

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

## **Posudek vedoucího práce**

na diplomovou práci Kristiana Urbana

### **Syntéza znělostních charakteristik českých exploziv a jejich předběžné percepční ověření**

Diplomová práce Kristiana Urbana se věnuje znělosti u exploziv a možnostem simulování různých stupňů znělosti pomocí parametrické syntézy řeči. Protože přesná kontrola akustických parametrů nezávisle na sobě je klíčová pro sestavování poslechových testů zaměřených na percepční dopad takových parametrů, představuje tento výzkum především z metodologického hlediska důležitý počín.

Předkládaná diplomová práce je rozdělena do sedmi hlavních kapitol. V první kapitole se autor stručně zabývá komunikací mezi člověkem a počítačem obecně, druhá kapitola se zaměřuje již přímo na řečovou syntézu. Diplomant nastiňuje historii řečové syntézy i hlavní moderní přístupy ke tvorbě řečového zvuku počítačem. Ve třetí kapitole se autor podrobněji věnuje znělosti: popisuje hlasivky a způsob jejich kmitání, koordinaci laryngálních svalů při přechodu ze znělých hlásek na neznělé a přibližuje i filtrovou teorii produkce řeči. Poslední kapitola teoretického úvodu nejprve odlišuje konsonantickou a vokalickou artikulace a následně se zaměřuje na explozivy, především na explozivy znělé, jejichž produkce vyžaduje specifická nastavení na úrovni glotální i supraglotální.

Úvodní kapitoly jsou psány živým a poutavým, někdy však až příliš epickým jazykem. Domnívám se, že popis syntézy řeči v druhé kapitole by místy byl pro laického čtenáře jen obtížně pochopitelný; autor některé základní koncepty často zmiňuje jen implicitně, jako by předpokládal jejich znalost na straně čtenáře. Hovoří například o difonové syntéze francouzštiny za použití 1200 jednotek (str. 16) a bez komentáře přechází k výběru jednotek; vztah mezi těmito dvěma formami inventáře jednotek z textu není zřejmý. Na straně 18 se píše o parametrické reprezentaci řečového signálu, avšak bez bližší zmínky o podobě této reprezentace; to může být pro čtenáře opět těžko představitelné. Z popisu na straně 19 jasně nevyplývá, proč je „každý typ afektované mluvy“ spojen s vytvářením nové databáze, zatímco u parametrické syntézy nikoli.

Třetí kapitolu zabývající se znělostí považuji za zdařilejší a mám k ní výtek méně. Z popisu na stranách 25–26 by se mohlo zdát, že Bernoulliho princip a myoelasticko-aerodynamická teorie jsou na sobě nezávislé koncepty, ačkoli van den Bergova teorie Bernoulliho princip přímo využívá. Na straně 31 je zřejmě vlivem nepozornosti chybně uvedeno, že „utlumeným frekvenčním pásmům se říká formanty“. Obrázek 5 představující demografickou křivku považuji spíše za redundantní; obrázky 5 a 6 jsou navíc číslovány obráceně. Obecně se zmíním ještě o způsobu citování zdrojů: uvádění celého názvu díla (občasné, nikoli systematické) neodpovídá konvencím, a text navíc znepréhledňuje.

Pátá kapitola diplomové práce Kristiana Urbana popisuje HLsyn, program pro hybridní parametrickou syntézu řeči, jehož autor využil pro tvorbu řečových vzorků. Detailně zde představuje jednotlivé parametry, jichž systém využívá. Chvályhodný je především oddíl 5.3, v němž důkladně popisuje způsob manipulace časovými údaji v programu; v manuálu k programu manipulace časové osy není popsána, což by mohlo navodit dojem, že ani není možná. Autor se s tímto nedostatkem programu elegantně vypořádal. Má jediná poznámka k této kapitole se týká parametru  $ue$ , u nějž je uvedena jednotka  $\text{cm}^3$ , správně je však  $\text{cm}^3/\text{s}$  (str. 41).

V šesté kapitole Kristian Urban přibližuje způsob syntézy jak nosné fráze, tak i cílových pseudoslov s explozivami v intervokální pozici. Detailně popisuje manipulaci třemi parametry, pomocí nichž z plně znělých exploziv vytvářel mezistupně na škále znělá–neznělá. Kapitola se rovněž věnuje sestavení percepčního testu, pomocí nějž byl vliv jednotlivých manipulací zkoumán. Autor na obrázku 9 (str. 48) ukazuje příklady syntetizovaných spojení, avšak jejich popis neodpovídá zobrazení: ani jeden oscilogram nelze považovat za znělý; znělosti se nejvíce blíží vlna b., která je však popsána jako neznělá. Práce bohužel neobsahuje CD se zvukovými ukázkami syntetizovaných spojení, a tak si čtenář nemůže jednotlivé položky porovnat. Na straně 49 by bylo vhodné znovu odkázat na Přílohu A se záznamovým archem percepčního testu.

V sedmé kapitole diplomant popisuje výsledky percepčního testu a následující kapitola výsledky shrnuje a komentuje. Prezentaci výsledků by výrazně pomohlo grafické zobrazení, tedy větší počet dílčích grafů než jen ten uvedený na straně 58.

K výsledkům a jejich interpretaci mám ještě jednu vážnější výhradu. Na straně 48 se hovoří, že by zkrácení okluzivní části explozivy mohlo způsobit percepční dojem ztráty znělosti. Protože výsledky naznačují tendenci opačnou – a především očekávatelnou – byl bych rád, aby diplomant při obhajobě motivaci pro svou původní hypotézu vysvětlil.

Uzavírám, že Kristian Urban provedl zajímavou studii za použití v českém prostředí málo známé metody, parametrické řečové syntézy. V percepčním experimentu zároveň ověřil vliv tří parametrů na vnímanou míru znělosti u exploziv. I přes zmíněné nedostatky předkládaná diplomová práce splňuje svým obsahem i rozsahem požadavky kladené na magisterskou práci. Vzhledem k výše uvedeným skutečnostem ji doporučuji k obhajobě a navrhuji hodnocení **velmi dobře**.

V Praze dne 9.9.2013

.....

doc. Mgr. Radek Skarnitzl, Ph.D.

Fonetický ústav FF UK