

Univerzita Karlova v Praze  
Filozofická fakulta  
Fonetický ústav  
Mgr. Radek Skarnitzl, Ph.D.

**Posudek vedoucího práce**  
na diplomovou práci Jitky Vaňkové  
**Spectral properties of the source signal as speaker-specific cues**

Diplomová práce Jitky Vaňkové se zabývá využitelností krátkodobých a dlouhodobých spektrálních charakteristik pro účely identifikace mluvčího. Z hlediska výběru tématu se jedná o dobře formulovaný cíl: výzkum vlastností řečového signálu využitelných pro identifikaci mluvčího je v současné době velmi aktuální nejen ve Fonetickém ústavu FF UK, ale celosvětově. Práce je psaná anglicky.

Předkládaná diplomová práce je rozdělena do šesti hlavních kapitol. První teoretická kapitola nastiňuje základní východiska forenzní fonetiky: autorka popisuje vývoj tohoto oboru, zdroje variability mezi jednotlivými mluvčími a řečové vlastnosti na různých rovinách, které lze podle dosavadního výzkumu využít k rozpoznání mluvčího. Druhá úvodní kapitola představuje dlouhodobé a krátkodobé spektrální vlastnosti řečového signálu, tedy různé parametrizace dlouhodobého spektra (LTAS) a různé způsoby zachycení krátkodobého spektrálního sklonu v rámci jednoho vokálu. Následující dvě kapitoly již uvádějí vlastní výzkum diplomantky, konkrétně jeho cíle a metodologické pozadí. Následuje obsáhlá kapitola popisující výsledky měření a závěrečné shrnutí a diskuse. V příloze jsou uvedeny texty nahrávek, z nichž analýzy vycházely.

Diplomovou práci Jitky Vaňkové považuji za velmi dobře provedenou studii, ačkoli obsahuje některé nedostatky. V úvodních kapitolách se jedná spíše o formulační neobratnosti. Na straně 22 například autorka hovoří o automatických metodách rozpoznávání mluvčího využívajících GMM; z formulace by se mohlo zdát, že Hollienův systém SAUSI patří právě k těmto metodám. Na straně 33 je zmíněn bayesovský způsob statistického vyhodnocování, což bez bližšího vysvětlení nepovažuji za zcela vhodné. Podobně se domnívám, že popis kvalifikačních požadavků na osoby provádějící rozpoznávání mluvčího (str. 26-27) do oddílu zaměřeného na sluchovou analýzu nepatří. V některých pasážích by textu prospělo jednotlivé části popisu více propojit. Autorka například do popisu LTAS na straně 39 zařadila parametr L1-L0, který je zároveň předmětem následujícího oddílu, ovšem v mírně změněné podobě (jako H1-A1). Na straně 21 jsou zmíněny odlišné způsoby vnímání známých a neznámých hlasů podle Kreimanové a Sidtisové (2011); považoval bych za vhodné se k tomuto poznatku vrátit na stranách 23-24, kde diplomantka popisuje přesnost sluchové identifikace mluvčího v závislosti na tom, zda posluchač mluvčího zná či nikoli – právě proto, že tradiční prameny zaměřené na forenzní fonetiku tento aspekt zcela pomíjejí.

Jinak ale představují teoretické kapitoly předložené diplomové práce velmi kvalitní uvedení do problematiky forenzní fonetiky a fonačních modifikací. Autorka postupuje logicky a systematicky, věnuje se klíčovým aspektům bez přílišné obsáhlosti. Jisté rezervy spatřuji jen v uvádění referencí:

v několika případech citovaný zdroj není uveden v seznamu literatury (namátkou Broaders, 2001 na str. 12 a 21; Fant, 1960 na str. 19; Laver, 1980 na str. 36; Pollák *et al.*, 2007 na str. 48). Dále jsou nesprávně citovány zdroje Kreiman a Gerratt (2003) na str. 35 a Hillenbrand *et al.* (1994) na straně 41; Bruce Gerratt je mimochodem uváděn jako Garrett.

Co se týče samotného výzkumu, který Jitka Vaňková v rámci diplomové práce realizovala, je především potřeba zdůraznit, že práce zahrnovala – kromě tradičně požadované segmentace – časově velmi náročnou manuální extrakci parametrů H1-H2, H1-A1 a H1-A3, při níž bylo nutné střídavě porovnávat několik zobrazení. Tohoto úkolu se diplomantka zhostila spolehlivě a přesně. Presentaci výsledků považuji za příkladně popsany výzkum, je důkladný a zároveň přehledný, autorka postupuje systematicky, používá odpovídajících statistických metod a grafické zobrazení výsledků je přehledné a nápadité. Výsledky jsou srovnávány s odpovídajícími výsledky z literatury. Pečlivé je i závěrečné srovnání krátkodobých a dlouhodobých parametrů. Popisu výsledků tak lze vytknout jen drobnosti: např. na straně 62 je uvedeno, že mluvčí DAMA se významně liší od sedmi mluvčích, a hned v následující větě je u též mluvčí uvedena hodnota pěti mluvčích. Na závěr zmíním vysokou kvalitu anglického textu - celá práce je až na drobné chyby napsána pěkným jazykem – a považuji za nutné vyzdvihnout pečlivý, poctivý a skromný přístup Jitky Vaňkové k výzkumu.

Byl bych rád, kdyby se diplomantka při obhajobě pokusila o obecnější interpretaci dvou dílčích výsledků: 1) Měla by vysvětlení pro vyšší úspěšnost diskriminace u vokálů /i e a/ než /o u/? 2) Jak by autorka dala do souvislosti účinnost vokálů pro rozlišení mluvčích v závislosti na povaze přízvuku (tedy přízvučná, popřízvučná, nepřízvučná) na jedné straně a současné znalosti o vlastnostech přízvuku v češtině na straně druhé?

Uzavírám, že předkládaná diplomová práce Jitky Vaňkové splňuje svým obsahem i rozsahem požadavky kladené na magisterskou práci. Zmíněné nedostatky jsou spíše marginální povahy. Vzhledem k výše uvedeným skutečnostem ji doporučuji k obhajobě a navrhuji hodnocení **v ý b o r n ě**.

V Praze dne 27.1.2012

.....

Mgr. Radek Skarnitzl, Ph.D.

Fonetický ústav FF UK