

---

**TECHNICKÁ UNIVERZITA V LIBERCI**  
Fakulta mechatroniky, informatiky a mezioborových studií

Studijní program: B2612 - Elektrotechnika a informatika  
Studijní obor: 2612R011 - Elektronické informační a řídicí systémy

**Jazyková adaptace systému MyVoice  
z češtiny do slovenštiny**

**Language adaptation of the MyVoice system  
from Czech to Slovak language**

**Bakalářská práce**

Autor: **Jan Pavlík**  
Vedoucí práce: Ing. Petr Červa, Ph.D.

V Liberci 25.5.2009

---

## Prohlášení

Byl(a) jsem seznámen(a) s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 o právu autorském, zejména § 60 (školní dílo).

Beru na vědomí, že TUL má právo na uzavření licenční smlouvy o užití mé bakalářské práce a prohlašuji, že **s o u h l a s í m** s případným užitím mé bakalářské práce (prodej, zapůjčení apod.).

Jsem si vědom(a) toho, že užít své bakalářské práce či poskytnout licenci k jejímu využití mohu jen se souhlasem TUL, která má právo ode mne požadovat přiměřený příspěvek na úhradu nákladů, vynaložených univerzitou na vytvoření díla (až do jejich skutečné výše).

Bakalářskou práci jsem vypracoval(a) samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Datum

Podpis

## **Poděkování**

Touto cestou bych chtěl vyslovit svoje poděkování vedoucímu mé bakalářské práce panu Ing. Petru Červovi, Ph.D. za odborné vedení, poskytnutí literatury, cenné rady a pomoc při vypracování mé bakalářské práce. Dále bych chtěl poděkovat Marie Solčanské a Františkovi Dritomskému za pomoc při překládání do slovenštiny a za namluvení velkého množství zvukových nahrávek.

## **Abstrakt**

Hlavním cílem této bakalářské práce bylo adaptování českého programu MyVoice na slovenský jazyk. Tento program je primárně určen pro handicapované lidi a slouží k ovládnutí počítače hlasem. Díky mezijazykové adaptaci dokážeme využít soubory a zdroje dostupné v jednom jazyce k počítačovému rozpoznávání v jiných jazycích, v tomto případě jsem tedy použil mezijazykovou adaptaci k přetvoření českého programu na jeho slovenskou verzi. Prvním a stěžejním krokem k vytvoření slovenské verze systému MyVoice bylo namapování slovenských specifických fonémů na již existující sadu českých fonémů. Dále bylo potřeba přeložit všechny hlasové povely systému MyVoice do slovenštiny a nakonec vytvořit zvukovou databázi slovenských slov na testování výsledného produktu.

Tuto práci jsem rozdělil na několik částí, v první jsem se zabýval teoretickými poznatky týkajícími se mezijazykové adaptace, fonetické transkripce, fonologie a jazykovými rozdíly mezi češtinou a slovenštinou. V další části jsem popsal vlastní proces tvorby slovenské verze systému MyVoice a v závěrečné části naleznete popis a výsledky testů, které ukázali jak úspěšně se MyVoice adaptoval na slovenský jazyk.

## **Abstract**

The main goal of this work was adapting Czech program MyVoice to Slovak language. This program is primarily determined to help handicapped people and it serves to voice control of a computer. Thanks to cross-lingual adaptation we can use files and resources available in one language to computer recognition in other languages, in this case I used cross-lingual adaptation to transform the Czech program into its Slovak version. The first and fundamental step to create the Slovak version of the system MyVoice was mapping Slovak-specific phonemes to the already existing set of Czech phonemes. Further, I had to translate every voice order used in MyVoice to Slovak language and finally I created a voice database of Slovak words used for testing of the final product.

I separated this work into several parts, in the first part I dealt with theoretical knowledge about cross-lingual adaptation, phonetic transcription, phonology and language differences between Czech and Slovak language. In the next part I described the process of creating the Slovak version of the system MyVoice and in the final part you can find a description and results of tests, which showed how successfully the system MyVoice was adapted to Slovak language.

## Obsah

Prohlášení .....	3
Poděkování .....	4
Abstrakt .....	5
Abstract.....	6
Obsah .....	7
Seznam obrázků .....	9
Seznam tabulek .....	10
1 Úvod.....	12
2 Fonetická abeceda a transkripce .....	13
2.1 Foném, alofón a fón .....	13
2.2 IPA a její modifikace .....	14
3 Čeština a slovenština v počítačovém rozpoznávání.....	18
3.1 Rozdíly a podobnosti ve slovnících.....	19
3.2 Rozdíly a podobnosti ve fonologii .....	19
4 MyVoice.....	22
4.1 Rozpoznávač české řeči použitý v systému MyVoice .....	22
4.2 Příkazové skupiny v systému MyVoice .....	23
4.3 Ovládání systému MyVoice.....	23
5 Tvorba slovenské verze MyVoicu .....	26
5.1 Překládání hlasových povelů .....	26
5.2 Mapovací tabulka a pravidla přepisu.....	27
5.3 Program na převod české verze MyVoicu na slovenskou.....	30
6 Testy.....	32
6.1 Sbírání dat .....	32
6.2 Průběh testů .....	33

6.3 Výsledky testů .....	34
7 Závěr .....	36
Seznam použité literatury .....	38

## **Seznam obrázků**

Obr. 3.1: Skupiny jazykových rodin pro češtinu a slovenštinu

Obr. 4.1: MyVoice – základní skupiny

Obr. 5.1: Struktura ini souboru

Obr. 6.1: Nahrávací program

Graf 6.1: Výsledky testů



## **Seznam tabulek**

Tab. 2.1: Inventář fonémů mluvené slovenštiny

Tab. 2.2: Fonetická abeceda ARPAbet

Tab. 2.3: PAS - Fonetická abeceda pro slovenštinu

Tab. 3.1: Inventář českých fonémů

Tab. 3.2: Inventář slovenských fonémů

Tab. 5.1: Mapovací tabulka

Tab. 6.1: Výsledky testů



## 1 Úvod

Jednou z věcí, kterou se člověk odlišuje od zvířat, je jeho snaha usnadnit si práci a vymýšlet pořád nové a nové věci. Hlavním produktem této snahy v nedávné minulosti byl vynález počítače. Příchod počítačů a jejich rychlý pokrok umožnil jejich využití v různých oblastech a nabízí širokou škálu uplatnění. Počítač je v současnosti největším pomocníkem při práci a velkým zdrojem zábavy pro většinu lidí na naší planetě, pomáhá nám při studiu, slouží jako zdroj informací nebo k bleskové komunikaci mezi lidmi. Život bez počítače je pro podstatnou část z nás již nepředstavitelný.

Ale co ti lidé, kteří mají nějaký handicap, kvůli kterému nemohou počítač ovládat standardním způsobem, ale přesto chtějí být moderními lidmi, kteří využívají všechny dostupné výhody počítače. Právě pro tyto lidi byl na Technické univerzitě v Liberci vyvinut systém MyVoice, který je českým ekvivalentem rozpoznávačů řeči, které byly vyvinuty pro světové jazyky jako je angličtina nebo francouzština. Tento program slouží jako hlasový ovladač, který dokáže nahradit běžné ovládací prvky počítače, jako je klávesnice a myš, a dovolí nám vykonávat všechny úkony pouze tím, že do mikrofону říkáme žádané příkazy. Protože byl program vyvinut v České republice, může prozatím sloužit jen lidem mluvícím česky. Díky mezijazykové adaptaci je však možné program předělat i na jiné jazyky. Toto již bylo testováno na španělštině a v této práci je popsána tvorba slovenské verze programu.

Mým úkolem bylo nejen přeložit veškeré hlasové povely do slovenského jazyka, ale i, kvůli existenci slovenských znaků, které čeština nemá, vymyslet vhodné namapování slovenské výslovnosti na již existující sadu českých fonémů. Dále jsem se zabýval experimenty, na které byly potřeba nahrávky slovenských slov. Pomocí nich jsem pak testoval úspěšnost rozpoznávání. Nakonec jsem testoval adaptaci systému MyVoice na jednoho konkrétního mluvčího, který nejprve nahrál adaptační slova, která se předložila rozpoznávači. Po adaptaci se zvýšila přesnost rozpoznávání povelů této konkrétní osoby. Experimenty jsem prováděl za pomoci dvou mluvčích, a to muže a ženy.

## **2 Fonetická abeceda a transkripce**

### **2.1 Foném, alofón a fón**

Tak jako se při psaní textu používá abeceda, která nám rozdělí celá slova na menší celky – písmena (grafémy), tak i při mluvení můžeme používat abecedu, která vyslovená slova rozdělí na menší části, které nazveme fonémy. Fonémy jsou vlastně reprezentací zvuku a jsou často označovány za nejmenší jednotku řeči. Změna jedné části slova z jedné fonémy na jinou může vyprodukovat úplně jiné slovo nebo smysl. Ve světových jazycích se běžně používá od 12 do 60 fonémů. Fonémy dále tvoří slabiky a slabiky tvoří slova.

Koartikulace je jev, kdy se foném může měnit podle kontextu, v kterém je vysloven. Konkrétní akustika dané fonémy potom závisí například na předcházejícím a následujícím fonému, rychlosti řeči, intonaci a podobně.

Foném použitý v různých slovech může být reprezentována jiným zvukem v závislosti na koartikulaci v jeho fonetickém okolí. Tuto variantu fonému nazýváme fón a všechny různé fóny jednoho fonému nazýváme alofón. Různé alofóny jsou použité podle fonetického kontextu a můžou například vymezovat hranice slova nebo slabiky.

Všeobecně je řeč neharmonický, nestacionární signál, proto je jen velmi těžko modelovatelná. Z mnohých pokusů však vyplývá, že po dobu krátkého časového intervalu můžeme řeč považovat za stacionární. Na fonémy, respektive na jejich alofóny, se můžeme dívat jako na kus stacionární informace, která nám může pomáhat nejen při klasifikaci slov, ale i při rozpoznávání a popisu jejich správné výslovnosti. [4, 7]

## 2.2 IPA a její modifikace

International Phonetic Alphabet (IPA) vznikla jako výsledek dlouholeté snahy jazykovědců celého světa. Jedná se o systém fonetické notace, jejíž cílem je poskytnout standardní a jednotný způsob reprezentace jakéhokoliv zvuku v jakémkoli jazyku. Ve svojí základní nejrozšířenější formě obsahuje přibližně 107 základních symbolů a 55 modifikátorů.

Fonetická transkripce je jednoznačný způsob zápisu zvuků hovorové řeči za pomoci fonetické abecedy, při tom IPA je pro tento účel tou nejvšeobecnější abecedou. Výběr vhodné fonetické abecedy závisí na tom, na kolik přesný popis chceme vytvářet – abeceda fonémů pro volnější popis nebo abeceda alofónů pro popis detailnější. Tento popis se používá na přípravu rozpoznávacích nebo trénovacích vzorů, které se potom používají k trénování rozpoznávání nebo na testování úspěšnosti rozpoznávání slov.

Slovenská modifikace systému IPA se jmenuje SPA (Slovak Phonetic Alphabet). Tato modifikace je vhodná pro potřeby podrobnějšího zapisování slovenštiny a jejím východiskem je inventář znaků slovenského pravopisného systému použitím některých znaků IPA. Inventář fonémů mluvené slovenštiny, zapsaný pomocí znaků SPA je uvedený v následující tabulce (Tab. 2.1).

Tab. 2.1: Inventář fonémů mluvené slovenštiny [7]

Fonéma	Příklad	Fonéma	Příklad	Fonéma	Příklad
a	papier	˘a	piatok	˘uo	vôľa, kôň
á	pás	˘e	spievať	p	popol
ã	pãč, mäso	˘u	cudzíu	r	para, Andrej
b	žaba	j	dvaja, prijímať	r	prst, vrch
c	cena, pocta	˘i	kraj, krajský, zjesť	˘r	vřba
č	oči, mačka	k	páka, krk	s	oscm, strom
d	voda, dom	l	skala, vlak	š	košefa, prášit'
ď	deti, háďa	l	vlk	t	vata, tento
ž	hádzat', medza	l	lipa, liek, ľad	˘t	Mat'o, platit'
ž	džem, Hodža	˘f	vľča	u	puto
e	pero	m	mama	ũ	ulicou, pravda, otcov
é	dcéra, nové	ŋ	amfiteáter	ú	púpava
f	fajka, včela, vpiť	n	rana, nos	v	slovo, v_oku
g	gágat', agát	ŋ	panský, inžinier	w	vdova, šeflekár
h	Praha, hrýzť	ŋ	tenký, cengat'	z	váza, zbierať
x	chata, pochvala	ɸ	bronchitída	ž	veža, žaba
γ	vrch_hory	ň	vaňa, niest'	>š	vyšši
i	pivo, dym	o	popol	>ň	denne
í	pískat', nový	ó	pól	>n	cenný

Jak IPA tak SPA používají pro zápis zvuků řeči znaky, které nekorespondují se znaky, které se používají v počítačovém zpracování textu. Pro potřeby analýzy a porovnávání psané a mluvené podoby jazyka byly proto v systémech automatického rozpoznávání řeči zavedené jiné způsoby fonetické transkripce, která vychází z mezinárodně standardizované „počítačové abecedy znaků“ ASCII. Příkladem takové fonetické abecedy je ARPAbet, kterou pro americkou angličtinu vyvinula agentura ARPA (Advanced Research Projects Agency), a která je v následující tabulce (Tab. 2.2).

Tab. 2.2: Fonetická abeceda ARPAbet[7]

IPA	ARPAb		Example	IPA	ARPAb		Example	IPA	ARPAb		Example
	single-symbol	upper-case			single-symbol	upper-case			single-symbol	upper-case	
i	i	IY	heed	i	X	IX	roses	m	m	M	mom
ɪ	I	IH	hid	p	p	P	pea	n	n	N	noon
a	e	EY	hayed	b	b	B	bat	n	G	NX	sing
ɛ	E	EH	head	t	t	T	tea	l	l	L	love
æ	@	AE	had	d	d	D	deep	l	L	EL	cattle
ɑ	a	AA	had	k	k	K	kick	m	M	EM	some
ɔ	c	AO	hawed	g	g	G	go	n	N	EN	son
o	o	OW	hoed	f	f	F	five	t	F	DX	batter
u	U	UH	hood	v	v	V	vice	ʔ	Q	Q	(glottal stop)
ʊ	u	UW	who'd	θ	T	TH	thing	w	w	W	want
ɚ	R	ER	heard	ð	D	DH	then	j	Y	Y	yard
e	x	AX	ago	s	s	S	so	r	r	R	race
ʌ	A	AH	mud	z	z	Z	zebra	tʃ	C	CH	church
aɪ	Y	AY	hide	ʃ	S	SH	show	dʒ	J	JH	just
aʊ	W	AW	how'd	ʒ	Z	ZH	measure	ɹ	H	WH	when
ɔɪ	O	OY	boy	h	h	HH	help				

ARPAbet existuje ve dvou verzích – jednoznakové a dvouznakové. Jednoznaková verze používá malá a velká písmena z tabulky ASCII. Dvojnaková verze používá jen velká písmena. Fonetická transkripce například textu „chocolate pudding“ pomocí jednoznakové verze ARPAbet je „CAKxllt pUdG“.

Fonetická abeceda ARPAbet sice vyhovuje pro angličtinu, inventářem svých znaků však nezahrnuje ostatní jazyky světa. Vývoj multilingvistických respektive národních systémů automatického rozpoznávání řeči si vyžaduje také mezinárodní fonetickou abecedu, která by obsahovala znaky pro všechny jazyky světa. Příkladem pokusů v uskutečnění tohoto cíle jsou fonetické abecedy WorldBEet a X-SAMPA .

SAMPA (Speech Assessment Methods Phonetic Alphabet) je fonetická abeceda, která byla navržena mezinárodní skupinou fonetiků v rámci programu ESPRIT v letech 1987-1989 pro jazyky Evropy. V prvním přiblížení pro jazyky zemí EU (dánština, holandská, angličtina, francouzština, němčina a italština). Za nedlouho byla doplněna o norštinu, švédštinu, řečtinu, portugalskou a španělskou. V rámci projektu BABEL byla v ostatních letech doplněna o jazyky některých států bývalého východního bloku, jako jsou bulharština, polština, estonština, maďarština, rumunština, chorvatština, ruština a slovinština.

Jak je však z výše uvedeného přehledu zřejmé, X-SAMPA neobsahuje všechny evropské jazyky. Mezi nimi je i čeština a slovenština. Proto je předpoklad, že v blízké budoucnosti bude X-SAMPA rozšířena i o tyto jazyky. Existují pokusy o zavedení vlastní, počítačově zpracovatelné fonetické abecedy. Příkladem těchto snah je PAC (Phonetic Alphabet for Czech) pro češtinu a z něj odvozená fonetická abeceda PAS (Phonetic Alphabet for Slovak) pro slovenštinu, uvedená v následující tabulce (Tab. 2.3).

Tab. 2.3: PAS - Fonetická abeceda pro slovenštinu[7]

Fonéma	PAS		Fonéma	PAS		Fonéma	PAS		Fonéma	PAS	
	I	II		I	II		I	II		I	II
a	a	a	x	X	x	m	M	mg	t	t	t
á	á	aa	ɣ	H	xh	n	n	n	t'	t'	tj
ä	ä	ae	i	i	i	ŋ		ns	u	u	u
b	b	b	í	í	ii	ɳ	N	ng	ü		
e	e	e	ja	A	ia	ʌ		nx	ú	ú	uu
ě	ě	eh	je	E	ie	ň	ň	nj	v	v	v
d	d	d	ju	U	iu	o	o	o	w	w	w
ď	ď	dj	j	j	j	ó	ó	oo	z	z	z
ɹ	C	dz	ɹ	y	y	õo	õ	uo	ž	ž	zh
ž	Č	dzh	k	k	k	p	p	p	-		
e	e	e	l	l	l	r	r	r	>š	Š	ssh
é	é	ee	l			ɹ	R	rx	>ň	Ň	nnj
f	f	f	ɹ	ɹ	lj	ɹ	ɹ	rr	>n	N	nn
g	g	g	l	l	ll	s	s	s			
h	h	h	m	m	m	š	š	sh			

Podobně jako ARPAbet i fonetické abecedy PAC a PAS obsahují dvě verze fonetického přepisu – jednoznakovou a dvouznavou. Jednoznaková verze používá znaky s diakritikou, takže fonetický přepis je čitelný i pro neoborníka. Dvouznaková verze používá jen malá písmena bez diakritiky. Fonetický přepis například čísla „dvadsať štyri“ je v jednoznakové verzi „dvacaťštiri“ a v dvouznavé verzi „dvacatjshtiri“. [4,7]



### 3 Čeština a slovenština v počítačovém rozpoznávání

Čeština a slovenština patří k západo-slovanské větvi evropských jazyků (Obr. 3.1). Jejich společný předek byl používán přibližně 15 miliony lidmi žijících na českém i slovenském území. Oba jazyky jsou považovány za velice podobné a blízce spřízněné, protože v minulosti byly oba oficiálními jazyky používanými uvnitř jednoho státu. Od roku 1993, kdy se Československo rozpadlo na dva samostatné státy, zde vyrůstá nová generace mladých lidí, kteří mají problémy porozumět řeči druhého národa. To ukazuje na to, že rozdíly mezi oběma řečmi jsou větší než se původně zdálo. Existuje ještě jeden podobný případ a to mezi španělštinou a portugalštinou. [5]



Obr. 3.1: Skupiny jazykových rodin pro češtinu a slovenštinu

### 3.1 Rozdíly a podobnosti ve slovnících

Za účelem určení kvality stupně nepodobnosti byly použity dva typy jazykových měření. Od doby kdy jsou Česká a Slovenská republika členy Evropské unie, můžeme najít mnoho EU dokumentů publikovaných v obou národních jazycích a můžeme je označit jako paralelní soubory. Z internetových stránek evropské unie si můžeme stáhnout několik významných textů a z nich vytvořit seznam slov používaných v jejich české a slovenské verzi. České texty obsahují 215 tisíc slov, z toho je 12434 jedinečných, zatímco slovenská verze obsahuje 218 tisíc slov, z toho 12217 jedinečných.

V dalších testech byly porovnávány české a slovenské slovníky zkompileované pro ASR (Automatic speech recognition) systémy. Oba slovníky měli přibližně stejnou velikost a to 170 tisíc slov vybraných podle kritéria frekvence výskytu slova. Po odebrání vlastních jmen byla velikost obou slovníků snížena na 145 tisíc položek, a z nich 29 tisíc (20%) bylo společných pro oba jazyky.

Ačkoli testy ukázaly, že 80% slovníkových inventářů je odlišných, aktuální měření nepodobnosti není tolik vysoké, provedeme-li detailnější pravopisné porovnání odpovídajících slovních párů. Mnoho rozdílů je pouze v jednu nebo ve dvou znacích, nebo v příponách kvůli mírným rozdílům morfologických vzorů. Jako příklad bych uvedl následující větu v české a ve slovenské verzi.

*CZ: Děti udělali přesně to, co jim řekli učitelé.*

*SK: Deti urobili presne to, čo im hovorili učitelia.*

[5]

### 3.2 Rozdíly a podobnosti ve fonologii

V mluvené češtině je obvykle rozlišováno 41 jednotlivých fonémů. Někteří fonetice rozpoznávají další tři dvojhlásky (tyto dvojhlásky mohou být nyní v systémech ASR modelovány pomocí kaskádních odpovídajících fonetických modelů bez ztráty rozpoznávací přesnosti).

Seznam českého fonetického inventáře se SAMPA symboly naleznete v následující tabulce (Tab. 3.1).

**Tab. 3.1: Inventář českých fonémů**

<b>Skupina</b>	<b>Znaky</b>
Samohlásky (11)	a, e, i, o, u, a:, e:, i:, o:, u:, @ (schwa)
Souhlásky (30)	p, b, t, d, c, J\, k, g,
	ts, dz, tS,dZ, r, l,
	f, v, s, z, S, Z, X, j, h\, Q\, P\
	m, n, N, J, F
Dvojhásky (3)	o_u, a_u, e_u

Pro slovenštinu existuje několik fonetických seznamů. Někteří rozlišují mezi 57 rozdílnými fonetickými jednotkami, jiní rozpoznávají 48 individuálních fonémů a 4 dvojhásky. V následující tabulce (Tab. 3.2) naleznete seznam slovenského fonetického inventáře, který je sestaven podle druhého přístupu, teda s 48 samostatnými fonémy a se 4 dvojháskami. Tabulka je postavena podobně jako tabulka předešla, takže můžeme vidět znaky, které se v obou jazycích liší (jsou vtištěné tučně).

Tab. 3.2: Inventář slovenských fonémů

Skupina	Znaky
Samohlásky (11)	a, e, i, o, u, {, a:, e:, i:, o:, u:,
Souhlásky (37)	p, b, t, d, c, J\, k, g,
	ts, dz, tS,dZ, r, l, r=, r=:, l=, l=:, L
	f, v, s, z, S, Z, X, j, h\, w, U_^, G, I_^
	m, n, N, J, F
Dvojhásky (4)	I_^a, I_^e, I_^U\, u_^o

Detailní analýza této tabulky ukáže, že symboly ‘r=’ a ‘l=’ reprezentují slabičnou verzi ‘r’ a ‘l’. Tyto dvě slabiky se vyskytují také v češtině, kde jsou považovány za alofony fonémů ‘r’ a ‘l’. Symboly ‘r=:’ a ‘l=:’ jsou pouze jejich delší formy. Podobně, slovenské fonémy označené jako ‘G’ and ‘I\_^’ jsou považovány za alofony fonémů ‘h\’ a ‘j’ v češtině. Zbývající slovenské specifické fonémy a dvojhásky můžou být přiblíženy k jejich zvukově blízkým českým protějškům. [5]

## 4 MyVoice

System MyVoice je program sloužící k hlasovému ovládní počítače. Vývoj tohoto systému byl motivován skutečností, že zde není žádný komerční nástroj tohoto typu pro handicapované uživatele. Tento případ je stejný pro mnoho jiných jazyků, pro které zatím nebyly vyvinuty systémy na rozpoznávání řeči, díky tomu, že vývoj takových systémů spotřebuje značné investice a zisky z něho jsou malé, protože počet členů cílové skupiny je značně omezený. Navzdory tomu se začali vyšetřovat možnosti přepracování softwarů jako je MyVoice na jiné jazyky. Lehčí je to samozřejmě na jazyky podobné češtině, jako je například slovenština, jsou zde ale i snahy o vytvoření komplexnějšího přístupu k přepracování systému na jazyky zvukově a jazykově mnohem rozdílnější, například Španělština. [6]

### 4.1 Rozpoznávač české řeči použitý v systému MyVoice

Rozpoznávač české řeči byl vyvíjen více než deset let na Technické univerzitě v Liberci. Jeho akustický model je založen na třístavovém levo-pravém skrytém Markově modelu (angl. Hidden Markov Model, zkr. HMM), který používá kontextově nezávislé řečové jednotky, ty jsou doplněny několika zvukovými modely. Rozpoznávač rozpoznává slovo po slově, nikoli souvislý text. V souladu s aplikačními podmínkami mohou být tyto modely buď nezávislé na mluvčím (SI = speaker independent), závislé na pohlaví (GD = gender-dependent), nebo takové, které se přizpůsobují mluvčímu (SA = speaker adapted). Rozpoznávací dekodovací jednotka používá slovník abecedně seřazených slov.

Tento rozpoznávač byl úspěšně použit pro vývoj systémů MyVoice a MyDictate. Tyto programy dovolí osobám s nějakým postižením rukou pracovat a počítačem pouze používáním několika stovek hlasových povelů. Aby toto bylo možné, MyVoice převádí mluvené příkazy na jednu nebo více základních akcí, jako například mačkání, držení a uvolňování tlačítek, nebo kombinování kláves klávesnice, dále pohyb kurzorem myši po obrazovce a následné klikání, jiné

příkazy nám umožní spouštět důležité programy, nebo psát celé věty. MyDictate je první diktovací program vyvinutý pro český jazyk. Program pracuje s velice velkým slovníkem, který zahrnuje 540 tisíc nejvíc frekventovaných českých slov a je primárně míněn jako pomoc pro pohybově handicapované uživatele. [6, 8]

## 4.2 Příkazové skupiny v systému MyVoice

MyVoice je rozdělen do několika příkazových skupin, každá z nich řeší specifický úkol. Například skupina, která řídí ovládání myši, je rozdílná od té, které nahrazuje klávesnici, ale mezi oběma můžeme snadno přecházet pomocí hlasových příkazů. Velikost těchto skupin se pohybuje v rozmezí 5 až 137 příkazy, přičemž největší z těchto skupin obsahuje hlavně jména písmen a kláves, která jsou na klávesnici počítače, což dělá rozpoznávání velice těžké, protože akustické rozdíly mezi nimi jsou velice malé. Nicméně, lepších výsledků při rozpoznávání je dosaženo, když jsou příkazy rozděleny do skupin, protože pak pro každý úkol může být definován specifický slovník. Kromě toho, dělení do skupin usnadní interakci, protože uživatel je neustále informován o tom, jaké příkazy může v daný použít. Software MyVoice v současnosti používá 60 handicapovaných lidí v České republice a jejich zprávy říkají, že chybovost rozpoznávání se pohybuje mezi 1% až 2%, tedy pouze v případě, že uživatel nemá nějakou vadu řeči. [6]

## 4.3 Ovládání systému MyVoice

Ovládání programu MyVoice je velice jednoduché. Po spuštění programu se nám na obrazovce objeví okno aplikace, kde je seznam všech dostupných základních skupin příkazů. Dále jsou zde dvě důležitá tlačítka „STOP“ a „START“, která slouží k zapnutí nebo vypnutí činnosti programu. Jak vypadá systém MyVoice hned po spuštění si můžete prohlédnout na následujícím obrázku (Obr. 4.2).



**Obr. 4.1: MyVoice – základní skupiny**

Nyní si hlasovým povelům můžete vybrat jednu ze skupin a tím určíte jaký program nebo periférii budete ovládat. Pod položkou „Spust' aplikaci“ naleznete seznam nejvíce používaných aplikací, které můžete jedním povelům spustit. Další skupina se jmenuje „Plocha“. Vyslovíte-li toto klíčové slovo dostanete se na plochu počítače. V základní skupině se dále nalézají položky sloužící k tomu, aby jste si mohli rychle otevřít nabídku start, nebo minimalizovat, maximalizovat a zavírat okna. Můžete se odtud rychle dostat k ovládání oblíbených multimedií jako je internet, televize, rádio anebo winamp.

Jednou z dalších funkcí MyVoicu je diktování textu. Diktování lze provádět třemi způsoby a to, buď odříkáváním celých slov, která jsou uvedena ve slovníku programu, nebo diktováním jednotlivých písmen, a nebo do třetice diktováním vlastních jmen, jejichž první

písmeno se zapíše do textu. Chceme-li teda zapsat písmeno “a“ použijeme vlastní jméno Adam. Tento třetí způsob je asi nejpřesnější, protože vlastní jména se snadno rozpoznávají. Samozřejmě můžeme ovládat i ostatní klávesy jako je enter, nebo backspace, čímž nám MyVoice nahradí celou klávesnici.

Další periferií, kterou nám MyVoice plnohodnotně zastoupí, je myš. Polohu kurzoru myši na obrazovce ovládáme pomocí dvojslovných povelů, první slovo udává směr a druhé o kolik pixelů se má kurzor myši posunout. Když se dostaneme na požadovanou pozici, stačí nám už jen vyslovit příkaz na požadovanou akci, jako je klik pravým nebo levým tlačítkem, či dvojklik.

System MyVoice je tedy komplexní nástroj pro ovládání počítače hlasem, jehož funkce by měli plně pokrýt všechny potřeby běžného uživatele.



## 5 Tvorba slovenské verze MyVoicu

K vytvoření jazykové adaptace systému MyVoice na slovenský jazyk je potřeba provést dva zásadní kroky. Prvním z nich je přeložení všech hlasových povelů z češtiny do slovenštiny. Druhým krokem je vytvoření mapovací tabulky, která slouží k převodu všech slovenských fonémů na odpovídající české fonémy. Jelikož má slovenština většinu znaků shodných z češtinou je nutné provést namapování jen rozdílných znaků.

### 5.1 Překládání hlasových povelů

System MyVoice obsahuje asi 500 hlasových povelů, které jsou rozděleny do skupin podle toho, jaký program nebo periferii ovládají. Každá skupina povelů je v instalačním adresáři MyVoicu interpretována jako „ini“ soubor. Soubory „ini“ jsou inicializační textové soubory sloužící hlavně k uchování nastavení nějakého programu. Mají velice specifickou strukturu, která je znázorněna na následujícím obrázku (Obr. 5.1).



```
[Main]
Count= 20
NextGroup=
InheritedGroup=
[word1]
word=Adresa
Pronunc=adresa
NextGroup=Klávesnice
RunAppPath=
KeyCount=2
Key1=115
Key2=27
[word2]
word=obľúbené
Pronunc=oblúbené
NextGroup=Klávesnice
```

Obr. 5.1: Struktura ini souboru

Na začátku každého inicializačního souboru je u položky „Count“ číslo, které nám říká kolik je v tomto souboru položek, tedy v našem případě kolik je zde hlasových povelů. Další řádky v „ini“ souboru, které byli pro moji práci důležité začínají klíčovým slovem „Word“, za touto položkou je vždy zapsán hlasový příkaz, jehož přesná výslovnost je v řádku hned pod ním. Výslovnost příkazu je uvozena klíčovým slovem „Pronunc“.

Jelikož se některé příkazy vyskytují v „ini“ souborech i vícekrát, použil jsem program, který mě vyextrahoval všechny jedinečné příkazy do jediného textového souboru. Poté jsem se již mohl zabývat překladem do slovenštiny. K překladu jsem používal hlavně česko-slovenský slovník. Protože se však v MyVoicu nevyskytují jen slova v základních tvarech, musel jsem si dohledat v gramatice slovenského jazyka jak se daná slova skloňují či časují. Nakonec jsem svůj překlad ukázal slovenskému rodilému mluvčí, který mě opravil drobné chyby.

## 5.2 Mapovací tabulka a pravidla přepisu

Jak už jsem se zmínil výše, veškeré příkazy v MyVoicu jsou uchovány v inicializačních souborech. Kromě názvu příkazů však ještě „ini“ soubory obsahují výslovnost těchto příkazů, kterou nalezneme za klíčovým slovem „Pronunc“. Abych mohl provést adaptaci systému MyVoice na slovenský jazyk, musel jsem zde napsat přesunout výslovnost všech do slovenštiny přeložených příkazů, ovšem za použití českých fonémů. Jak je vidět z obrázku (Obr. 3.1) čeština a slovenština patří do stejné větve slovanských jazyků, takže jsou si oba jazyky velmi podobné, mají mnoho stejných písmen, tudíž i stejných fonémů. Ale mají i několik rozdílných znaků, a ty jsem musel namapovat na české fonémy. Kompletní inventář slovenských znaků i s namapováním na české fonémy naleznete v následující tabulce (Tab. 5.1). Slovenská výslovnost je zapsaná pomocí SAMPA symbolů. V této tabulce si hlavně všimněte tučně zvýrazněných položek, jsou to znaky, které se v obou jazycích liší a musel jsem jejich slovenský foném nahradit co nejpodobnějším českým fonémem.

**Tab. 5.1: Mapovací tabulka (část A)**

	Slovenské písmeno	Slovenský foném	Výslovnost použitá v MyVoicu
Samohlásky:	a	a	a
	e	e	e
	i	i	i
	o	o	o
	u	u	u
	ä	{	e
	á	a:	á
	é	e:	é
	í	i:	í
	ó	o:	ó
	ô	u_^o	uo
	ú	u:	ú

V této první části tabulky, kde jsou zobrazeny všechny slovenské souhlásky, jsou pouze dva od češtiny odlišné znaky. Prvním z nich je **ä**, které se při vyslovení nejvíce podobá českému **e**, takže například slovo **spät'** je za pomoci české fonetické sady z MyVoicu zapsané jako **spet'**. Dalším rozdílným znakem je **ô**, to sem namapoval na dva české znaky **uo**, tento znak by mohl být tedy uveden i v části tabulky s dvojhláskami. Příklad slova s tímto znakem je například **môže**, jehož výslovnost se tedy zapíše **muože**.

**Tab. 5.1: Mapovací tabulka (část B)**

	Slovenské písmeno	Slovenský foném	Výslovnost použitá v MyVoicu
Souhlásky:	p	p	p
	b	b	b
	t	t	t
	d	d	d
	ť	c	ť
	ď	J\	ď
	k	k	k
	g	g	g
	c	ts	c
	č	tS	č
	dz	dz	dz
	dž	dZ	dž
	r	r	r
	l	l	l

V předchozí části tabulky jsou všechny znaky shodné s českými.

**Tab. 5.1: Mapovací tabulka (část C)**

	Slovenské písmeno	Slovenský foném	Výslovnost použitá v MyVoicu
Souhlásky:	ř	r=	r
	ř	r=:	r
	l	l=	l
	l	l=:	l
	l	L	l

Slovenština obsahuje několik verzí písmen **r** a **l**, které se ve výslovnosti skoro neliší, proto jsem je všechny namapoval na obyčejné české **r** a **l**, jak můžete vidět v předcházející tabulce. Zde je několik příkladů s těmito rozdílnými znaky: **vřch**, **vřba**, **vřk**, **vřča**, **lřad**.

**Tab. 5.1: Mapovací tabulka (část D)**

	Slovenské písmeno	Slovenský foném	Výslovnost použitá v MyVoicu
Souhlásky:	f	f	f
	v	v	v
	<b>v</b>	<b>U_^</b>	<b>u</b>
	s	s	s
	z	z	z
	š	S	š
	ž	Z	ž
	ch	X	X
	j	j	j
	<b>j</b>	<b>I_^</b>	<b>j</b>
	h	h\	h
	<b>v</b>	<b>w</b>	<b>v</b>
	<b>h</b>	<b>G</b>	<b>h</b>
	m	m	m
	n	n	n
	n	N	n
	ň	J	ň
	f	F	f

V této předposlední části tabulky je stěžejní písmeno **v**, které má ve slovenštině dvě

podoby. První odpovídá klasickému českému **v** a druhé, které se vyslovuje jako **u**, například slovo **Gustáv** se vysloví jako **Gustáu**. V prvním sloupci tabulky si můžete všimnout, že písmena jako **j**, **n** a **v** se zde objevují vícekrát, což je způsobeno tím, že v psané podobě mají písmena shodný znak, ale jejich výslovnost, tedy i fonetický zápis, je jiný, ale ne natolik abych je musel namapovat na různé české znaky .

**Tab. 5.1: Mapovací tabulka (část E)**

	Slovenské písmeno	Slovenský foném	Výslovnost použitá v MyVoicu
Dvojhlásky:	ia	l_^a	ja
	ia	l_^a	ija
	ie	l_^e	je
	ie	l_^e	ije
	iu	l_^U\	ju
	iu	l_^U\	iju

. Největším problémem byly slovenské dvojhlásky **ia**, **ie**, **iu**. Jejich výslovnost se totiž liší podle toho, kde se ve slově vyskytují. Jsou-li někde uvnitř slova jejich výslovnost jsem zvolil takto: **ja**, **je**, **ju**. Takže například slovo **koniec** se vyslovuje jako **konjec**. Naproti tomu jsou-li tyto dvojhlásky na konci slova jejich výslovnost je jiná: **ija**, **ije**, **iju**. Příklad této výslovnosti je třeba ve slově **diktovanie**, které se foneticky zapíše jako **diktovaňije**.

### 5.3 Program na převod české verze MyVoicu na slovenskou

K rychlému převodu české verze systému MyVoice na slovenskou jsem si vytvořil program, který funguje tak, že na vstupu má české „ini“ soubory a na výstupu jejich překlad do slovenštiny. Program dále pracuje s pomocným textovým souborem, v kterém je tabulka se čtyřmi sloupci, v prvním je český tvar slova, v druhém pak jeho česká výslovnost zapsaná pomocí českých fonémů, ve třetím sloupci je slovenský překlad daného slova a ve čtvrtém sloupci je slovenská výslovnost zapsaná pomocí slovenských fonémů. Program prochází vstupní „ini“ soubory a kopíruje jejich obsah, narazily na příkaz, který následuje za klíčovým slovem

„Word“, vyhledá si ho v pomocném souboru, kde nalezne jeho slovenské znění a tím ho nahradí. Na dalším řádku následuje výslovnost příkazu, tu program opět nahradí slovenským tvarem a pak podle mapovací tabulky a předem určených pravidel nahradí všechny znaky slova českými fonémy z fonetické sady použité v české verzi MyVoicu (třetí sloupec v Tab. 5.1). Tím na výstupu dostaneme „ini“ soubory se slovenskými příkazy a správnou výslovností.

K vytvoření tohoto programu jsem použil Perl. Perl je interpretovaný programovací jazyk, který byl vytvořen Larrym Walsem v roce 1987. Tento programovací jazyk slouží hlavně pro práci s texty. K jeho hlavním rysům patří to, že se v něm nemusí deklarovat proměnné ani jejich datový typ, proměnné vznikají při jejich prvním použití a datový typ se určí podle toho, co do proměnných vložíme. Dalším klíčovým znakem jazyka Perl jsou regulární výrazy, které představují velice silný nástroj pro práci s textovými informacemi, a které nám umožňují rychlé hledání či nahrazování částí textů. V neposlední řadě Perl nám umožňuje pracovat s asociativním polem neboli hashem, což je speciální typ pole, v němž nám jako index (neboli klíč) slouží řetězec znaků, což v praxi znamená velké usnadnění vyhledávání. [3]

Tím, že se slovenské „ini“ soubory tvoří programově a ne jen prostým ručním přepisováním, zde dávám prostor pro pozdější změny a úpravy například mého namapování, které by vedly k lepšímu rozpoznávání slovenských mluvčích.

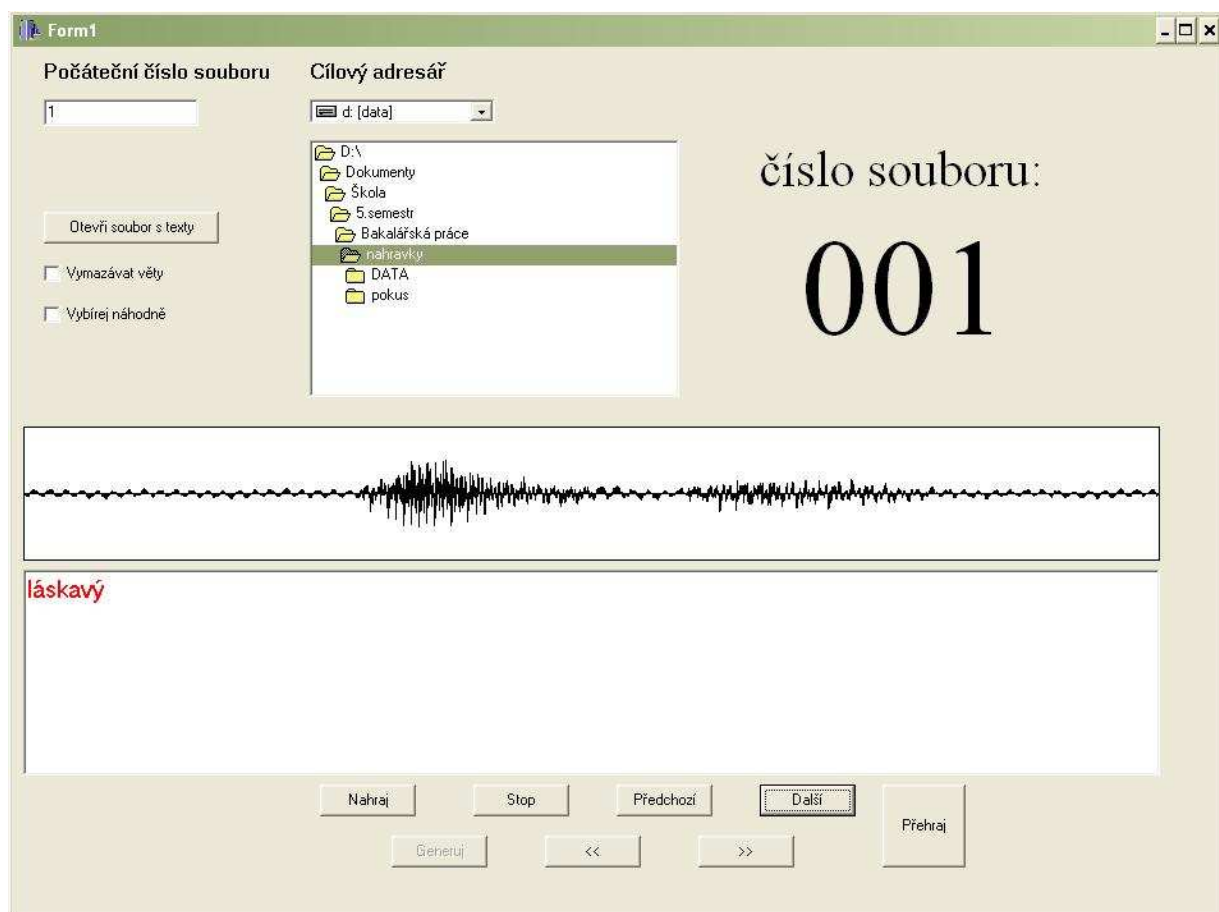
## 6 Testy

Tato kapitola obsahuje popis sbírání dat a popis samotného testování funkčnosti slovenské verze systému MyVoice. Testy jsou rozděleny do dvou částí, v první se testy týkají přesnosti, s kterou MyVoice rozpoznává slovenské příkazy, v druhé části testů se zkoumá adaptace programu na jednoho konkrétního mluvčího. Na oba typy testů jsou použiti dva mluvčí, a to muž a ženu.

### 6.1 Sbíráání dat

Než jsem přikročil k samotným testům, musel jsem shromáždit dostatečné množství zvukových souborů obsahujících slovenská slova namluvená od rodilých slovenských mluvčích. Pro první část testů jsem potřeboval namluvit většinu hlasových povelů z MyVoicu, jejich počet se blíží k 500 položkám. K druhé části textů, čili k adaptaci na jednoho konkrétního mluvčího, bylo nutné získat nahrávky zhruba 400 adaptačních slov.

K nahrávání jsem použil program, který mě dodal vedoucí mé bakalářské práce. Jeho ovládání bylo jednoduché, nejprve je nutné do programu nahrát soubor, který obsahuje slova, jež se mají nahrát. Program tento soubor čte po řádcích a zobrazuje po jednom jejich obsah, proto jsem si vstupní soubor upravil tak, aby na jednom řádku bylo vždy jedno slovo. Dále si musíme zvolit adresář, do kterého chceme ukládat zvukové soubory. Poté již můžeme začít nahrávat. Jak už sem řekl, v rozhraní programu se zobrazuje slovo, které chceme nahrát. Po stisku tlačítka nahraj přečteme slovo a pak stiskneme tlačítko stop, čímž ukončíme nahrávání. Program uloží nahrávku ve formátu wav a textový soubor s nahraným slovem do cílového adresáře. Jména obou souborů jsou shodná a odpovídají vždy číslu pořadí nahrávaného slova. Program dále obsahuje tlačítka pro listování mezi slovy vstupního souboru a tlačítka pro listování mezi již nahranými zvukovými soubory. Ještě zmíním užitečné tlačítko přehraj, po jehož stisku si můžeme okamžitě zkontrolovat kvalitu nahrávky. Jak nahrávací program vypadá si můžete prohlédnout na následující obrázku (Obr. 6.1).



**Obr. 6.1: Nahrávací program**

Tento program plně vyhovoval mým požadavkům a zvukové soubory, které program vytvořil, byly bez dalších úprav použitelné pro testování.

## 6.2 Průběh testů

Vlastní testy probíhaly poměrně jednoduše. V první fázi se zvukové nahrávky slov ze základního slovníku MyVoicu předložili rozpoznávači z MyVoicu a ten se je snažil správně rozpoznat. Výsledkem byl textový soubor, který obsahoval všechna testovaná slova a to, jestli byla rozpoznána správně či špatně. Na konci souboru byla zapsána procentuální úspěšnost rozpoznávání.



V druhé části testů jsem se zabýval adaptací na konkrétního mluvčího. K tomu jsem využil zvukové soubory s adaptačními slovy, ty byly předloženy rozpoznávači a ten se pak adaptoval na mluvčího, který tyto adaptační slova namluvil. Pak se opět rozpoznávači předložily zvukové nahrávky slov z MyVoicu namluvené stejným člověkem. Procentuální přesnost rozpoznávání se pak zlepšila.

### 6.3 Výsledky testů

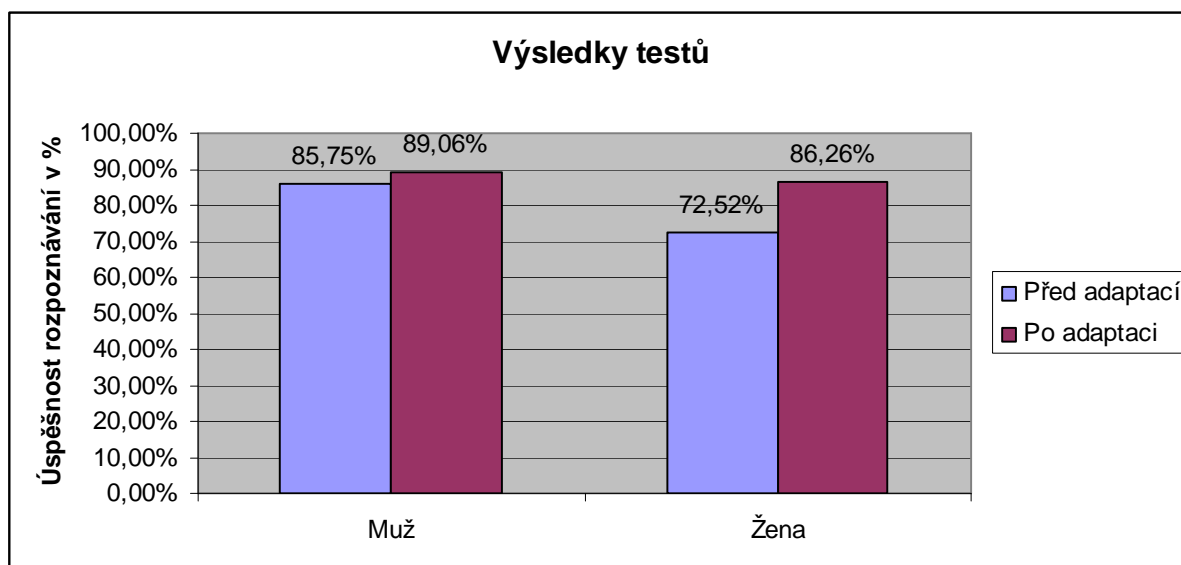
Výsledky testů jsem zpracoval do následující tabulky (Tab. 6.1).

**Tab. 6.1: Výsledky testů**

	<b>Muž</b>	<b>Žena</b>
<b>Před adaptací</b>	85.75 %	72.52 %
<b>Po adaptaci</b>	89.06 %	86.26 %

Jak je vidět z tabulky úspěšnost rozpoznávání před adaptací dopadla výrazně lépe pro muže. O to větší pak byl pokrok v kvalitě rozpoznávání po adaptaci pro ženu. Testy dopadli samozřejmě hůře než když byl MyVoice testován ve své české verzi, ale to je způsobeno tím, že byl MyVoice primárně vyvíjen pro české mluvčí. I tak si myslím, že jsou testy uspokojivé a slovenská verze systému MyVoice je plně funkční a může tak sloužit handicapovaným Slovákům.

Pro přehlednost jsem ještě výsledky zpracoval do následujícího grafu (Graf 6.1).



Graf 6.1: Výsledky testů

## 7 Závěr

Cílem této práce bylo vytvoření adaptace systému MyVoice na slovenský jazyk. V první fázi jsem přeložil všechny hlasové povely MyVoicu do slovenštiny, pak jsem vymyslel vhodné namapování a pravidla pro převod slovenských specifických znaků na sadu českých fonémů, s kterou MyVoice pracuje ve své české verzi. V konečné fázi jsem za pomoci dvou rodilých slovenských mluvčích nashromáždil dostatečné množství zvukových nahrávek, které jsem posléze použil k testování funkčnosti a přesnosti rozpoznávání slovenské verze systému MyVoice.

Z dosažených výsledků je vidět, že rozpoznávání není tak přesné jako u české verze programu, kde se tato přesnost pohybuje kolem 98%. Toto je způsobeno hlavně tím, že testy byly prováděny na českém zvukovém modelu. Cestou k dosažení lepší přesnosti rozpoznávání je určitě adaptování systému na konkrétního mluvčího, kde se procentuální přesnost rozpoznávání zlepšuje.

Hlavním přínosem této práce je nejen vytvoření správně fungující slovenské verze systému MyVoice, která je schopna pomáhat handicapovaným Slovákům ovládat počítač, ale hlavně provedení a otestování mezijazykové adaptace ze zdrojového na cílový jazyk. Mezijazyková adaptace se jeví jako schopný nástroj k přetvoření programů podobných MyVoicu tak, aby je mohli používat i lidé mluvící jinými jazyky.

Samozřejmě mě tato práce pomohla i k vlastnímu sebevzdělání, v jejím průběhu jsem se seznámil s termíny fonetické abecedy a fonetická transkripce. Dále jsem se zabýval společnými a rozdílnými znaky češtiny a slovenštiny. Samozřejmě jsem se musel seznámit s ovládáním systému MyVoice. Díky této práci jsem si také musel osvojit základy programovacího jazyku Perl. V dnešní době, kdy už mladá generace čím dál méně rozumí slovenštině, jsem rád, že mě tato práce pomohla aspoň k částečnému zlepšení vnímání slovenštiny.

Tato práce však obsahuje prostor i pro její rozšíření. Například by se dalo vyzkoušet i jiné namapování slovenských specifických symbolů, což by mohlo vést ke změně přesnosti rozpoznávání. Dále by bylo možné nasbírat větší množství slovenských zvukových nahrávek a

předělat ryze český akustický zvukový model na smíšený, což by výrazně zlepšilo rozpoznávání. Systém MyVoice obsahuje rozsáhlý slovník čítající asi 10 tisíc slov, který slouží k diktování textu, z časových důvodů jsem tento slovník do slovenštiny nepřekládal. V neposlední řadě by bylo potřeba přeložit i interface programu, aby byla tato verze programu kompletně slovenská. Prostoru pro pokračování v mé práci je dost.

## Seznam použité literatury

- [1] BUZÁSSYOVÁ, Klára – HORÁK, Gejza. *Česko-slovenský slovník*, první vydání, Bratislava, 1979.
- [2] PAULINY, Pauliny. *Slovenská gramatika: Opis jazykového systému*, první vydání, Bratislava, 1981.
- [3] SATRAPA, Pavel. *Perl pro zelenáče*, Praha, Neokortex, 2000, 224 s., ISBN 80–6330–02–8.
- [4] NOUZA, Jan. *Počítačové zpracování řeči: Cíle, problémy, metody a aplikace*, Liberec, 2001, ISBN 80–7083–551–6.
- [5] Nouza, Jan – Silovský, Jan – Červa, Petr. *Czech-to-Slovak adapted broadcast news transcription system*, Liberec, 2008.
- [6] Callejas, Zoraida – Nouza, Jan – Červa, Petr. *Cost-efficient cross-lingual adaptation of a speech recognition system*, Liberec, 2008.
- [7] JUHÁR, Jozef. *Spracovanie signálov v systémoch automatického rozpoznávanie řeči*, Košice, 1999, Habilitační práce. Katedra elektroniky a multimediálních telekomunikací. Technická univerzita v Košiciach.
- [8] PSUTKA, Josef – MATOUŠEK, Jindřich. *Mluvíme s počítačem česky*, první vydání, Praha, 2006, ISBN 80–200–1309–1.