72.82	19.72	246.64	41.71	40.01	152.79
<b>t</b>	7105	<b>71.97</b>	67.24	11.58	253.70	27.93	41.22	107.87
<b>prétura</b>	5193	<b>67.69</b>	65.88	14.01	189.76	20.92	42.47	94.57
<b>koda</b>	1912	<b>83.58</b>	74.73	11.58	253.70	39.07	39.05	137.51
<b>ť</b>	1166	<b>73.62</b>	72.00	19.74	170.38	20.51	48.65	99.85
<b>prétura</b>	1086	<b>73.80</b>	72.00	19.74	148.34	19.73	50.56	99.24
<b>koda</b>	80	<b>71.07</b>	63.74	24.00	170.38	29.13	40.38	117.56
<b>v</b>	2397	<b>45.98</b>	44.69	10.91	149.00	14.78	28.33	64.09
<b>prétura</b>	2256	<b>45.71</b>	44.29	10.91	149.00	14.83	27.84	63.96
<b>koda</b>	141	<b>50.37</b>	49.99	23.86	93.17	13.33	33.10	68.02
<b>x</b>	859	<b>74.98</b>	68.95	11.83	253.41	32.84	39.33	117.14
<b>prétura</b>	370	<b>75.33</b>	73.32	11.83	173.18	29.25	39.36	113.20
<b>koda</b>	489	<b>74.72</b>	64.38	22.22	253.41	35.35	39.26	122.06
<b>z</b>	1830	<b>59.99</b>	58.50	9.15	191.07	19.91	36.94	84.22
<b>prétura</b>	1763	<b>59.65</b>	58.33	9.15	191.07	19.56	36.94	83.64
<b>koda</b>	67	<b>68.91</b>	64.65	23.49	131.54	26.15	37.32	106.27
<b>ž</b>	1083	<b>56.49</b>	55.59	11.66	151.89	21.05	30.03	82.60
<b>prétura</b>	674	<b>60.08</b>	58.80	11.66	151.89	20.90	35.12	85.72
<b>koda</b>	409	<b>50.58</b>	48.21	12.36	118.92	19.96	25.86	77.85



## Příloha 5: Koeficienty temporálního modelu

**5.1 Koeficienty pozice ve slabice.** Vokály nabývají pouze hodnoty „nukleus“, nejsou zde tedy zobrazeny. Extrapolované hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob extrapolace.

Hláška	Koeficienty		Trvání (ms)		Poznámka
	Prétura	Koda	Prétura	Koda	
p	1.03	0.96	80.7	74.9	podle f a rozdílu mezi výchozími hodnotami
b	1.00	1.06	58.7	62.5	podle p a rozdílu mezi t a d
t	1.00	0.83	68.4	56.2	
d	0.92	1.31	30.7	43.8	
ť	1.16	0.84	83.6	60.5	
d'	1.10	0.88	57.1	45.8	
k	1.15	0.87	73.8	56.2	
g	1.09	1.05	47.3	45.4	podle k a rozdílu mezi ť a d'
f	1.37	0.95	81.9	56.4	
v	1.05	0.96	43.7	40.1	podle f a rozdílu mezi znělými a neznělými (protože rozdíl s-z vede k příliš extrémní hodnotě)
s	1.15	1.00	93.2	81.2	
z	1.15	0.91	62.6	49.5	
š	1.16	0.82	96.2	68.4	
ž	1.20	0.92	63.6	48.7	
ch	1.26	0.81	83.2	53.5	
h	0.98	0.61	54.2	33.8	podle ch a rozdílu mezi š a ž
c	1.05	1.04	92.1	91.2	
dz	1	1			nemůže být v préture
č	1.03	0.93	104.8	94.1	
dž	1.05	0.88	67.0	56.3	podle č a rozdílu mezi výchozími hodnotami
m	1.00	0.94	57.9	54.4	
m̲	1	1			nemůže být v préture
n	1.03	1.16	44.6	50.2	
ň	1.06	1.19	53.5	60.2	
ŋ	1	1			nemůže být v préture
r	0.93	1.06	38.6	44.0	
r sl.	1	1			jen nukleus
ř	0.99	1.12	53.4	60.5	
l	1.00	1.20	37.7	45.4	
j	1.00	1.46	28.6	41.8	
?	1	1			nemůže být v kodě

**5.2 Koeficienty pozice slova v prozodické frázi.** Extrapolované a post-hoc dopočítané hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob.

Hláška	Koeficienty			Trvání (ms)			Poznámka
	Ini	Med	Fin	Ini	Med	Fin	
a	0.99	0.99	1.30	57.4	57.4	75.4	<i>iniciální stejné jako mediální</i>
e	1.00	1.00	1.55	49.9	49.9	77.4	
i	0.99	0.99	1.50	45.1	45.1	68.6	
o	1.00	1.00	1.38	46.9	46.9	64.8	
u	0.98	0.98	1.66	44.1	44.1	74.5	
a:	0.93	0.93	1.44	96.6	96.6	149.0	
e:	1.00	1.00	1.30	99.4	99.4	129.7	
i:	1.01	1.01	1.61	54.6	54.6	86.4	
o:	1.00	1.00	2.10	83.9	83.9	176.2	<i>podle o a rozdílu ve výchozích hodnotách</i>
u:	1.15	1.15	1.21	73.0	73.0	77.3	
ou	1.01	1.01	1.60	81.1	81.1	128.7	
p	0.81	0.94	0.98	63.4	73.1	76.9	
b	1.08	0.97	1.15	63.4	56.9	67.4	<i>podle p a rozdílu mezi znělými a neznělými</i>
t	0.97	0.87	1.34	66.0	59.5	91.3	
d	1.15	1.02	1.33	38.7	34.1	44.5	
ť	0.86	1.03	1.13	62.0	74.1	81.1	
d'	0.81	0.97	1.17	42.0	50.6	61.1	<i>podle ť a rozdílu mezi k a g</i>
k	1.03	1.00	1.44	66.4	64.4	92.5	
g	1.07	1.08	1.67	46.4	47.0	72.5	<i>podle k a rozdílu mezi ť a d'</i>
f	0.94	0.94	1.52	55.9	55.9	90.5	<i>frikativy - iniciální a mediální slova stejné</i>
v	0.99	0.99	1.16	41.0	41.0	48.4	
s	0.95	0.95	1.26	77.1	77.1	102.7	
z	0.94	0.94	1.18	51.3	51.3	64.8	
š	0.88	0.88	1.50	73.6	73.6	124.8	
ž	1.01	1.01	1.64	53.7	53.7	86.9	<i>podle š a rozdílu mezi s a z</i>
ch	0.94	0.94	1.54	61.7	61.7	101.5	<i>podle h a rozdílu mezi š a ž</i>
h	1.09	0.91	1.15	60.2	50.3	63.6	<i>post-hoc ze všech neseskupených dat</i>
c	0.97	0.97	1.57	85.1	85.1	138.4	<i>afrikáty - iniciální a mediální slova stejné</i>
dz	1.28	1.28	2.07	85.1	85.1	138.4	<i>podle c a rozdílu mezi znělými a neznělými</i>
č	1.01	1.01	1.12	102.3	102.3	114.1	
dž	1.61	1.61	1.79	102.3	102.3	114.1	<i>podle č a rozdílu ve výchozích hodnotách</i>
m	0.89	0.92	1.19	51.7	53.5	69.2	
m̩	0.89	0.92	1.19				<i>stejně jako m</i>
n	0.96	0.97	1.08	41.45	42.14	47.00	
ň	1.00	0.91	1.23	50.60	46.21	61.98	
ŋ	0.99	1.00	1.41	60.83	61.43	86.9	<i>podle g a rozdílu ve finálních hodnotách</i>
r	1.11	1.03	1.14	46.20	42.61	47.49	
r sl.	1	0.98	1.00		69.72	70.75	<i>malá citlivost na pozici, necháváme 1</i>
ř	1.19	1.05	1.26	64.06	56.65	68.01	
l	0.97	1.03	1.53	36.52	38.97	57.60	
j	1.49	0.98	1.72	42.62	27.96	49.06	
?	1	1	1				<i>není citlivý na pozici, iniciální je irelevantní</i>

**5.3 Koeficienty prodlužování finální slabiky finálních slov.** Extrapolované a post-hoc dopočítané hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob.

Hláška	Koeficient	Trvání (ms)	Poznámka
a	1.39	104.40	
e	1.19	92.32	
i	1.28	88.01	
o	1.64	106.22	
u	1.12	83.25	
a:	<b>2.16</b>	<b>322.31</b>	<i>post-hoc ze všech neseskupených dat</i>
e:	1.12	162.28	
i:	1.10	95.23	
o:	0.71	125.97	
u:	1.51	116.48	
ou	<b>1.21</b>	<b>156.32</b>	<i>post-hoc ze všech neseskupených dat</i>
p	<b>1.61</b>	<b>124.11</b>	<i>podle b a rozdílu mezi t a d</i>
b	1.04	70.44	
t	<b>1.10</b>	<b>100.61</b>	<i>post-hoc ze všech neseskupených dat</i>
d	<b>1.27</b>	<b>46.94</b>	<i>post-hoc ze všech neseskupených dat</i>
ť	<b>1.29</b>	<b>104.41</b>	<i>podle t a rozdílu ve výchozích hodnotách</i>
d'	<b>1</b>	<b>65.50</b>	<i>podle d a rozdílu ve výchozích hodnotách</i>
k	1.04	96.60	
g	<b>1</b>	<b>56.80</b>	<i>podle d' a rozdílu ve výchozích hodnotách (rozdíl ť-d' by vedl k příliš extrémní hodnotě)</i>
f	<b>1.34</b>	<b>121.50</b>	<i>post-hoc ze všech neseskupených dat</i>
v	1.09	52.83	
s	1.19	121.82	
z	<b>1.05</b>	<b>68.15</b>	<i>podle s a rozdílu mezi t a d</i>
š	1.34	167.40	
ž	<b>1.31</b>	<b>113.73</b>	<i>podle š a rozdílu mezi s a z</i>
ch	<b>1</b>	<b>98.00</b>	<i>podle k a rozdílu ve výchozích hodnotách</i>
h	1.08	68.80	<i>podle g a rozdílu ve výchozích hodnotách</i>
c	1.15	146.57	
dz	<b>1</b>	<b>94.35</b>	<i>podle c a rozdílu mezi znělými a neznělými</i>
č	1.18	145.11	
dž	1.09	92.89	<i>podle č a rozdílu mezi znělými a neznělými</i>
m	1.32	91.64	
ŋ	<b>1.32</b>	<b>91.64</b>	<i>stejně jako m</i>
n	1.09	51.21	
ň	1.18	73.23	
ŋ	1.09	94.60	<i>podle finálního slova</i>
r	1.14	53.95	
r sl.	<b>1</b>	<b>71.10</b>	<i>nedostatek dat a malá citlivost na změnu pozice</i>
ř	<b>1</b>	<b>66.25</b>	<i>podle r a rozdílu v referenčních hodnotách</i>
l	1	55.04	
j	1.20	58.92	

**5.4 Koeficienty hláskového okolí.** Extrapolované a post-hoc dopočítané hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob. Vokály, ráz a slabičné [r], pro něž je tento koeficient irelevantní, nejsou zobrazeny. Zkratka IV označuje intervokalickou pozici. Mimo shluk označuje pozici, kde v sousedství konsonantu je další konsonant, který ale nepatří do stejné slabiky.

Hláška	Koeficienty			Trvání (ms)			Poznámka
	Mimo shluk	IV	Ve shluku	Mimo shluk	IV	Ve shluku	
p	0.86	1.09	0.83	67.34	85.18	64.92	post-hoc dopočítáno ze všech neseskupených dat
b	0.84	1.00	0.78	49.48	58.59	45.52	podle p a rozdílu mezi výchozími hodnotami (protože d se chová jinak)
t	1.10	0.98	0.92	75.18	66.43	62.34	post-hoc dopočítáno ze všech neseskupených dat
d	1.30	0.91	1.09	43.58	30.44	36.70	
ť	1.02	1.07	0.92	73.67	76.63	65.82	podle d a rozdílu mezi k a g
d'	0.85	1.07	1.06	44.33	55.63	55.26	podle d a rozdílu mezi výchozími hodnotami
k	0.96	1.13	0.89	61.61	72.92	57.49	podle g a rozdílu mezi ch a h
g	0.74	1.20	1.08	32.27	51.92	46.94	podle k a rozdílu mezi výchozími hodnotami
f	0.97	1.31	1.00	57.69	77.87	59.53	podle v a rozdílu mezi s a z
v	0.93	1.06	1.02	38.78	44.17	42.49	
s	1.02	1.14	0.83	83.24	92.47	67.09	post-hoc dopočítáno ze všech neseskupených dat
z	0.99	1.08	0.92	53.91	58.77	50.05	post-hoc dopočítáno ze všech neseskupených dat
š	0.89	1.09	0.79	74.12	91.03	65.56	podle ž a rozdílu mezi s a z
ž	0.92	1.08	0.91	48.99	57.33	48.52	podle z a rozdílu mezi výchozími hodnotami
ch	0.86	1.12	0.89	56.47	73.55	58.69	podle š a rozdílu mezi výchozími hodnotami
h	0.83	1.14	0.87	46.07	63.15	48.13	podle ch a rozdílu mezi výchozími hodnotami
c	0.83	1.05	1.01	72.80	91.88	88.60	
dz	1.00	1.00	1.07	66.70		71.56	podle c a rozdílu mezi s a z, nemůže být intervokalické
č	1.00	1.02	1.01	101.34	103.05	102.20	podle c a rozdílu mezi výchozími hodnotami
dž	1.20	1.09	1.34	76.21	69.35	85.16	podle č a rozdílu mezi š a ž
m	0.90	1.00	0.83	52.19	57.75	48.41	
ɱ	0.90	1.00	0.83				stejně jako m
n	1.00	0.98	0.84	43.30	42.42	36.58	
ň	0.91	1.00	0.75	46.17	50.73	37.86	
ɲ	1.00	1	0.89	61.34		54.61	podle n a rozdílu ve výchozích hodnotách, nemůže být intervokalické
r	1.00	0.93	1.06	41.66	38.53	43.85	
ř	0.90	0.96	1.10	48.65	51.73	58.92	
l	1.25	0.99	0.95	47.20	37.38	35.84	
j	0.97	1.01	0.94	27.64	28.87	27.02	

**5.5 Koeficienty trvání vokálů po rázu.** Byly vypočítány pouze pro krátké vokály, dlouhé se v korpusu nevyskytují. Extrapolovaná hodnota je zobrazena červeně, v poznámce je uveden způsob extrapolace.

Hláška	Koeficient	Trvání (ms)	Poznámka
a	0.69	45.50	
e	0.69	34.43	podle a
i	0.8	37.90	
o	0.74	42.63	
u	0.58	40.80	

**5.6 Koeficienty délky slova ve slabikách.** Extrapolované a post-hoc dopočítané hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob. Čísla v názvech sloupců označují počet slabik ve slově.

Hláška	Koeficienty				Trvání (ms)				Poznámka
	1	2	3	4	1	2	3	4	
a	0.99	1.01	0.94	1.06	57.15	58.31	54.25	61.72	
e	0.97	1.02	1.05	0.92	48.61	50.76	52.23	45.74	
i	1.04	0.95	0.93	0.88	47.50	43.40	42.57	40.26	
o	0.99	1.02	1.06	0.96	46.59	48.10	49.83	45.11	
u	1.12	0.96	1.01	1.03	50.26	43.06	45.27	46.46	
a:	0.85	0.91	1.04	1.12	87.70	93.87	107.86	116.44	
e:	1.01	1.01	0.83	0.98	100.36	100.66	82.74	97.49	podle e a rozdílu výchozích hodnot
i:	0.94	1.06	1.01	1.01	50.50	57.28	54.12	54.35	
o:	1.36	1.01	1.03	0.98	114.45	85.09	86.81	82.10	podle o a rozdílu výchozích hodnot
u:	1.08	0.99	1.00	1.02	68.92	63.04	63.92	65.11	podle u a rozdílu výchozích hodnot
ou	0.81	1.04	0.91	0.87	64.72	83.60	72.89	69.68	
p	0.89	0.95	1.03	0.89	69.30	73.84	80.70	69.34	post-hoc přepočítáno z neseskupených všech dat
b	0.85	1.06	1.07	1.05	49.83	62.24	62.52	61.70	
t	0.97	1.11	1.05	1.17	65.85	75.64	71.28	79.74	post-hoc přepočítáno z neseskupených všech dat
d	1.13	0.97	1.02	1	37.85	32.69	34.16		4slabičné nemodelujeme, kde není dostatek dat
ť	1.04	0.91	1.16	1	74.69	65.21	83.11		
d'	0.93	0.96	1.06	1	48.41	49.82	55.18		
k	0.92	0.99	1.04	1.34	59.38	63.72	66.99	86.53	
g	1.09	1.01	0.79	1	47.31	44.02	34.16		post-hoc přepočítáno z neseskupených všech dat
f	1.07	1.10	0.84	1.22	63.86	65.60	49.78	72.75	podle v a rozdílu mezi s a z
v	0.94	0.96	1.13	0.90	39.11	39.91	46.96	37.58	
s	0.93	0.98	0.99	1.10	75.65	79.54	80.39	89.66	
z	0.93	0.95	0.97	0.82	50.90	51.76	52.89	44.75	
š	0.87	0.90	0.94	1	72.43	74.60	78.42		

ž	0.97	0.98	1.46	1	51.78	52.00	77.82		
ch	0.77	1.05	0.70	1	50.36	69.08	46.27		podle h a rozdílu mezi š a ž
h	0.93	0.93	0.82	1	51.58	51.25	45.67		
c	0.96	1.06	1.24	0.98	83.95	93.06	109.21	85.85	podle č a rozdílu výchozích hodnot
dz	0.96	1.01	1.29	1	64.09	67.66	86.36		podle c a rozdílu mezi znělými a neznělými
č	1.05	1.04	0.94	1	106.58	106.05	95.61		
dž	0.84	1.39	1.21	1	53.54	88.32	77.21		
m	0.88	0.98	1.01	0.90	51.01	56.98	58.73	52.48	
ŋ	0.88	0.98	1.01	0.90					stejně jako m
n	0.97	0.94	0.92	1.24	41.90	40.77	40.13	53.85	
ň	0.84	0.95	0.92	1.18	42.28	47.89	46.63	59.61	
ŋ	0.98	0.96	0.95	1.17	59.93	58.80	58.16	71.88	podle n a rozdílu výchozích hodnot
r	1.03	1.04	0.99	1.04	42.69	43.00	41.16	43.19	
r sl.	1.02	1.02	1.00	1.02	72.29	72.60	70.76	72.79	podle r a rozdílu výchozích hodnot
ř	1.13	0.99	1.06	1	60.73	53.12	57.24		
l	1.37	0.98	1.11	1.11	51.50	36.90	41.67	41.82	
j	0.96	1.04	0.99	1	27.60	29.64	28.39		
?	1.02	1.14	0.79	0.73	65.11	72.82	50.28	46.56	

**5.7 Koefficienty délky prozodické fráze.** Extrapolované hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob extrapolace. Čísla v názvech sloupců označují počet slov v prozodické frázi.

Hláška	Koefficienty								Trvání (ms)								Poznámka
	3	4	5	6	7	8	9	3	4	5	6	7	8	9			
a	1.06	1.01	0.97	0.96	0.99	0.92	1.18	61.4	58.8	56.3	55.9	57.6	53.5	68.4			
e	1.08	1.00	1.04	0.98	0.91	0.96	0.96	53.7	49.9	51.8	49.1	45.4	48.0	47.8			
i	1.06	1.08	0.96	1.00	0.97	1.03	0.86	48.3	49.3	43.6	45.7	44.4	47.0	39.3			
o	1.11	1.01	0.97	1.01	1.05	0.95	1.08	52.0	47.7	45.5	47.5	49.5	44.7	50.7			
u	1.04	0.95	1.01	1.03	0.97	0.97	0.94	47.0	42.6	45.3	46.2	43.6	43.7	42.3	podle o a rozdílu mezi i a e		
a:	0.74	0.93	0.95	1.06	0.91	0.80	0.83	77.0	96.6	98.1	110.2	94.2	82.6	86.5			
e:	1.19	0.94	0.93	0.91	0.95	0.98	0.98	118.8	93.5	92.5	91.3	95.3	97.9	97.7	podle e a rozdílu výchozích hodnot		
i:	1.00	1.07	1.01	1.01	1.01	0.90	1.13	53.8	57.8	54.6	54.1	54.2	48.5	60.7			
o:	1.06	1.01	0.98	1.01	1.03	0.97	1.04	89.0	84.7	82.5	84.5	86.5	81.6	87.7	podle o a rozdílu výchozích hodnot		
u:	1.03	0.96	1.00	1.39	0.71	0.55	0.73	65.6	61.3	64.0	88.7	45.5	35.1	46.6	podle u a rozdílu výchozích hodnot		
ou	1.10	0.93	1.02	1.04	0.88	0.96	1.01	88.6	74.7	81.7	83.4	70.6	77.2	81.2	poměrná část součtu trvání o a u (podle výchozích hodnot)		
p	0.94	0.98	0.99	0.86	0.87	0.82	0.83	73.1	76.4	77.2	67.0	68.0	64.3	64.5			
b	0.95	1.00	0.91	1.02	1.03	0.94	0.83	55.6	58.4	53.7	59.6	60.5	55.2	48.9			
t	0.88	0.91	0.81	0.90	0.87	0.83	0.88	60.0	61.8	54.9	61.6	58.9	56.8	60.0			

d	1.00	1.01	0.95	1.13	1.18	1.17	0.99	33.5	33.9	32.0	37.9	39.7	39.3	33.3	
ť	1.21	0.88	1.13	0.96	1.05	0.84	0.86	86.7	63.0	81.0	69.3	75.4	60.7	61.9	
d'	1.26	0.90	0.98	0.97	0.95	1.47	0.83	65.9	46.9	51.1	50.5	49.6	76.7	43.5	
k	1.04	0.98	1.01	1.06	0.94	1.02	1.13	66.7	62.8	64.7	68.0	60.8	65.4	73.1	
g	1.16	1.05	1.12	1.14	0.79	0.72	1.26	50.5	45.5	48.5	49.7	34.1	31.4	54.7	podle k a rozdílu mezi <i>ť</i> a <i>d'</i>
f	0.74	0.91	0.92	0.87	1.19	1.09	1.18	44.0	54.3	54.6	51.9	70.7	65.0	70.1	
v	1.15	1.06	1.01	0.91	1.10	0.92	1.03	48.0	44.1	42.0	37.9	45.6	38.1	42.7	
s	0.81	0.96	0.97	1.01	0.87	0.88	0.98	65.4	78.0	79.1	81.8	71.1	71.7	79.8	
z	1.00	0.98	0.89	0.88	1.12	0.98	0.74	54.6	53.7	48.6	48.2	61.2	53.8	40.7	
š	1.14	0.91	1.11	0.76	0.86	0.79	0.73	95.3	76.1	92.3	63.5	71.3	65.6	61.2	
ž	1.24	1.12	1.03	1.02	0.98	0.82	1.09	65.8	59.4	54.6	54.3	51.9	43.6	57.8	
ch	1.58	0.99	0.77	1.08	0.76	0.78	0.89	104.1	65.2	50.6	71.3	50.1	51.3	58.4	
h	1.05	0.93	0.92	0.92	0.79	0.80	0.99	58.1	51.6	50.9	51.1	43.7	44.1	55.0	podle ch a rozdílu mezi <i>š</i> a <i>ž</i>
c	1.03	1.02	0.92	0.99	1.08	0.96	1.11	90.6	90.0	80.8	87.0	94.6	84.0	97.6	
dz	1.03	1.03	1.00	0.97	0.77	1.00	1.12	68.4	68.9	66.8	64.9	51.6	66.8	75.0	podle c a rozdílu mezi <i>znělými</i> a <i>neznělými</i>
č	0.76	0.92	1.01	1.03	0.96	1.10	1.10	77.3	93.5	102.9	104.3	97.7	112.1	111.2	podle c a rozdílu <i>výchozích</i> hodnot
dž	1.03	1.03	1.00	0.97	0.76	1.00	1.13	65.4	65.9	63.8	61.9	48.6	63.8	72.0	podle dž a <i>rozdílu</i> <i>výchozích</i> hodnot
m	0.85	0.96	0.92	0.91	0.92	0.93	0.98	49.2	55.5	53.6	52.5	53.4	54.1	56.7	
mj	0.85	0.96	0.92	0.91	0.92	0.93	0.98								stejně jako m
n	1.07	0.97	0.98	0.92	1.00	1.05	1.07	46.4	41.9	42.5	39.9	43.2	45.4	46.4	
ň	0.93	0.86	0.93	0.84	1.17	0.85	0.79	47.2	43.7	47.1	42.2	58.8	42.7	39.7	
ŋ	1.05	0.98	0.98	0.94	1.00	1.03	1.05	64.4	59.9	60.5	57.9	61.2	63.5	64.5	podle n a rozdílu <i>výchozích</i> hodnot
r	0.88	1.02	1.05	1.01	1.05	1.06	1.06	36.4	42.3	43.7	41.8	43.4	44.2	44.1	
r sl.	0.93	1.01	1.03	1.00	1.03	1.04	1.04	66.0	71.9	73.3	71.4	73.0	73.8	73.7	podle r a rozdílu <i>výchozích</i> hodnot
ř	1.03	0.93	1.00	1.15	1.16	0.97	1.00	55.7	50.2	54.1	61.7	62.3	52.4	54.0	
l	0.87	1.07	1.06	1.00	1.04	0.93	0.92	32.6	40.3	39.8	37.9	39.1	35.1	34.7	
j	1.13	0.94	0.92	0.89	1.13	1.27	0.87	32.2	26.8	26.3	25.5	32.4	36.4	24.9	
?	1.07	0.75	0.75	0.79	0.75	0.73	0.51	68.1	48.1	47.7	50.6	47.9	46.5	32.7	

**5.8 Koeficienty slabičné struktury.** Vypočítané pouze pro vokály. Extrapolované a post-hoc dopočítané hodnoty jsou zobrazeny červeně, v poznámce je uveden způsob. Názvy sloupců označují strukturu dané slabiky.

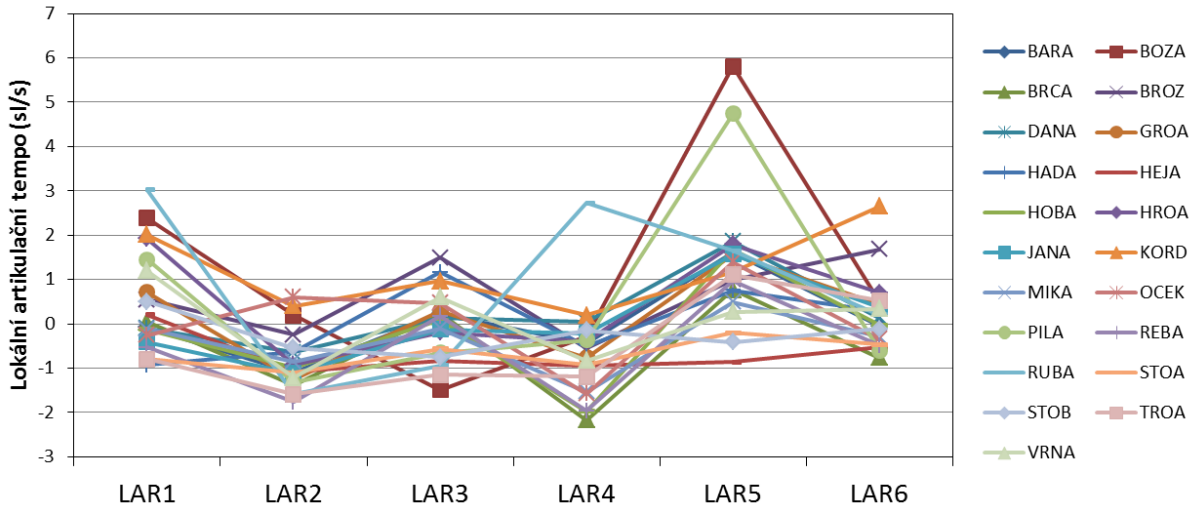
Hláška	Koeficienty							Trvání (ms)							Poznámka
	V	VC	CV	CVC	CCV	CCVC	CCCV	V	VC	CV	CVC	CCV	CCVC	CCCV	
a	1.10	1.01	1.00	0.96	0.90	1.01	0.93	63.97	58.80	57.81	55.77	52.06	58.32	53.67	
e	1.10	0.91	0.99	1.06	0.96	0.97	0.97	54.75	45.61	49.64	53.09	47.77	48.48	48.22	
i	1.11	0.95	0.99	0.96	1.07	1.00	0.96	50.49	43.14	45.09	43.87	48.88	45.56	43.92	podle e a rozdílu mezi referenčními hodnotami
o	1.15	1.16	1.00	1.04	0.95	0.91	1.08	54.13	54.47	47.12	48.74	44.85	42.63	50.85	podle u a rozdílu mezi referenčními hodnotami
u	1.04	1.17	0.93	1.16	1.02	0.88	1.15	46.96	52.47	41.97	52.05	45.73	39.70	51.72	podle o a rozdílu mezi i a e
a:	1.06	0.75	1.03	0.84	0.95	0.91	0.85	109.61	78.23	106.50	87.07	97.94	94.47	88.45	podle a a rozdílu v referenčních hodnotách
e:	1.05	0.96	1.00	1.03	0.98	0.99	0.98	104.65	95.51	99.54	102.99	97.67	98.38	98.12	podle e a rozdílu mezi referenčními hodnotami
i:	2.05	0.95	1.05	0.96	1.05	0.88	0.97	110.55	51.34	56.72	51.84	56.30	47.54	52.12	podle i a rozdílu v referenčních hodnotách
o:	1.08	1.09	1.00	1.02	0.97	0.95	1.05	91.11	91.45	84.10	85.72	81.83	79.61	87.83	podle o a rozdílu v referenčních hodnotách
u:	1.03	1.12	0.74	1	0.59	0.92	1.11	65.61	71.12	46.88	63.70	37.35	58.35	70.38	podle u a rozdílu v referenčních hodnotách; CVC stanovena post-hoc, viz text
ou	1.03	1.16	1.03	1.10	0.93	0.89	1.11	82.40	93.34	82.56	87.98	74.89	71.86	89.53	poměrná část součtu trvání o a u (podle referenčních hodnot)



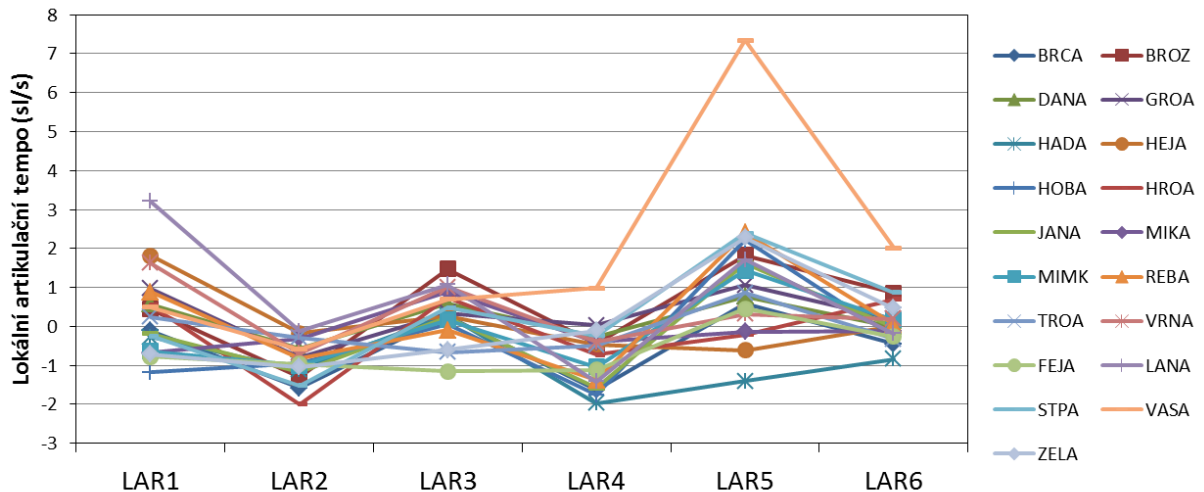
## Příloha 6: Grafy reziduí lokálního artikulačního tempa

Rezidua v grafech jsou vypočítána jako rozdíl modelových a reálných hodnoty LAR. Nula tedy zastupuje modelové hodnoty LAR, kladná čísla znamenají zrychlení a záporná zpomalení lokálního artikulačního tempa oproti modelu. Každá barevná čára znázorňuje jednoho mluvčího.

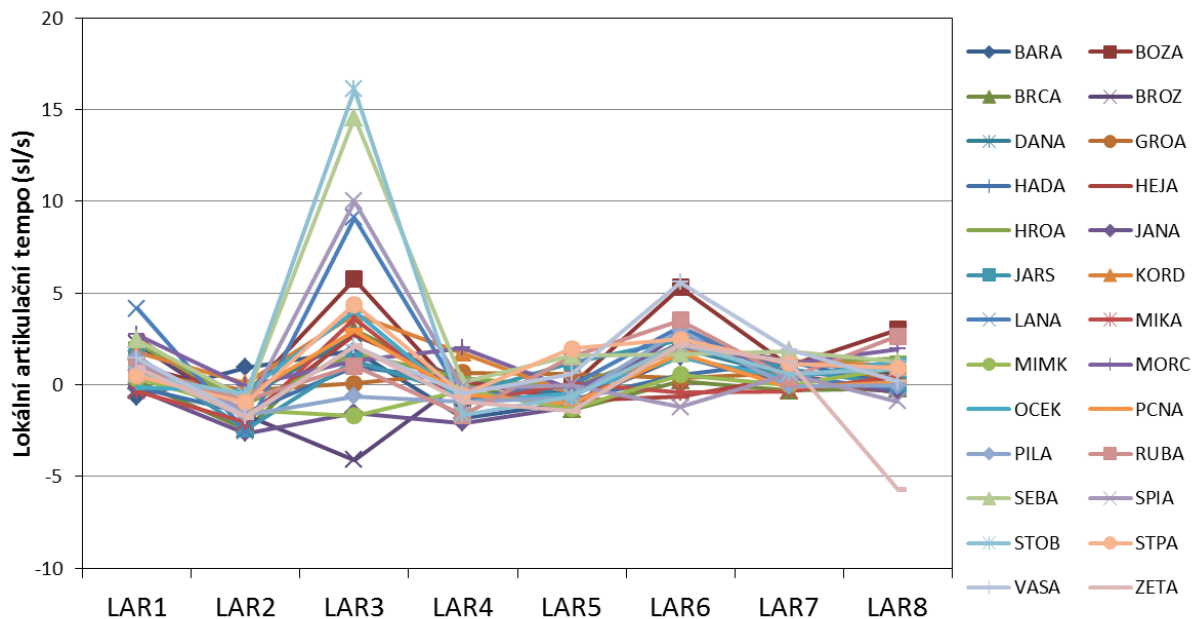
**Graf H1a\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Řekneš jim, co si myslíš.***



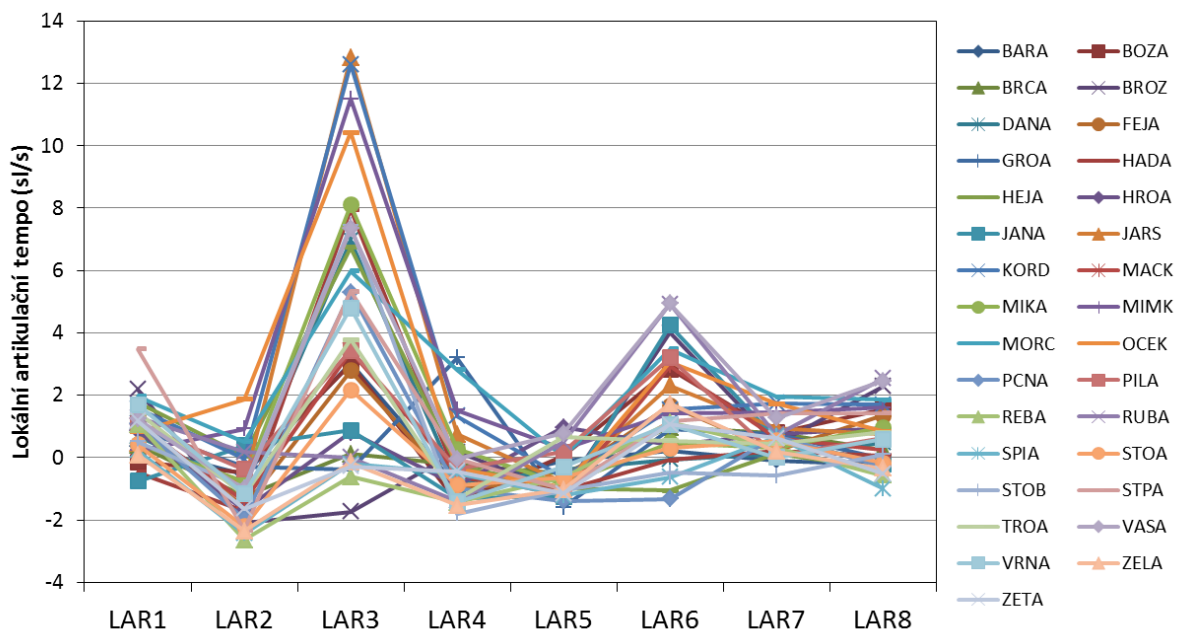
**Graf H1b\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Řekneš jim, co si myslíš?***



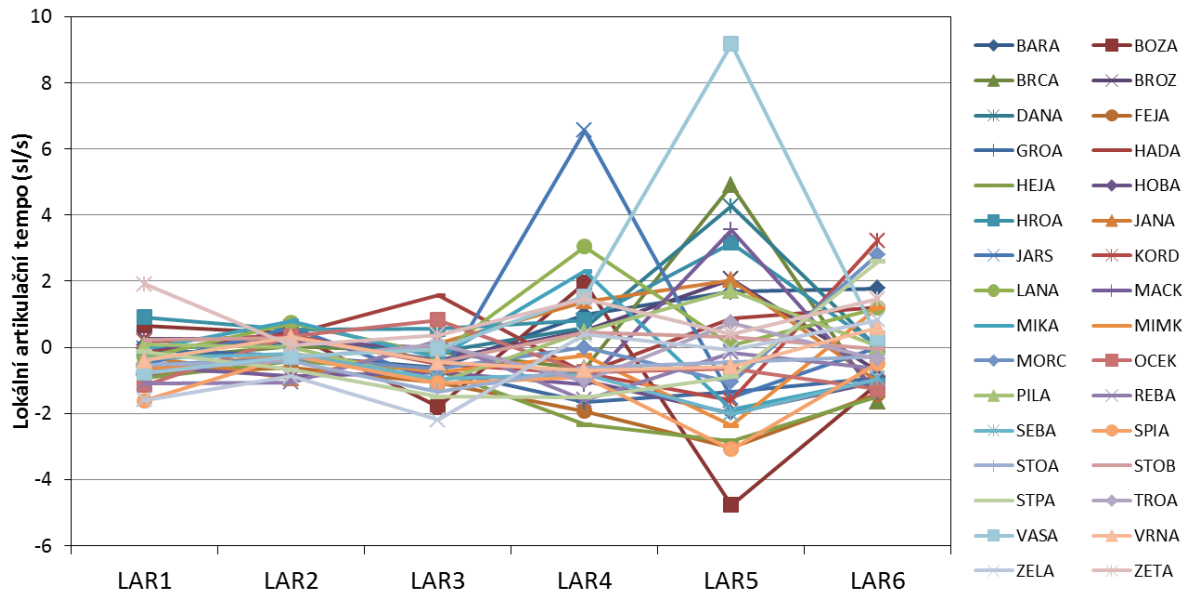
**Graf H1c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Řeknete jim, co si myslíte.***



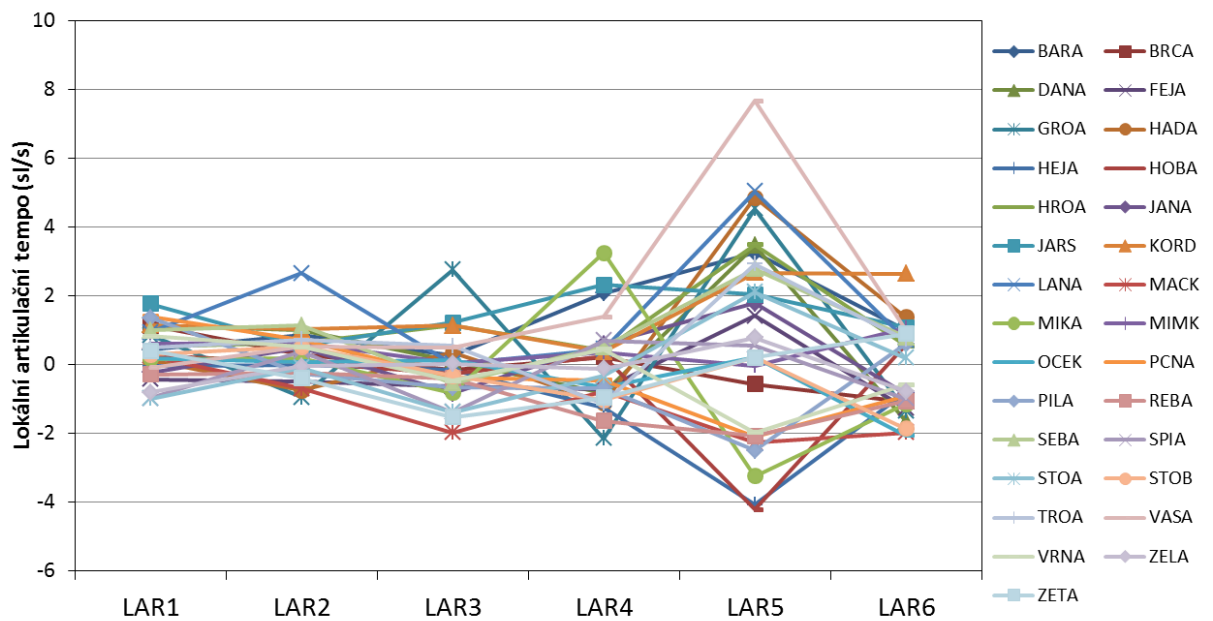
**Graf H1d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Řeknete jim, co si myslíte?***



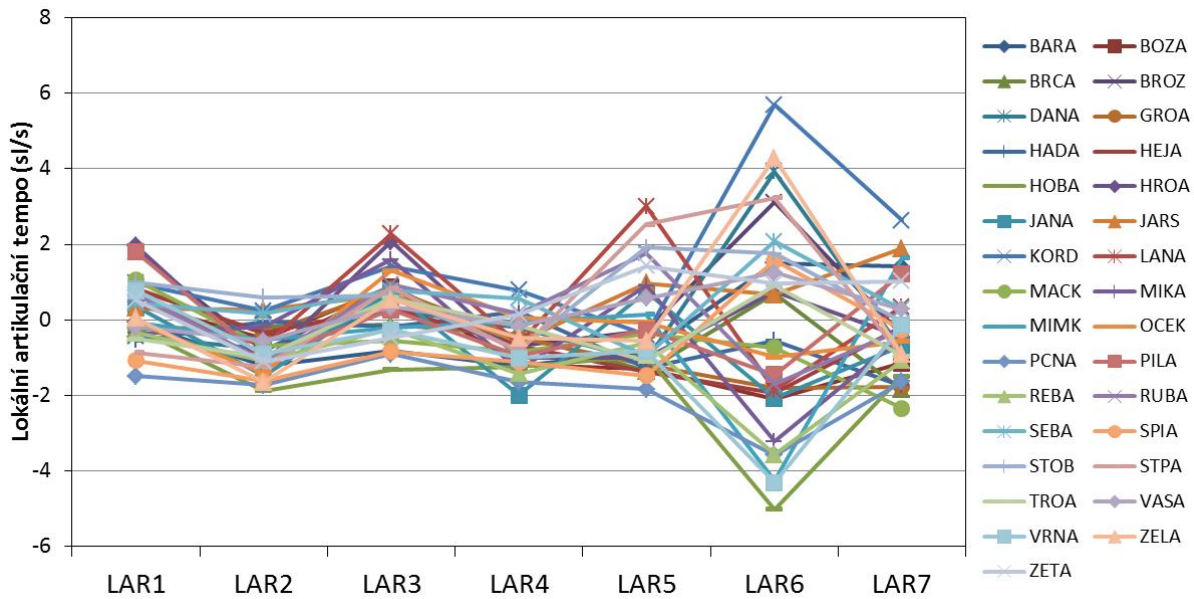
**Graf H2a\_3: Kontury reziduí lokálního artikulačního tempo pro větu *Zeptáš se, co bude dál.***



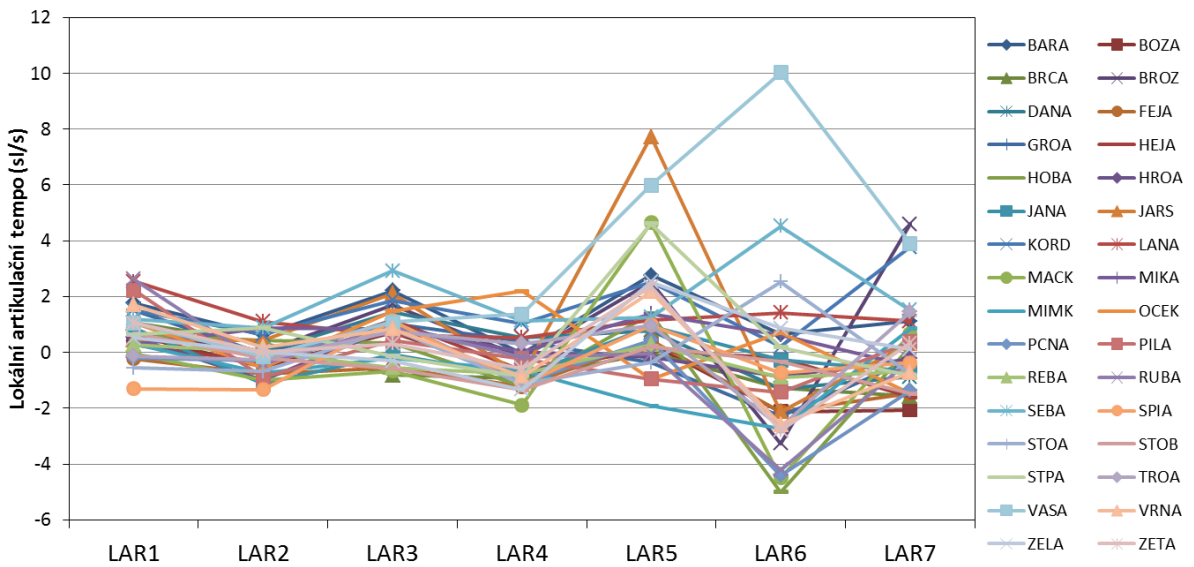
**Graf H2b\_3: Kontury reziduí lokálního artikulačního tempo pro větu *Zeptáš se, co bude dál?***



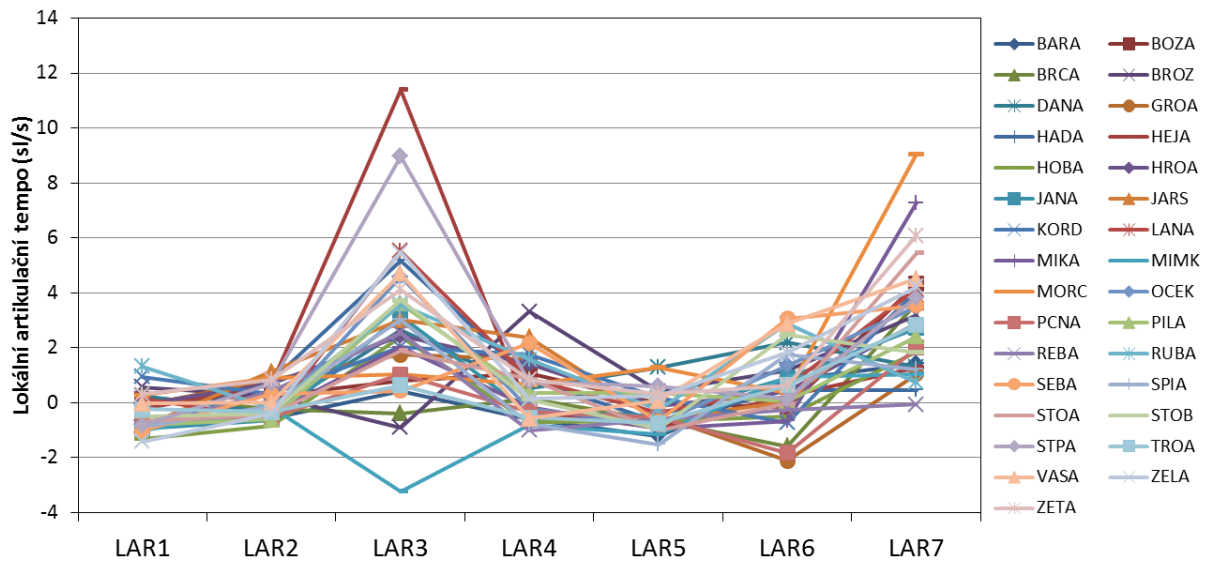
**Graf H2c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zeptáte se, co bude dál.***



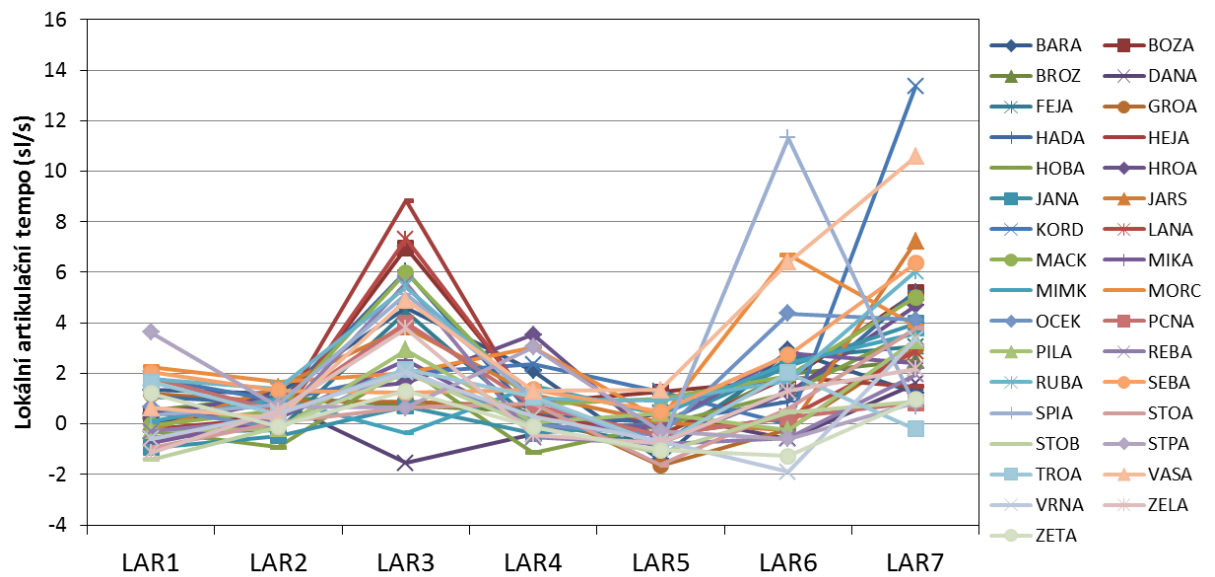
**Graf H2d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zeptáte se, co bude dál?***



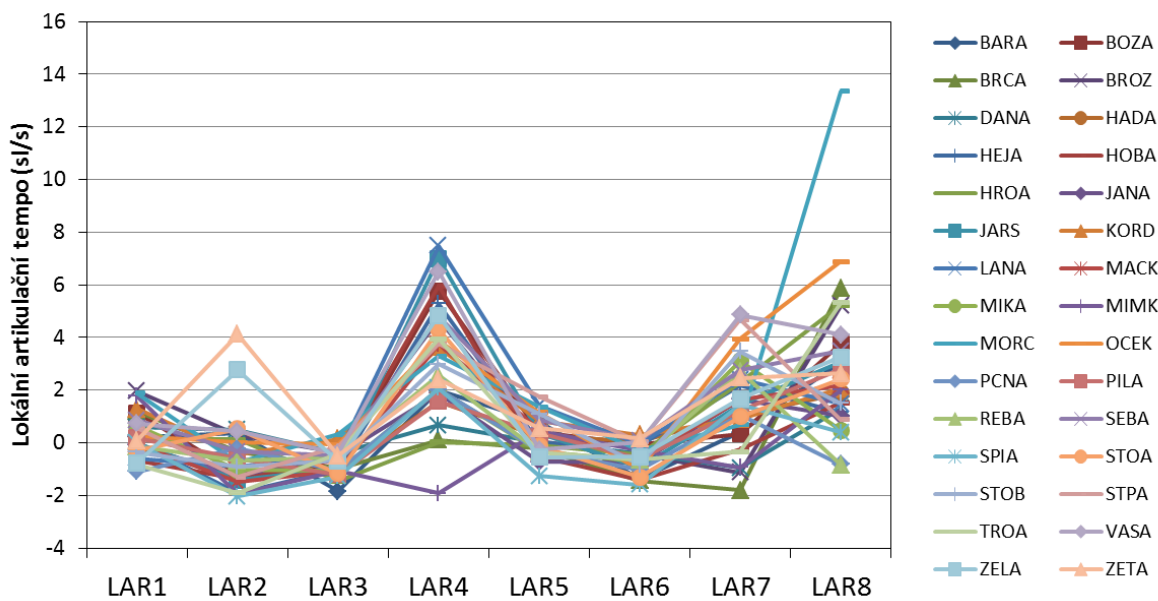
**Graf H3a\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zjistíš, v kolik to pojede.***



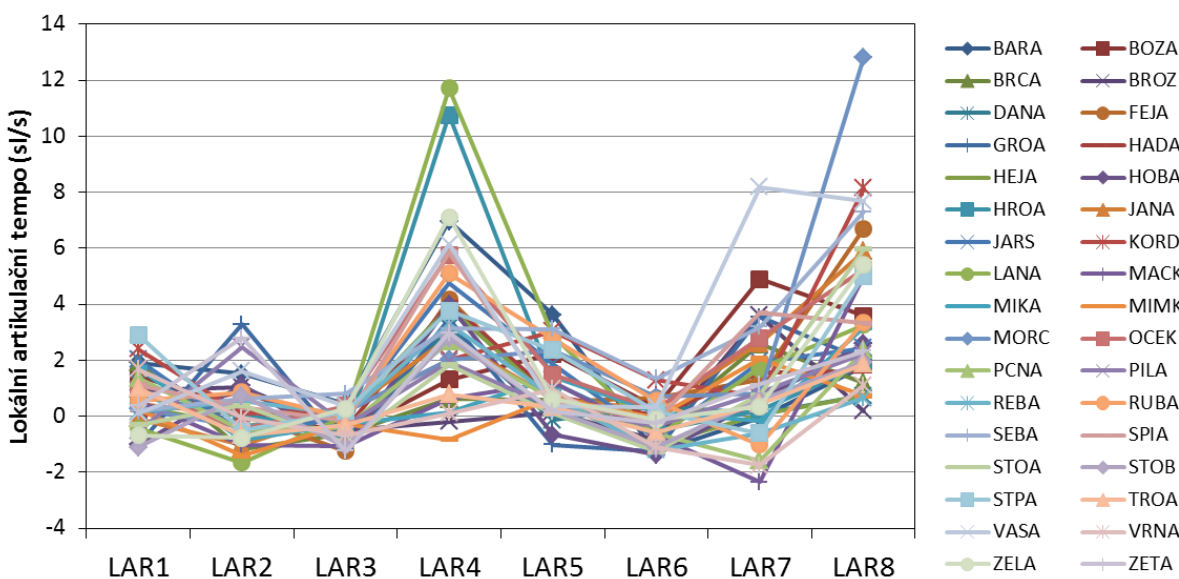
**Graf H3b\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zjistíš, v kolik to pojede?***



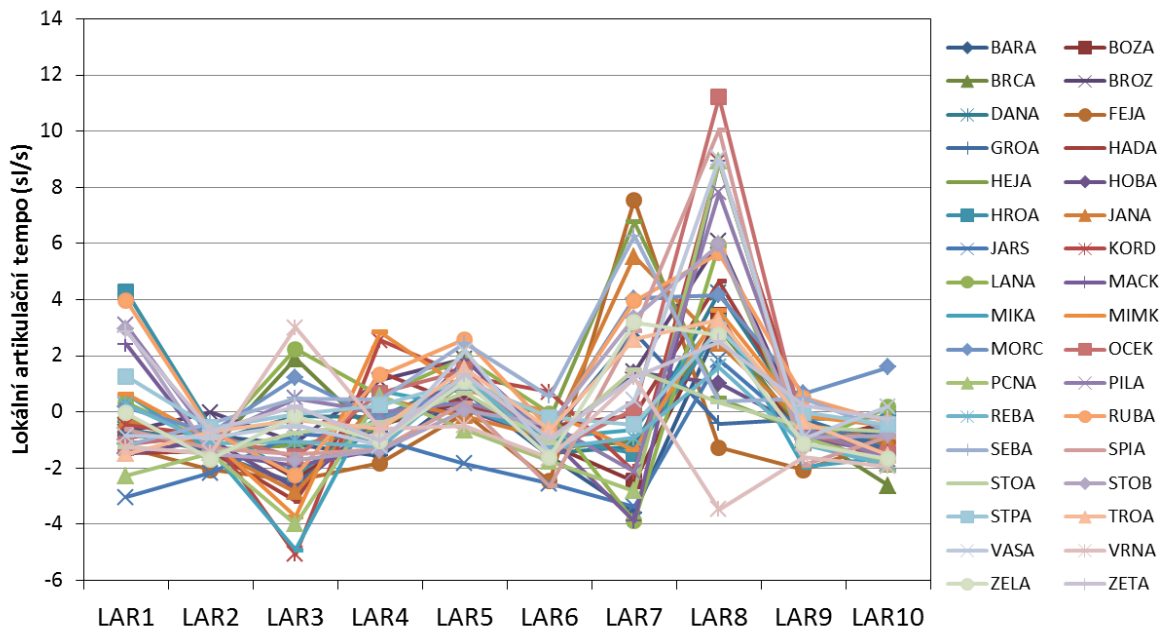
**Graf H3c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zjistíte, v kolik to pojede.***



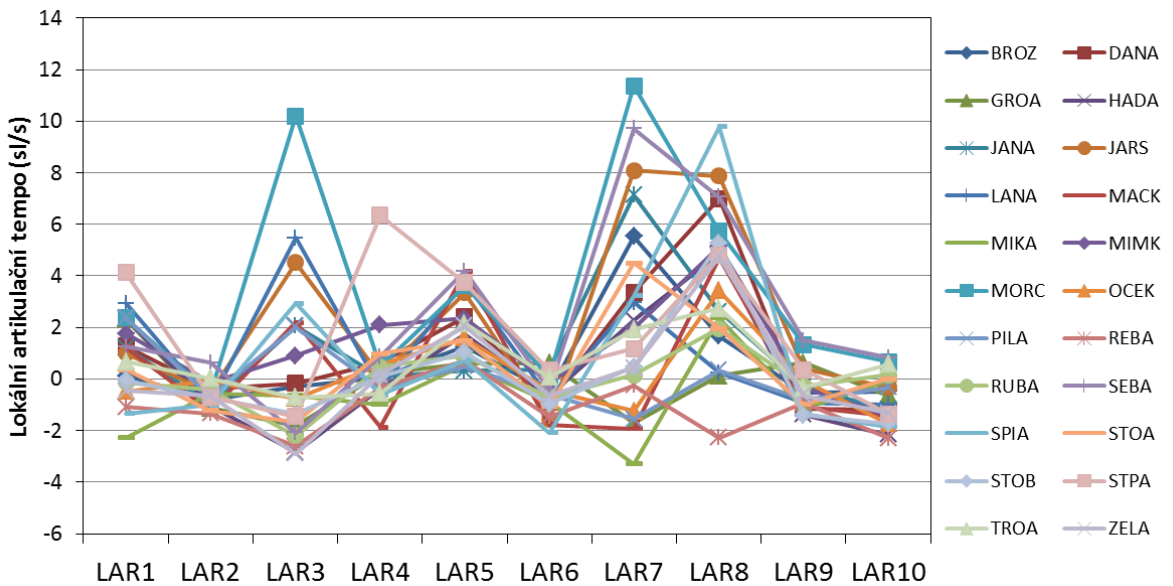
**Graf H3d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Zjistíte, v kolik to pojede?***



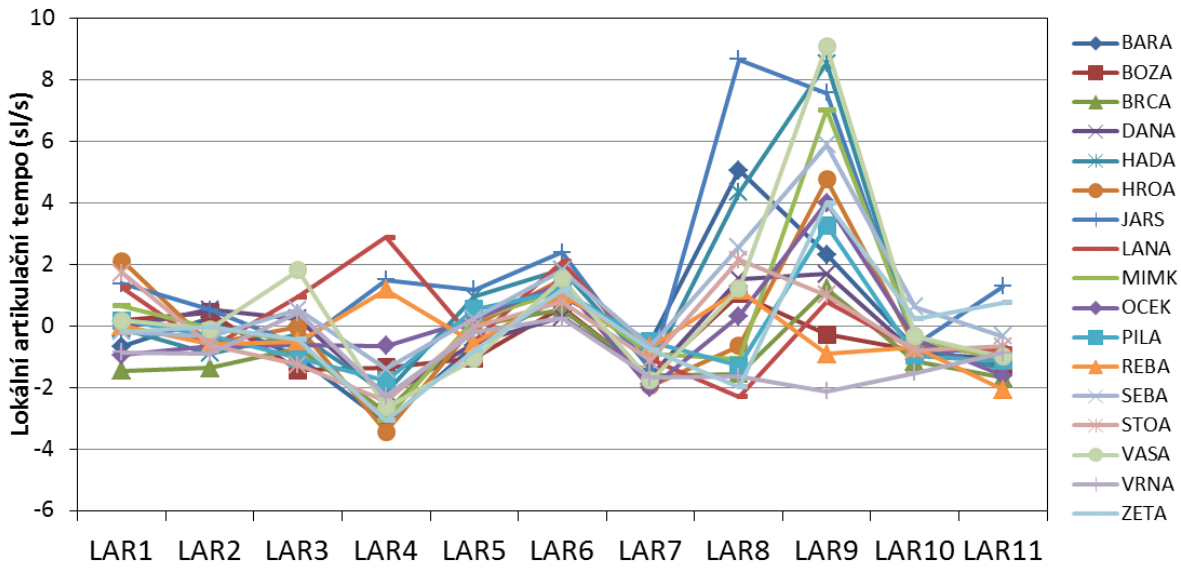
**Graf H4a\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Nemáš poněti, kdo pojede vzadu.***



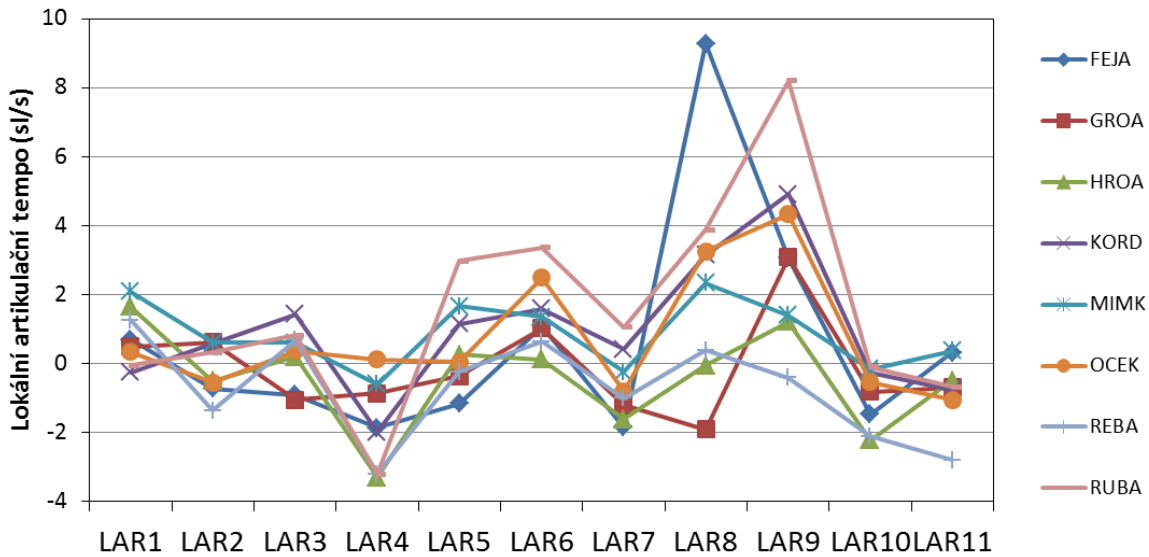
**Graf H4b\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Nemáš poněti, kdo pojede vzadu?***



**Graf H4c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Nemáte poněť, kdo pojede vzadu.***

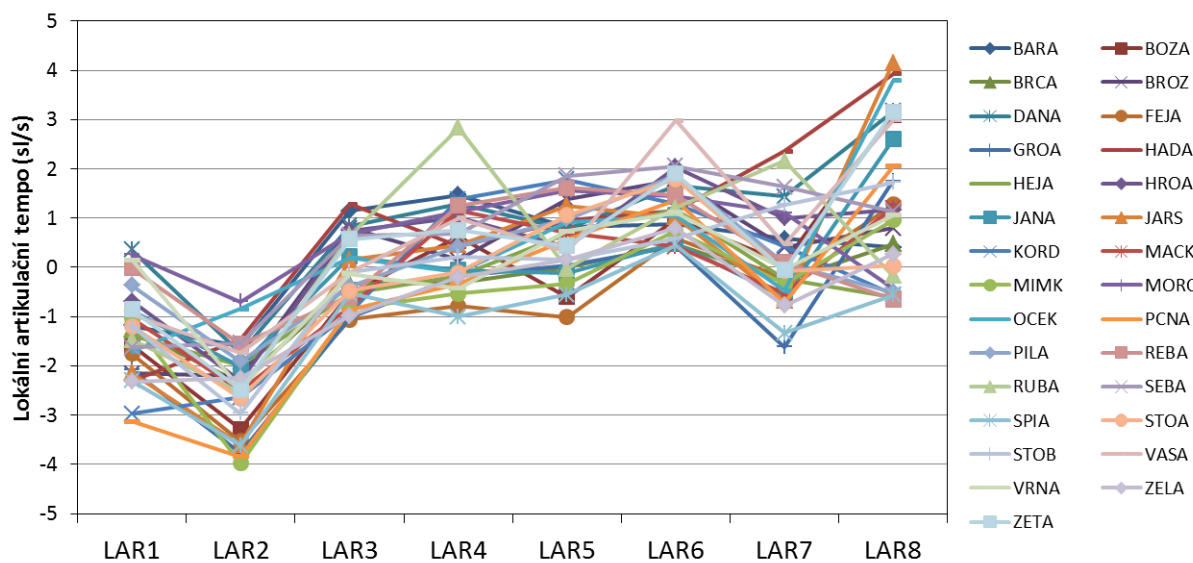


**Graf H4d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Nemáte poněť, kdo pojede vzadu?***

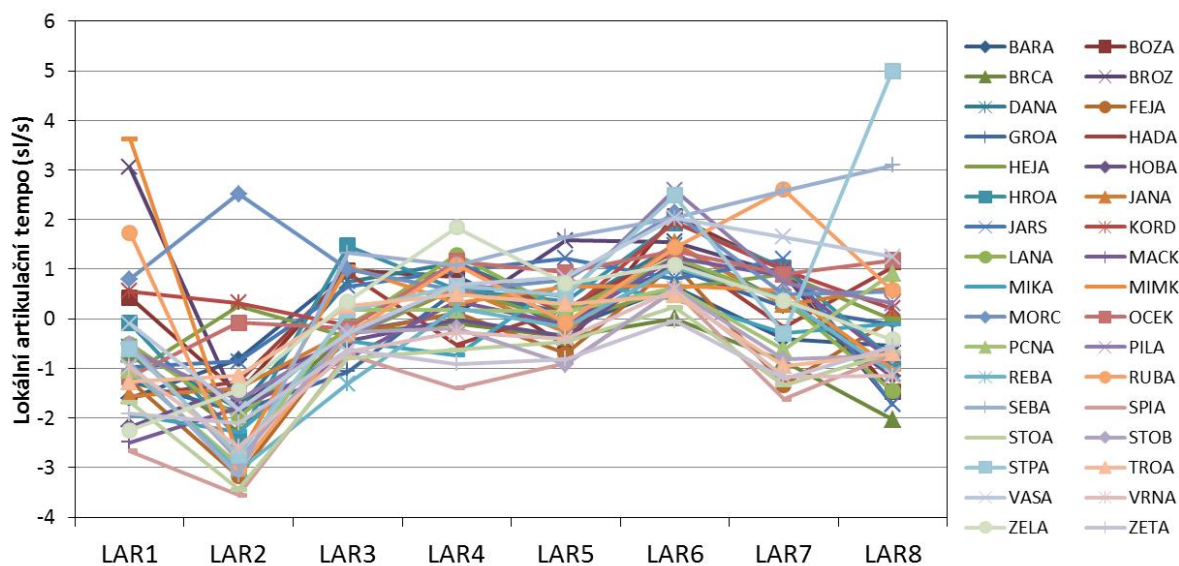




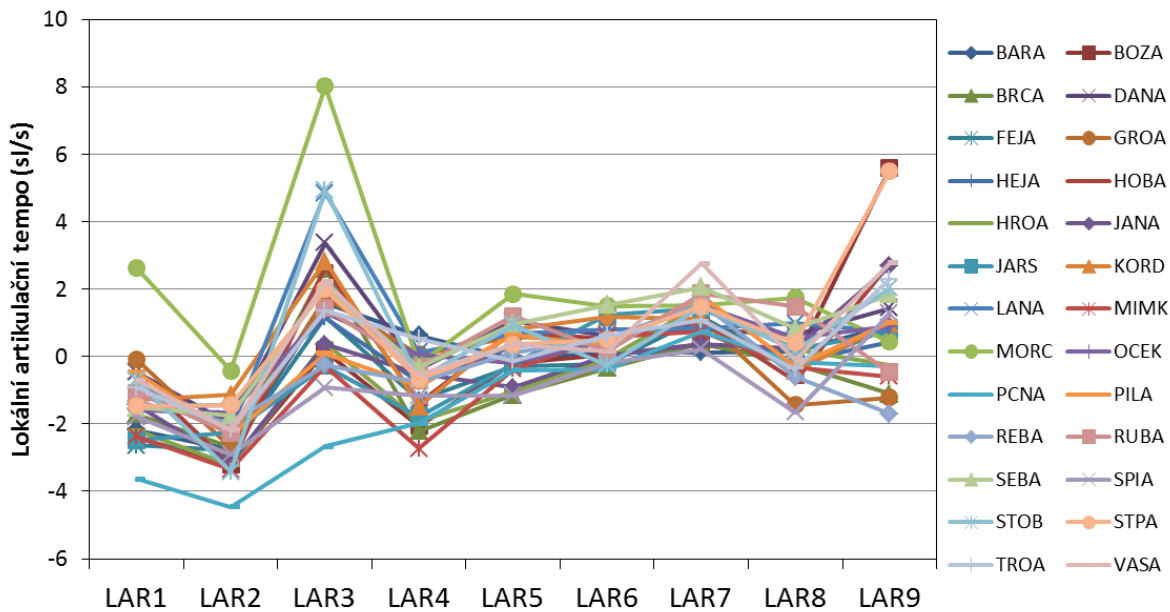
**Graf H5a\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Netušíš, kam to přesouvají.***



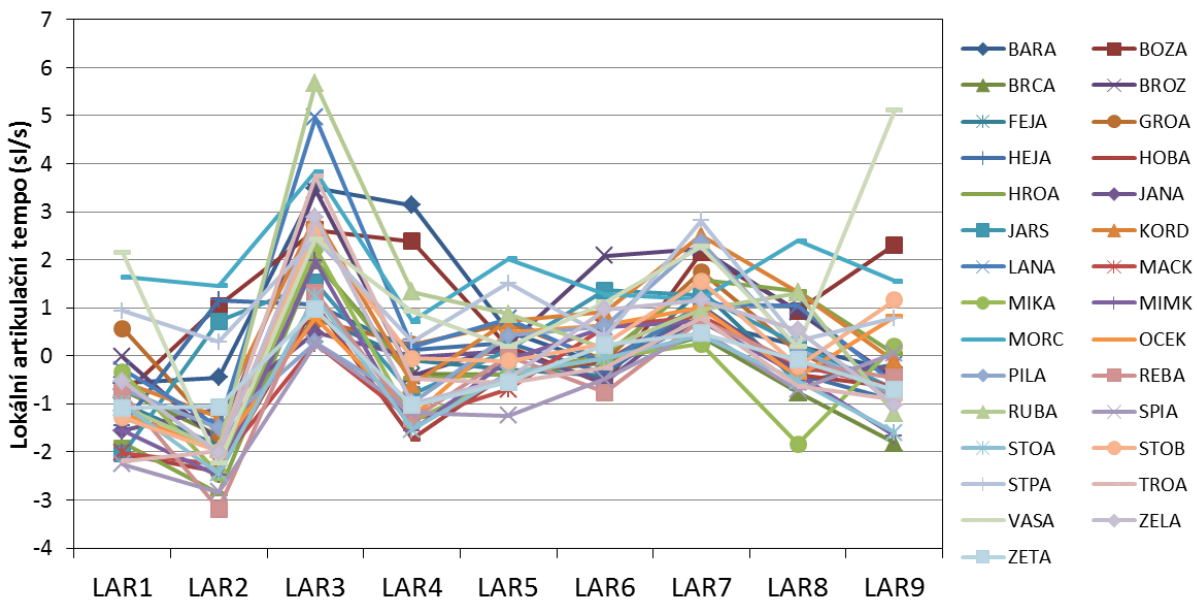
**Graf H5b\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Netušíš, kam to přesouvají?***



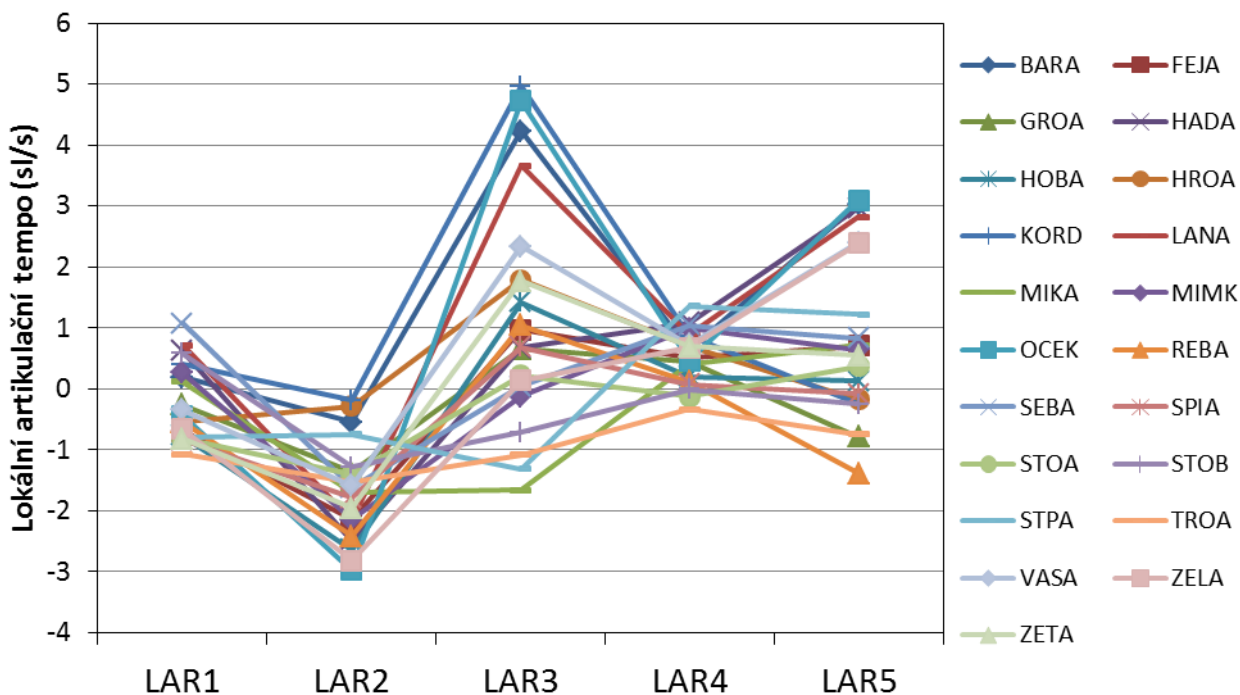
**Graf H5c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Netušíte, kam to přesouvají.***



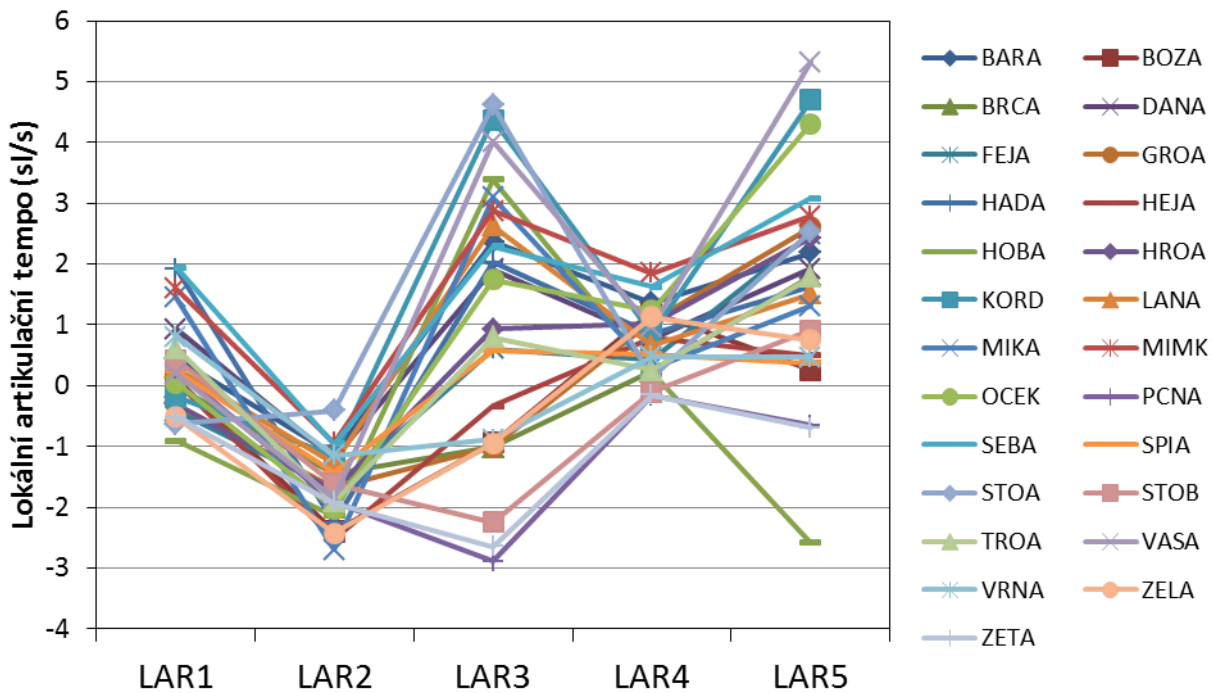
**Graf H5d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Netušíte, kam to přesouvají?***



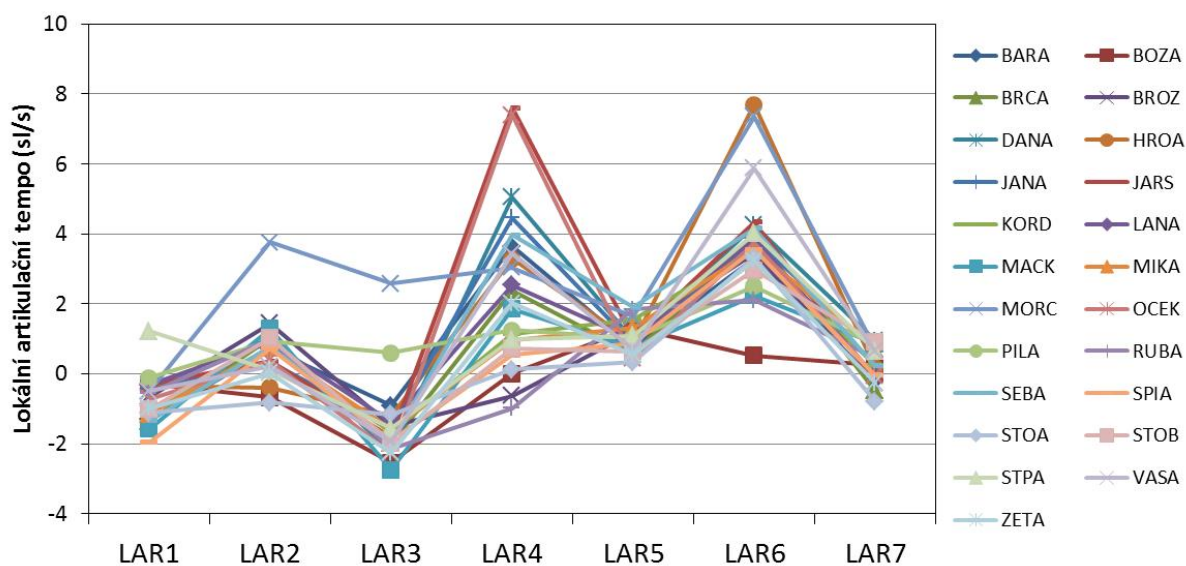
Graf H6a\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Víš, kdy to dostaneš.*



Graf H6b\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Víš, kdy to dostaneš?*



**Graf H6c\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Víte, kdy to dostanete.***



**Graf H6d\_3: Kontury reziduí lokálního artikulačního tempa pro větu *Víte, kdy to dostanete?***

