



# Posudek diplomové práce

předložené ve Fonetickém ústavu Filozofické fakulty Univerzity Karlovy

**Jméno a tituly posuzujícího vedoucího/opponenta:** Ing. Tomáš Božil, Ph.D.

posudek vedoucího/vedoucí  posudek oponenta/oponentky

**Autor/autorka:** Bc. Michaela Rabanová

**Název práce:** Intra- and inter-speaker variability of Czech noise segments

**Rok odevzdání:** 2024

**Odborná úroveň práce:**

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

**Členění práce:**

logické  vykazuje mírné nedostatky  vykazuje závažnější nedostatky  nevhodné

**Věcné chyby:**

téměř žádné  vzhledem k rozsahu přiměřený počet  méně podstatné četné  závažné

**Zvolená metoda:**

původní a adekvátní  vhodně zvolená  nepřilíš vhodná  nevhodně zvolená

**Rozsah práce:**

příliš velký  přiměřený tématu  dostatečný  nedostatečný

**Rozsah použité literatury:**

nadprůměrný  průměrný  podprůměrný  nevyhovující

**Způsob používání odborné literatury:**

adekvátní  částečně nevyhovující – viz podrobné zhodnocení  podezření na plagiát

**Grafická a formální úroveň:**

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

**Jazyková úroveň:**

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

**Tiskové chyby:**

téměř žádné  vzhledem k rozsahu a tématu přiměřený počet  četné

**Celková úroveň práce:**

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

**Zadání práce:**

mimořádně náročné  náročnější  průměrně náročné  lehčí  nedostatečně náročné

splněno  splněno s drobnými výhradami  splněno s většími výhradami  nesplněno



### **Podrobné zhodnocení:**

Ačkoliv jsem zastáncem širšího uvedení do tématu včetně vysvětlení obecně známých skutečností, které pomohou pochopit zpracované téma diplomové práce i čtenářům z jiných oborů, zde je poměr „učebnicového textu“ ze začátku možná přece jen již příliš veliký. Prvních 15 stran je věnováno obecně problémům forenzní fonetiky porovnávání hlasů, přičemž z velké části se jedná v podstatě o lehký výtah základních tezí převážně ze dvou prací – Jessen 2008 a Rose 2002. V zásadě by se dalo shrnout, že pro cíle této diplomové práce je z tohoto úvodního textu pouze důležitý závěr, že pro forenzní fonetiku je potřeba brát v potaz nejenom podobnost nahrávek, ale také jejich vyjimečnost.

Ale i následujících 11 stran obsahuje obecně velice známé fakty o frikativách a k prvnímu bodu zadání – prostudovat relevantní literaturu týkající se role šumových segmentů při srovnávání mluvcích – se autorka dostává až na str. 26. Nebylo by to špatně, pokud by zde nebyl zjevný nepoměr, protože tomuto bodu se věnuje skutečně jen na několika stranách a rozsah literatury je zde velice malý.

K teoretické části uvádím několik poznámek:

- Str. 11: uvedení vzorce pro likelihood ratio jako obrázek, navíc silně rastrovaný, není vhodné. Nikde dále v práci se navíc s likelihood ratio nepočítá.
- Figure 2 – 5 a 7 – 8 jsou evidentně bez úpravy 100% převzaty z děl, která jsou sice v textu citovaná, nicméně v popisu obrázků musí být v takovém případě převzetí výslovně uvedeno.
- Str. 14: chapter 1.2.2 -> pravděpodobně 2.2.2.
- Str. 17, kde se hovoří o fonému /v/ v češtině, by bylo vhodné zmínit, že často bývá realizován také jako aproximanta.
- Str. 18: citovaná práce Laufer 1991 chybí v seznamu referencí.
- Str. 23: section 1.1 -> opět pravděpodobně špatné číslování.
- Str. 26:  $k = 0, 0.5, 1, 0.5 \dots 1/2(N-1)$  – místo druhého 0.5 mělo být 1.5.
- “any sinusoid with no phase is a cosine wave” – nesouhlasím, to by byla sine wave.

Praktická část není příliš silná – autorka dostala k dispozici nahrávky a skripty pro extrakci parametrů. Přínos je tedy v manuálním zarovnání hranic třetiny nahrávek a tvorba deskriptivních grafů, které jsou výhradně zaměřeny na kvantily (vypsání následně i formou tabulek). Další statistika a hlubší interpretace chybí. K této části opět uvádím seznam poznámek:

- Str. 33: nahrávky na mobilních telefonech – není uvedena šířka pásma těchto originálních nahrávek. Může být u každé osoby jiná a často ani nemusí odpovídat plnému rozsahu danému vzorkovací frekvencí, která též není uvedena.
- Str. 34: “30 ms window around the temporal midpoint” – není uvedeno, jaké segmentační okno bylo zvoleno. To má přitom zásadní dopad na spektrální vlastnosti. “Transformed into a spectrum” – není uveden konkrétní postup a parametry metody. Chybí zde tedy důležité informace pro replikaci, případně přiložený skript.
- Str. 37: obsírný popis, proč nebyly pro hlásky v grafech použity IPA značky, musel dát jistě více práce, než tyto symboly ve výsledných grafech upravit. Rozhodnutí ale respektuji, ovšem pokud se už použily značky š a ř, nerozumím, proč bylo použito x místo ch.
- Str. 38: Figure 12 zobrazuje mimo jiné PSD bez vysvětlení této zkratky a bez uvedení parametrů této metody; ty mají přitom výrazný dopad na interpretaci daného grafu.
- Str. 41 – celá sekce 4.1: hovoří se zde o populačních charakteristikách, které by mohly být užitečné pro forenzní analýzy. O samotné populaci se v práci ovšem nic neuvádí – o jak staré lidi



se jednalo, jakého vzdělání, z jaké oblasti? V textu práce se akorát dozvídáme, že se jedná o muže a že nahrávky autorka dostala od svého vedoucího.

- Str. 42 – celá sekce 4.1: vzhledem k neuvedenému frekvenčnímu rozsahu nahrávek pořízených mobilem je těžké brát tyto grafy jako obecně použitelné, jelikož frekvenční rozsah nahrávek má na tyto parametry zásadní dopad. Škoda také, že nebylo zohledněno segmentální okolí, protože to může mít také vliv na uvedené charakteristiky.
- Str. 58 – Fig. 31: zobrazovaná veličina evidentně vykazuje výraznou nesymetrii. Má smysl se u takovýchto dat bavit o kvantilech a outliers? Transformace logaritmem by jistě velice pomohla.
- Str. 59 – Fig. 32: vzhledem k GDPR nepovažuji za vhodné zobrazovat zkratky jmen mluvčích. Obrázku by též velice pomohlo seřadit mluvčí dle jejich střední hodnoty než dle abecedy jejich zkratek.
- Str. 70 a 71 – Fig. 42 a 43: zobrazované veličiny opět evidentně vykazují výraznou nesymetrii. Obzvláště ve Fig. 43 není prakticky nic vidět, skupiny nelze opticky porovnávat. Opět bych doporučil transformaci logaritmem.

#### **Otázky k obhajobě a náměty do diskuze:**

- Na str. 25 je uvedeno, že pozitivní čtvrtý moment indikuje relativně vysokou špičatost. Co je zde referencí, tj. co odpovídá nulové špičatosti?
- V textu k Figure 12 je uvedeno, že DCT0 reflektuje průměrnou amplitudu spektra, DCT1 jeho lineární sklon atd. V grafu je vidět u každého DCT „koeficientu“ frekvenčně závislá křivka v decibelech. Jak máme takový graf číst? Kde zde nalezneme konkrétní koeficient a jaká je tedy jeho hodnota? To samé i pro Fig. 13 – 15. V textu práce toto vůbec není vysvětleno či jakkoliv rozvedeno.
- Fig. 40 a text na str. 67: čím si vysvětlujete jiné pořadí mediánů CoG v telefonním přenosu NB?

#### **Práci**

doporučuji  nedoporučuji uznat jako diplomovou.

#### **Navrhuji hodnocení stupněm:**

výborně  velmi dobře  dobře  neprospěl/a

V Českém Brodě, 5. září 2024

Ing. Tomáš Bořil, Ph.D.