
TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky, informatiky a mezioborových studií

Studijní program: X2612 – Elektrotechnika a informatika

Studijní obor: 1802T007 – Informační technologie

Expediční kompletace distribuovaného rozpoznávacího systému

Expeditionary competition of distributed recognition system

Diplomová práce

Autor:	Bc. Tomáš Klucho
Vedoucí práce:	Ing. Miroslav Holada, Ph.D.
Konzultant:	Ing. Josef Chaloupka, Ph.D.

Prohlášení

Byl(a) jsem seznámen(a) s tím, že na mou diplomovou 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 diplomové práce a prohlašuji, že **s o u h l a s í m** s případným užitím mé diplomové práce (prodej, zapůjčení apod.).

Jsem si vědom(a) toho, že užít své diplomové 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).

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

Datum:

Podpis:

Děkuji svému vedoucímu diplomové práce Ing. Miroslavu Holadovi, Ph.D. za užitečné konzultace, rady, připomínky k této práci a také za jeho trpělivost a především pak za poskytnutí již hotového programu k rozpoznávacímu systému Dundis.

Abstrakt

Tato práce se zabývá expediční kompletací distribuovaného hlasového rozpoznávacího systému. Jedním z cílů této práce bylo vytvoření komplexního klientského zařízení. Následně vytvoření jednoduchého a pohodlného způsobu instalace klienta na počítač. Dalším cílem bylo zaměřením se na distribuovaný systém. Bylo potřeba vytvořit aplikaci, která zavede služby rozpoznávacího systému do operačního systému libovolného počítače, které umožní zpracování řeči. Dále bylo nutno navrhnout způsob instalace nastavení rozpoznávače. Dalším cílem bylo navrhnout a naprogramovat webové prostředí, kde budou poskytnuty potřebné části pro rozpoznávání řeči. Posledním cílem bylo provést sérii instalací a krátkých zkušebních testů na různých počítačích a různých operačních systémech.

Pro účely kvalitního rozpoznávání řeči byl použit distribuovaný rozpoznávací systém, který pro aplikaci umožnil využít Ústav Informačních technologií a elektroniky, Technické univerzity v Liberci. Rozpoznávací systém je využit k rozpoznávání izolovaných slov či krátkých frází.

Práce je rozdělena tematicky do dvou částí. V první je uveden teoretický přehled problematiky hlasových systémů, informace o modulech sloužící ke složení dialogového systému a informace o zpracování řeči v jednotlivých zařízeních (klientské, serverové a slavové) při rozpoznávání řeči pomocí distribuovaného systému.

Druhá část je zaměřena na vlastní návrh a realizaci kompletace distribuovaného hlasového rozpoznávacího systému. Vytvořené aplikace byly otestovány řadou testů na různých konfiguracích počítačů, kde byla ověřena funkčnost a také možnost využití na jednotlivých operačních systémech.

Klíčová slova: distribuovaný hlasový rozpoznávací systém, klientská a serverová část distribuovaného systému, rozpoznávání řeči, rozpoznávání izolovaných slov a krátkých frází, dialogový systém

Abstract

This work deals with expeditionary completion of distributed speech recognition system. One of the objectives of the work was to create a complex client system with a simple and convenient installation procedure on client computer. Another aim was to focus on the distributed system. It was necessary to create an application, which loads services of the recognition system into any computer's operation system. It was also necessary to propose how to install the recognizer settings. Another objective was to design a web environment, which will be provided necessary components of the speech recognition. The last goal of the work was to perform series of installations and short experimental test on various computers and operation systems.

For purposes high-quality speech recognition it was use complete recognition system, which for application enabled use Institute of Information technology and electronics, Technical university of Liberec. The recognition system is used to recognize isolated words or short phrase.

The work is divided thematically into two parts. The covers a theoretical survey of the problem of voice system, information on the modules used for composition of the dialog system and information about speech recognition by various devices (client, server, slave) in the distributed system.

The second part is focused on the design and implementation of the completion of a distributed speech recognition system. Developed applications have been tested in a series of tests on various configurations of computers, where the functionality and usage on different operating systems was verified.

Key words: distributed speech recognition system, client and server part of distributed system, speech recognition, recognize isolated words and short phrases, dialog system

Obsah

Prohlášení	3
Poděkování	4
Abstrakt (český)	5
Abstrakt (anglický)	6
Úvod	9
1. Úvod do problematiky hlasových systémů	11
1.1 Míra vlivu řeči při rozpoznávání slov	12
2. Složení hlasového dialogového systému	14
2.1 Rozpoznávání řeči	14
2.1.1 Princip rozpoznávání řeči	16
2.1.2 Detekce začátku a konce slova	17
2.1.3 Parametrizace	18
2.1.4 Porovnání s modely	19
2.1.4.1 Princip porovnávání se vzory	19
2.1.4.2 Statistická metoda	19
2.2 Generování odezvy	20
3. Distribuovaný rozpoznávací systém	21
3.1 Výchozí rozpoznávací systém	21
3.2 Realizace DSR Dundis	22
4. Klientské zařízení	24
4.1 Volba rozpoznávacího systému	24
4.2 Výchozí nastavení klientského zařízení	26
4.3 Realizace klientského zařízení	27
4.3.1 Úvodní nastavení	28
4.3.2 Odolnost proti výpadkům spojení	31
4.3.3 Hlavní panel Komunikátor Sluha	31
4.3.4 Ovládací ikona – nahrávací ikona	33
4.3.5 Konfigurace připojení klienta a rozpoznávacího systému	34
4.3.6 Editor slovníku	35
4.3.7 Výběr prohlížeče internetových stránek	38
4.3.8 Přehled používaných slovníků	39
4.4 Instalace klientského zařízení	39
5. Konfigurace rozpoznávacího systému	43
5.1 Master - server	43
5.2 Slave	45
5.3 Instalace	46
6. Webové prostředí	48
7. Testování klientského zařízení	50
7.1 Vyhodnocení výsledků	50
7.2 Podporované verze operačního systému	51
7.3 Shrnutí	51
Závěr	53

Seznam použité literatury	54
Příloha A – manuál aplikace klientského zařízení.....	55
Příloha B – obsah přiloženého DVD	66

Úvod

Problematika hlasového zpracování řeči si získává stále větší pozornost a stále více se dostává do podvědomí lidí. V současné době je snahou vytvořit takový hlasový dialogový systém, který by bylo možné použít ve více odvětvích současně, například v oblasti soudnictví, lékařství a školství. Jedná se o velmi těžkou úlohu, zejména pak v českém jazyce, který nabízí několik milionů různých slovních tvarů, což zvyšuje nároky na slovník. V těchto oblastech je kromě obecně používaných slov i spousta termínů, což by vedlo k nárůstu počtu slov ve slovníku. V současnosti se vývoj hlasového zpracování řeči a i související komerční aplikace na českém trhu posunul a to zejména díky nárůstu výkonu výpočetní techniky, který umožní v rozumném čase zpracovat větší množství informací.

Dalším z důvodů výzkumu hlasového dialogového systému je i fakt, že je veřejně prospěšný pro uživatele, kteří jsou zdravotně handicapovaní s pohybovými či zrakovými obtížemi, případně pro ty, kteří jsou plně zaměstnaní například v lékařství (patolog). Systémy hlasové syntézy mohou zprostředkovat zrakově postiženým občanům přístup k informacím, které by jim byly jinak nedostupné. Systém rozpoznávání řeči může být zase využit pohybově postiženými lidmi k hlasem řízenému ovládání světel, televize, rádia apod.

Tato práce se zabývá problematikou hlasových systémů k realizaci expediční kompletace distribuovaného hlasového systému. Práce se soustředí na distribuované rozpoznávání izolovaných slov a krátkých frází a z části se opírá o poznatky a závěry autorovi bakalářské práce [1], kde se zabýval návrhem interaktivního hlasového rozhraní mezi člověkem a počítačem. Pro účely kvalitního rozpoznávání řeči byl použit již hotový rozpoznávací systém, který pro aplikaci umožnil využívat Ústav Informačních technologií a elektroniky, Technické univerzity v Liberci. Cílovou platformu tvoří operační systémy Microsoft Windows 2000 a vyšší. Zdrojové kódy jsou psány v prostředí Microsoft Visual Studio 2008 v jazyce C++.

Cílem práce bylo, aby výsledný systém vyhovoval následujícím požadavkům:

1. jednoduchý český instalátor klientského zařízení
2. možnost konfigurace klientského zařízení
3. intuitivní ovládání klientského zařízení

4. pohodlná tvorba slovníku se správnou strukturou pro rozpoznávací systém
5. odolnost klientského zařízení vůči výpadkům sítě, při obnovení přerušeného spojení se celý systém automaticky uvede do provozu sám
6. jednoduchý český instalátor nastavení rozpoznávacího systému na vlastním počítači
7. správné zavádění služeb rozpoznávacího systému do operačního systému
8. možnost konfigurace rozpoznávacího systému
9. správné propojení klientského zařízení a rozpoznávacího systému ať už pomocí webové adresy s rozpoznávacím systémem nebo s vlastním rozpoznávacím systémem ve vlastní lokální síti

Výsledek práce tak uživatelům nabídne možnost využívat aplikaci k ovládání chodu počítače, případně k vyhledávání informací z databáze. Veškeré programy pro hlasové ovládání počítače je možné získat z vytvořeného webového prostředí.

1 Úvod do problematiky hlasových systémů

Hlasové dialogové systémy už podle názvu využívají lidský hlas. Pomocí něho je možné předávat informaci do rozpoznávacího systému, který informaci zpracuje a následně odpoví podle rozpoznatého požadavku. Na tuto problematiku je potřeba se dívat ze dvou pohledů. Prvním je tvorba a správná realizace jednotlivých částí hlasového dialogového systému. Pro úspěšné vedení komunikace člověka s hlasovým dialogovým systémem je potřeba mít dobře navržené a kvalitní moduly **rozpoznávání řeči, porozumění mluvenému jazyku, syntézy řeči** a modul zajišťující **řízení dialogu**. Tyto části zpracovávají informaci (zvuk) tak, aby došlo ke správnému rozpoznání. Druhým pohledem je tvorba rozhraní mezi počítačem a uživatelem, který využívá vlastní hlas k předávání informací hlasovému systému.

V dnešní době je celá řada kvalitních systémů hlasového dialogu, které rozpoznávají s velkou přesností. Dají se dělit od jednodušších, kde se využívá malého počtu slov a jsou vyslovována izolovaně až po systémy, které reagují na souvislou řeč. Problematika hlasového zpracování řeči si získává stále větší pozornost a bylo již dosaženo obrovského posunu i v závislosti na vývoji výpočetní techniky, která umožní v rozumném čase zpracovat větší množství informací. Jedná se o velmi těžkou úlohu, zejména pak v českém jazyce. Český jazyk je specifický, jelikož nabízí několik milionů různých slovních tvarů a i přes pokročilé komponenty výpočetní techniky není možné vytvořit takový hlasový systém, který by v reálném čase zpracovával takovéto množství záznamů. To vede k tomu, že je složité vytvořit hlasový dialogový systém, který by se dal využít ve více oblastech současně například v oblasti lékařství, informační technologii a soudnictví.

Systémy je možné použít např. v oblasti vyhledávání v databázích, vyhledání informace z obsáhlejšího telefonního seznamu nebo vyhledání autobusového, vlakového či leteckého spoje. Dále je možné systémy využít pro menší bankovní operace, hlasové ovládání nebo v oblasti, kde je člověk plně zaměstnán například v lékařství při operacích. V neposlední řadě je možné využít systémy pro zdravotně handicapované s pohybovými či zrakovými obtížemi. Systémy mohou zprostředkovat zrakově postiženým občanům přístup k informacím, které by jim byly jinak nedostupné. Systém rozpoznávání řeči může být zase využit pohybově postiženými lidmi k ovládání světel, televize, rádia apod.

1.1 Míra vlivu řeči při rozpoznávání slov

Při zpracování řeči, jako vstupu do dialogového systému, je potřeba uvážit důsledek využití řeči jako komunikačního média. Řeč jako prostředek přenosu informace má své klady i zápory.

Existuje mnoho oblastí, kde je vhodné využít komunikaci s počítačem pomocí řeči. Jde hlavně o situace, kdy:

- Oči i ruce uživatele jsou plně zaměstnány jinými úkoly.
- Uživatel potřebuje vyhledat informaci z obsáhlé databáze
- Uživatel je vzdálen od systému a může využít pouze hlasové komunikace přes běžné telefonní nebo rádiové spoje
- Uživatel je zdravotně handicapovaná osoba s pohybovými, popřípadě zrakovými obtížemi.

Na druhé straně mezi nevýhody patří zejména podoba řeči v odpovědích systému. V případě, že je odpověď v textové podobě na monitoru a uživatel si ji nepřeje vidět, může ji jednoduše ignorovat. Ovšem v případě, že je odpověď v podobě zvuku, je obtížné jednoduše tuto skutečnost ignorovat. Z tohoto důvodu není možné využít zvukové odpovědi například v pracovním prostředí, protože by mohlo informovat ostatní pracovníky o důležitých informacích nebo je rušit. Možnost odstranění tohoto problému je využitím sluchátek s mikrofonom. Z pohledu rychlosti odpovědi zajišťuje řečový výstup menší rychlost přenosu informace, než je rychlost přenosu vizuální informace.

Pár důležitých rad a zásad, které je vhodné akceptovat při návrhu hlasového dialogového systému, aby se dosáhlo aplikace, která bude pozitivně akceptována uživateli a stala se prospěšnou:

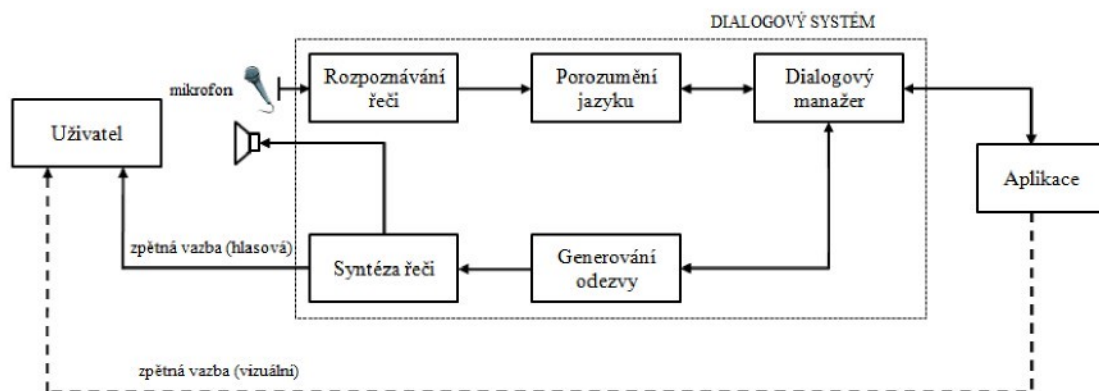
- Systém by měl být uživatelský příjemný, přehledný a robustní ve způsobu interakce s uživatelem.
- Užití systému by mělo být přirozené a aktuálně prospěšné pro uživatele.
- Ideální systém by měl dovolit uživateli zeptat se na to, co v dané situaci považuje za vhodné.
- Funkce rozpoznávání řeči musí být spolehlivá, tj. přesnost rozpoznání slov by měla být co nejvyšší, aby se dialog neprotahoval zbytečným opakováním a ověřováním nespolehlivě rozpoznávaných promluv. Spolehlivý a rychlý přístup

k informacím či rychlé a bezpečné vyřešení dané úlohy má pozitivní vliv na uživatele.

Výsledkem uplatnění lidského faktoru při návrhu hlasového dialogu by měla být spokojenost potenciálního uživatele s pohodlím a snadností obsluhy. Lidský faktor vstupuje do úlohy prostřednictvím rozumného návrhu na řízení dialogu, volbou přirozeného slovníku, využitím nápověd, časováním, užitím zpětné vazby k uživateli apod.

2 Složení hlasového dialogového systému

Na obrázku 2.1 je znázorněno blokové schéma pro tvorbu hlasového dialogového systému. Schéma se skládá z jednotlivých modulů – **rozpoznávání řeči**, **porozumění jazyku**, **generování odezvy** a také **syntézy řeči**. Hlavním a nejdůležitějším modulem je **dialogový manažer**, který řídí interaktivní strategii a také je zprostředkovatelem různých informací a znalostí získaných jinými moduly dialogového systému tak, aby byla úspěšně dokončena úloha, která je řešená dialogem. V této práci se využívaly pouze moduly – rozpoznávání řeči, generování odezvy a dialogový manažer, jelikož při rozpoznávání izolovaných slov není potřeba využívat všechny moduly. Více informací o jednotlivých modulech hlasového dialogového systému lze nalézt ve [2].



Obr. 2.1: Blokové schéma hlasového dialogového systému

2.1 Rozpoznávání řeči

Počítačové rozpoznávání mluvené řeči je předmětem zájmu výzkumných laboratoří již velkou řadu let. Velký vliv na obtížnost rozpoznávání řeči má například typ řečníka, prostředí, ve kterém je promluva řečníka rozpoznávána, vyřčení a způsob promluvy.

Stručnější popis jednotlivých důvodů, které mají vliv na rozpoznání řeči:

- Hlas jakékoliv osoby je odlišný od jiné, jelikož každý uživatel má odlišné hlasové ústrojí a také artikulace je u všech uživatelů různá. Z toho plyne, že každý má jinou barvu hlasu, jiný přízvuk, odlišné tempo řeči atd. Systémy rozpoznávání řeči proto můžeme dělit na **systémy závislé na řečnickovi** a na **systémy nezávislé na řečnickovi**.

- Navíc hlas stejné osoby se také může měnit v různých okamžicích. Promluva se mění, pokud uživatel vysloví dotaz normálním hlasem, potichu, nahlas atd. Proto je nemožné, aby stejný řečník sdělil promluvu stejného slova stejným způsobem v různých situacích. Toto zapříčiňuje jev **prozodie řeči** – ovlivňuje jiné časování vyslovení dotazu v různých situacích. Také změna frekvence hlasu je jiná.
- V souvislé řeči se projevuje jev **koartikulace**, který může pozměnit fonetické vlastnosti začátku a konce slova v závislosti na kontextu okolních slov.
- Dalším parametrem je různorodost prostředí. V promluvě uživatele se projevuje vliv okolního prostředí.
- Projevuje se i nekvalita a rušení přenosového zařízení, které snižuje kvalitu přenášeného výrazu.
- Vliv na správnou funkci rozpoznávání řeči má i způsob vyřčení promluvy. U **rozpoznávání izolovaných slov** (jednoslovných výrazu či krátkých frází) z relativně malého slovníku bude rozpoznávání lehčí, než u **rozpoznávání diskrétního diktátu** (slova jsou vyslovována izolovaně s krátkou mezislovní pauzou). Nejtěžší úlohou je **rozpoznávání souvislé řeči**, kdy má slovník desetitisíce až statisíce slov.
- Dalším důležitým vlivem na správnou funkci rozpoznávání řeči je způsob promluvy. Zda jde o řeč **čtenou** nebo o **spontánně pronášenou promluvu**. U čtené řeči by mělo dojít k vyšší pravděpodobnosti správného rozpoznávání. Gramatická stavba vět by měla vyhovovat spisovné češtině. Neměly by se projevit tzv. **neřečové události** (hlasité váhání, nádechy apod.) a promluva by měla být plynulá. Naopak u spontánně pronášené promluvy řečník velmi často vkládá neřečové události, opakuje se, vysloví část slova či věty a poté se snaží o jinou formulaci. Navíc dochází k vyslovování velkého množství hovorových slov a nespisovných gramatických vazeb.

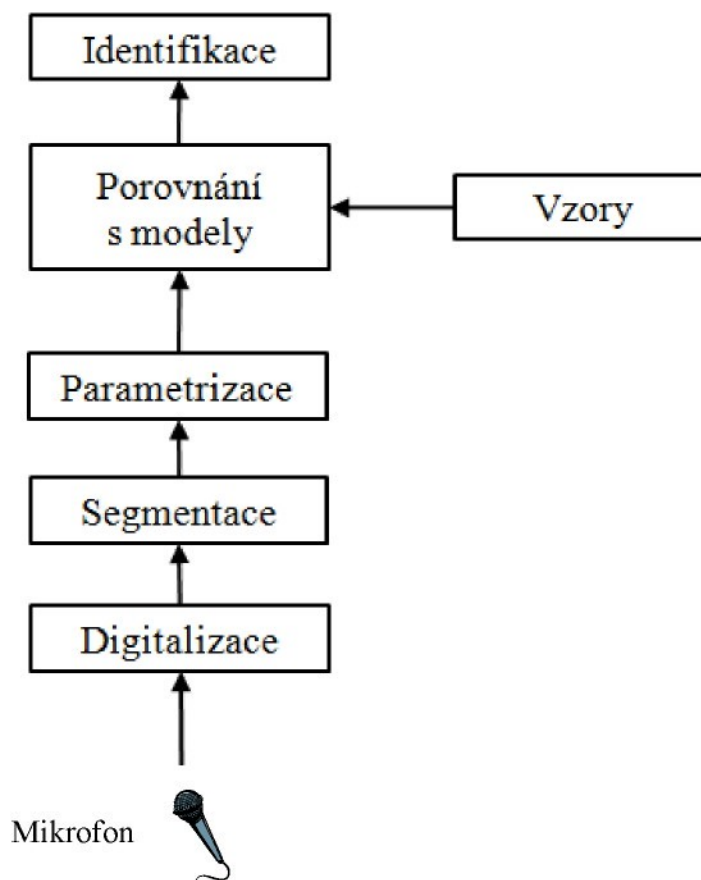
Pro potlačení těchto nežádoucích vlivů při rozpoznávání řeči lze provést úkony pro zvýšení robustnosti:

- Při návrhu rozpoznávání řeči, kde se očekává negativní vliv v podobě hluku na pozadí nebo nekvalita a rušení přenosového zařízení, lze navrhnout způsob na potlačení nebo zmírnění vstupního šumu či vlivu zkreslení přenosového kanálu.

- Akustický model je potřeba trénovat z takových dat, která budou odpovídat prostředí, pro které je aplikace připravována.
- Krátké „nedořeky“ a nové zahajování komunikace lze zachytit modelem neřečových událostí.
- Využití vhodného jazykového modelu při využití rozpoznávání řeči, kde se očekává dialog veden spontánní řečí.

2.1.1 Princip rozpoznávání řeči

Část této práce se zabývá problematikou rozpoznávání izolovaných slov a krátkých frází. Na obrázku 2.2 je znázorněno blokové schéma takového rozpoznávání. Schéma se skládá z několika částí, kterými daná promluva v rozpoznávacím systému prochází. Těmito částmi jsou **digitalizace**, **segmentace**, **parametrizace**, **porovnání s modely** a **identifikace**.



Obr. 2.2: Blokové schéma rozpoznávání izolovaných slov

Vstupem do hlasového dialogového systému potažmo do modulu zpracování řeči je **promluva uživatele** v podobě analogového systému. Tato informace je snímána pomocí **mikrofonu**. Analogový signál se moduluje do číslicové podoby (**digitalizace**) automaticky pomocí zvukové karty. Některé zvukové karty generují nuly na začátku a konci vzorkovacího procesu a to může způsobit problémy při výpočtu některých parametrů, které používají operace dělení či logaritmování. Z tohoto důvodu je vhodné přičíst malý šum v podobě hodnoty -1 nebo +1. Některé zvukové karty mají nenulovou stejnoměrnou složku. Tento nežádoucí vliv se dá odstranit tzv. **preemfázovým filtrem**. Následně se číslicový signál rozdělí na úseky o stejné délce (**segmentace**). Toto dělení se dělá pravidelně a délka úseku by měla být menší než trvání nejkratších fonémů (obvykle 10 až 25ms). Úseky se většinou definují tak, aby se částečně překrývaly. Tím je docíleno hladkého průběhu parametrů (**příznaků**) počítaných v jednotlivých úsecích. Zpracovávaný signál obsahuje i neřečové úseky v podobě ticha. Pomocí **detektoru začátku a konce** dojde k odstranění nepotřebné informace signálu. V tuto chvíli signál obsahuje pouze důležité informace pro rozpoznávání. V jednotlivých úsecích dojde k vypočtení parametrů (**parametrizace**), které vytvoří **příznakový vektor** – soubor příznaků, který se určuje v každém úseku. Reprezentace neznámého slova pomocí sekvence příkazových vektorů je porovnávána se všemi modely jednotlivých slov (**porovnání s modely**) a slovo, které má největší shodu se stává rozpoznávaným slovem (**identifikace**).

2.1.2 Detekce začátku a konce slova

Ve snímaném signálu se objevují neřečové úseky v podobě ticha na začátku a na konci promluvy. Úkolem detektorů, při rozpoznávání izolovaných slov či krátkých frází, je odstranění nepotřebné informace signálu a ponechání úseků s nosnou informací.

Předpokládá se, že energie signálu v pozadí v neřečovém úseku (ticho) je menší, než energie v řečových úsecích. Tudíž nejjednodušší detektory začátku a konce slova pracují na principu sledování energie signálu a detekci vzestupu energie na začátku slova a poklesu energie na jeho konci.

Složitější detektory začátku a konce slova zkoumají nejen energii signálu, ale i jeho obsah s využitím kepsrální analýzy, kdy hledají charakteristické rysy lidské promluvy. Výhodou těchto detektorů je rozlišení hlasitého šumu na pozadí od užitečné informace v podobě řeči.

2.1.3 Parametrizace

Tento proces se provádí na úsecích nesoucích informace rozpoznávané promluvy. Jednotlivé parametry se počítají v každém úseku a jsou uloženy v příznakovém vektoru uspořádaných n -tic čísel, které reprezentují obsah signálu z pohledu lidské řeči.

Pro zachycení důležitých a podstatných informací ze signálu stačí frekvenční rozsah 8kHz a tudíž vzorkovací frekvence 16kHz a 16-bitová hloubka, aby nedocházelo ke zkreslení tichých nahrávek.

Reprezentace řečových signálů se neustále vyvíjí. Vývoj začínal od reprezentace signálu pomocí parametrů energie a počtu průchodů nulou, přes spektrum a spektrální příznaky a lineárně prediktivní koeficienty až po v dnešní době nejvyužívanější kepstrum, kepstrální příznaky a delta příznaky. I přestože ve vývoji nejvhodnější reprezentace signálu se vytvořily nové metody jako různé modifikace kepstrálních příznaků, tak se ve své podstatě moc neuchytily a proto se v praxi převážně používají právě kepstrální příznaky.

V současnosti jsou nejvíce používány pro reprezentaci řečových signálů kepstrum, kepstrální příznaky a delta příznaky. Kepstrální analýza je založena na určování krátkodobého spektra signálu a jeho transformaci sloužící k oddělení důležité informace od nepodstatné. Spektrum signálu se měří na úsecích o délce 25ms, což představuje nejkratší dobu trvání fonémů. Posun mezi úseky je 10ms z důvodu hladšího průběhu parametrů a to vše při řeči vzorkované na 16kHz. Dále se pro reprezentaci signálů používají delta příznaky, které zastupují dynamické příznaky a charakterizují změny mezi stejnými příznaky v sousedních úsecích. V moderních systémech se používají i delta-delta příznaky, které se počítají ze změn delta příznaků mezi stejnými příznaky v sousedních úsecích jako při výpočtu delta příznaku.

Pro výpočet kepstra se využívají dvě metody. První metoda získání vektoru příznaků je založena na výpočtu lineárně prediktivních koeficientů (LPC). Algoritmus založený na této metodě je rychlejší na výpočet, jednodušejí implementovatelný, ale v současnosti se již moc nevyužívá. V současné době a v praxi se využívá metoda založená na výpočtu podle definice přes FFT, logaritmus výkonu a IFFT pomocí křivky lidského vnímání frekvencí (využití mel stupnice). Označení této metody je MFCC

(Mel-Frequency Cepstral Coefficients). Tato metoda vykazuje lepší výsledky při rozpoznávání a v dnešní době již není problém s implementací.

Více informací o parametrizaci a reprezentaci řečových signálů lze vyhledat ve [3] a [4].

2.1.4 Porovnání s modely

Nejdůležitější částí rozpoznávacího systému je porovnání s referenčními modely. Porovnává se neznámé slovo s databází referencí (vzorů) a rozhoduje o určení, do jaké třídy slovo patří, a tudíž má neznámé slovo největší pravděpodobnost shody. Jednotlivými třídami jsou přímo konkrétní slova a jsou shromážděna ve slovníku.

Pro porovnávání s modely se využívají algoritmy pro měření podobnosti různě dlouhých řad. Algoritmy pracují na principu **porovnávání se vzory** a na rozpoznávací systémy pracující s využitím **statistických metod**.

2.1.3.1 Princip porovnávání se vzory

Tato metoda se nejvíce vyvíjela v 70. a 80. letech, kdy byla hlavním modulem rozpoznání řeči pro izolovaně vyslovovaná slova. Slovo se zpracovává jako celek a jeho správné vyhodnocení se určuje pomocí nejkratší kumulované vzdálenosti (Mahalanobisova vzdálenost) mezi příznakovými vektory neznámého slova a referenční posloupností příznakových vektorů. Rozhodující je určení správné vzdálenosti, která je určena pomocí metody dynamického programování. Tato metoda pracuje s efektem nelineární časové normalizace, kde je kolísání v časové ose modelováno časově nelineární „bortivou“ funkcí s přesně danými vlastnostmi. Časové rozdíly mezi dvěma řečovými obrazy jsou přitom eliminovány „borcením“ jedné z časových os tak, aby bylo dosaženo maximální shody s druhým obrazem. Tato metoda se nazývá **dynamické borcení času** (DTW – Dynamic Time Warping).

2.1.3.2 Statistická metoda

Tato metoda rozpoznávání řeči je založena na statických metodách, kde jsou slova a celé promluvy modelovány pomocí tzv. **skrytých Markovových modelů** (HMM – Hidden Markov Model). Využívá se modelování pravděpodobnostních stavových automatů. Každý referenční model má jeden automat. Míra podobnosti je

určena pravděpodobností mezi referencí a promluvou. Hledá se nejvyšší pravděpodobnost.

Vytvoření referencí slov ze slovníku je možné provést dvěma způsoby. První způsob je modelování slova jako jednoho celku jedním skrytým Markovovým modelem slova. To ovšem s sebou nese nevýhody, jelikož je snahou mít univerzální rozpoznávací systém nezávislý na mluvčím. Z tohoto důvodu je potřeba mít ke každému slovu velké množství mluvčích. Z toho plyne, že přidání nového slova s sebou nese nutnost natrénování nového slova stejnou skupinou mluvčích. Druhým a používanějším způsobem je modelování slova pomocí skrytých Markovových modelů subslovních jednotek (fonémů, difonů, trifonů). V tomto případě stačí mít vytvořenou databázi referencí těchto subslovních jednotek a vzor libovolného slova lze složit z těchto referencí. Toto umožňuje jednoduché přidání nového slova do slovníku.

Více informací o porovnávání s modelem lze najít ve [3] a [4].

2.2 Generování odezvy

Generování odezvy umožňuje, aby uživateli hlasového dialogového systému byla sdělována správná odpověď na danou promluvu v odpovídající podobě. Hlavním úkolem je určit, jak danou odpověď podat. Odpověď, která se má vytvořit, může být v textové nebo zvukové podobě. V současné době se většinou využívá pro aplikace hlasových dialogových systémů konstrukce promluv založených na sémantických rámcích. V odpovědi jsou nevyplněné části, do kterých se vkládají odpovídající informace, aby odpověď byla celistvá a měla všechny náležitosti, které řečník promluvou zadal. Vhodné je zajistit, aby vložené informace byly ve správném kontextu s okolními slovy.

Když je počet generových odezev konečný, je možné odpovědi předem namluvit a uložit do slovníku jako jeden celek. To zajišťuje správnou gramatickou vazbu mezi okolními slovy. V opačném případě je odpověď složena z předem připravené věty s doplněnými informacemi, aby poskytla řečníkovi odpovídající odpověď. V případě, že je některá z volných pozic nevyplněná nebo špatným způsobem rozpoznána, musí existovat dotaz, aby došlo k doplnění volné nebo chybně rozpoznané promluvy a nové vložení rozpoznané informace do odpovědi.

3 Distribuovaný rozpoznávací systém

Problém v oblasti nedostatečného hardwaru pro řešení úloh lze zmírnit přenesením určité části úlohy na jiné zařízení, které disponuje takovou výpočetní technikou, která je vhodná k vyřešení úlohy. Obecně se takový přístup nazývá distribuované zpracování. V případě, že se jedná o hlasové rozpoznávání, tak je to označováno jako distribuované rozpoznávání řeči (DSR – Distributed Speech Recognition).

Jedná se o systémy používající architekturu klient - server. Informace přijatá na straně klienta je zaslána pomocí datového spojení serveru, který zpracuje odpovídající část úlohy a odešle výsledek nazpět klientovi.

Výhodou této realizace je snížení hardwarových požadavků zařízení na klientské straně, jelikož o podstatně náročnější výpočet se stará právě strana serveru. Naopak nevýhodou v této realizaci je nutnost síťového spojení a prodleva při zasílání informací na server. Ovšem v dnešní době se již předpokládá, že i běžní uživatelé mají přístup k internetu s dostatečným datovým přenosem.

V této práci byl využit distribuovaný rozpoznávací systém pod názvem Dundis (<http://dundis.ite.tul.cz>), který provozuje Technická univerzita v Liberci.

3.1 Výchozí rozpoznávací systém

V této práci byl použit distribuovaný rozpoznávací systém, který pro aplikaci umožnil využívat Ústav Informačních technologií a elektroniky, Technické univerzity v Liberci. Jedná se o kvalitní rozpoznávací systém používající získané informace z mikrofону potažmo zvukové karty počítače.

Základní parametry rozpoznávacího systému:

- vstupní data – vzorkovací frekvence 16kHz a 16bitová hloubka signálu
- úseky pro parametrizaci dlouhé 25ms a posun mezi jednotlivými úseky je 10ms
- detektor začátku a konce slova je založen na sledování vývoje energie signálu
- pro parametrizaci se používají kepsrální koeficienty, které jsou založeny na metodě MFCC, používá se 13 kepsrálních příznaků + delta a delta-delta příznaky

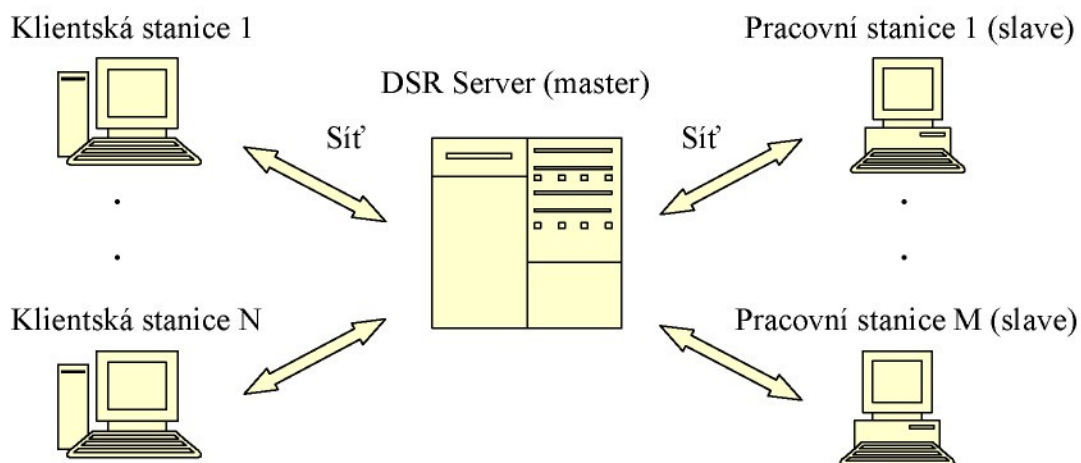
- porovnání s modely je založeno na multimixturových skrytých markovských modelech
- reference vzorů (slov ve slovníku) se skládají řetěžením jednotlivých subslovních jednotek (fonémů)
- slovníky a fonémové modely jsou uloženy v souborech se speciální příponou a jsou načítány při inicializaci rozpoznávače

Jelikož se jedná o distribuovaný rozpoznávací systém, který předává náročnější výpočetní část jinému zařízení, tak na straně klienta nedochází k úplnému vyřešení úlohy. Z důvodu nižších požadavků na propustnost sítě se na straně klienta provede pouze digitalizace a výpočet 13 keprstrálních příznaků a tyto informace se předají jinému zařízení, které dokončí vyhodnocení informace a předá odpověď klientovi.

3.2 Realizace DSR Dundis

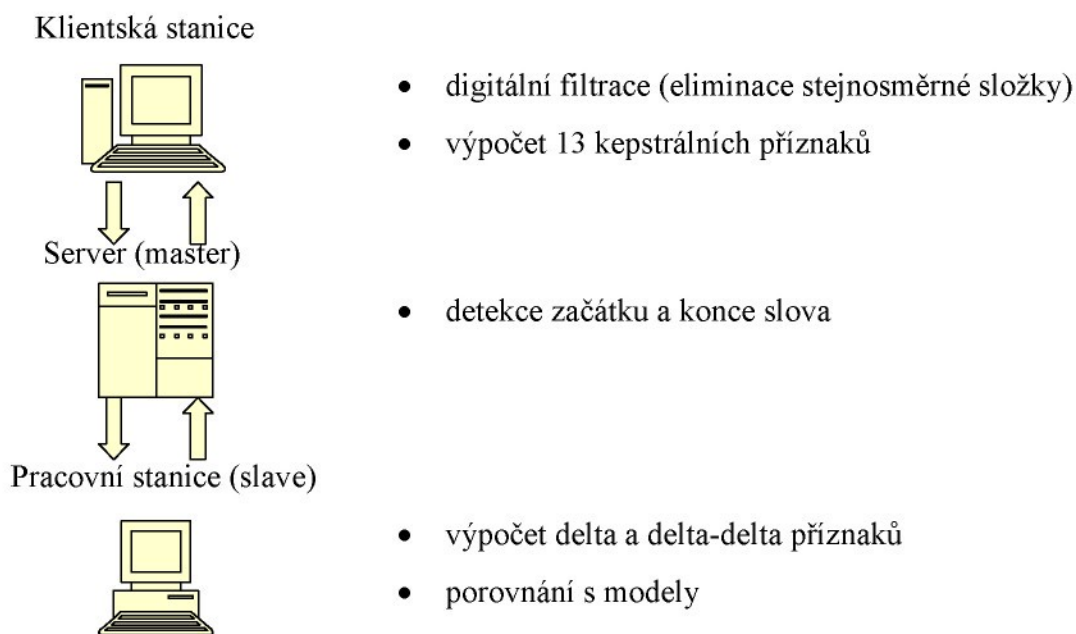
Právě jeden z počítačů bude řešit veškeré řízení. Je označen jako **master** (pán) a ostatním říká, co mají dělat. Zbylé počítače označeny jako **slave** (otrok) se k tomuto hlavnímu počítači připojují a budou vykonávat veškerou práci, kterou jim master zadá.

Komunikace probíhá od uživatele (klienta), který zadá požadavek v podobě promluvy k masteru, který jej pomocí datového přenosu přijme a následně předá slave počítačům. Připojené slave počítače čekají na zadání úlohy ke zpracování masterem, realizují algoritmus řešení a odesílají výsledek zpět masteru. Ten výsledek pošle zpět na klientské zařízení. Vnitřní organizace mezi masterem a slavem je založena na systému fronty úloh (volné a zaměstnané slave počítače). Na obrázku 3.1 je znázorněno blokové schéma DSR systému Dundis.



Obr. 3.1: Blokové schéma DSR systému Dundis

Na obrázku 3.2 je ukázán postup při zpracování řeči a předávání náročnější výpočetní části serveru.



Obr. 3.2: Postup zpracování promluvy v jednotlivých zařízeních

Master a slave jsou realizovány jako autonomní objekty a využívají asynchronní komunikaci po síti s využitím vlastního vlákna. Připojení slave počítače k masteru probíhá následujícím způsobem. Proběhne inicializace pomocí zaslání dat od mastera pro vzájemnou synchronizaci. Poté se slave počítač může dostat do fáze, kdy nepracuje a čeká na přidělení úlohy od mastera nebo pracuje a to zpracovává úlohu získanou od mastera. V objektu mastera jsou připojeni všechny slave počítače. Fronty úloh jsou realizovány pomocí dvojcestného spojového seznamu, což neomezuje délku fronty.

Více informací ohledně využití jednotlivých částí distribuovaného systému lze nalézt v [5].

4 Klientské zařízení

První částí cyklu rozpoznávání řeči je klientské zařízení (klient). Zde probíhá první etapa zpracování řeči v podobě zaznamenání zvuku pomocí mikrofону respektive zvukové karty. Následně dojde k digitalizaci nasnímaného signálu a odstranění nežádoucí stejnosměrné složky. Posledním bodem u klienta je parametrizace signálu pomocí 13 keprstrálních příznaků. Tato oblast ohledně daného rozpoznávání není předmětem zkoumání této diplomové práce.

V této práci byla snaha dokončit navrženou aplikaci, která vytváří interaktivní rozhraní mezi počítačem a uživatelem. V diplomové práci byla využita aplikace navržená a vytvořená autorem bakalářské práce [1], která neobsahovala veškeré očekávané funkce, které by mělo ideální hlasové rozhraní mít. Navíc zde neexistovala možnost uživatelského nastavení a přizpůsobení aplikace, které se v dnešní době očekává u vytvářených aplikací. To vedlo k domněnce, že aplikace není příliš uživatelsky příjemná. Ovladatelnost, přehlednost a snadnost práce s aplikací ovšem byla dostačující, jak je možné vidět na vyhodnocení ovladatelnosti aplikace z grafu 7.1. Tudíž bylo potřeba vytvořit uživatelsky zajímavější aplikaci, kde si uživatel může sám nastavit určité parametry, které se zaznamenají, a při opětovném spuštění dojde k jejich načtení. Také se v ní vyskytly nežádoucí chyby, které vedly ke špatnému chodu aplikace a v některých případech až k úplnému spadnutí.

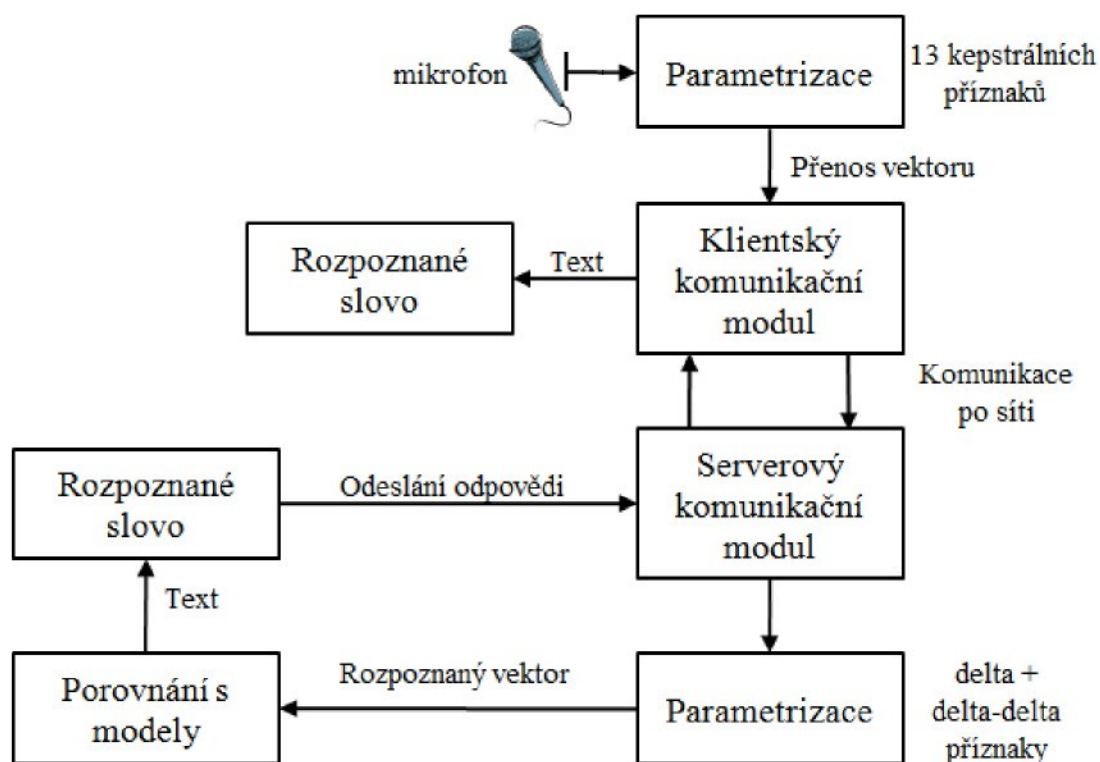
Aplikace vytvořená v bakalářské práci byla realizována v prostředí Microsoft Visual Studio C++ .NET a proto bylo použito stejné prostředí s novou verzí 2008.

4.1 Volba rozpoznávacího systému

Pro účely co nejlepšího a pokud možno bezchybného rozpoznávání řeči byl využit již hotový rozpoznávací systém, který je oblastí zkoumání a vývoje již dlouhou řadu let. Využití systému umožnil Ústav Informačních technologií a elektroniky, Technické univerzity v Liberci. Rozpoznávací systém byl využit k rozpoznávání izolovaných slov nebo krátkých frází. Tento systém je vybaven řadou užitečných funkcí. Jednou z nich je, že dochází k hlídání šumu a hluku. V okamžiku, kdy je zapnuté nahrávání a v pozadí je nějaký šum či hluk, nedochází k náhodnému vygenerování odpovědi, jelikož je sledován vývoj energie v zaznamenaném signálu.

Další výhodou tohoto systému je jeho univerzálnost a možnost využití pro velké množství mluvčích, jelikož reference vzorů zaznamenaných ve slovníku jsou složeny z modelů jednotlivých fonémů. Další funkcí systému je možnost pracovat s dvěma typy slovníků – serverový, uživatelský. Serverový slovník je nahrán při inicializaci rozpoznávacího systému a je na straně serveru. Tento slovník není možné uživatelem měnit a jeho kapacita může být desetitisíce slov. Uživatelský slovník si vytváří uživatel sám pomocí Editoru slovníku, který je součástí této aplikace. Informace k tvorbě slovníku je v bodě 4.3.6 a příloze A. Slovník takto vytvořený má omezenou kapacitu a je možné využít maximálně dvou tisíc slov.

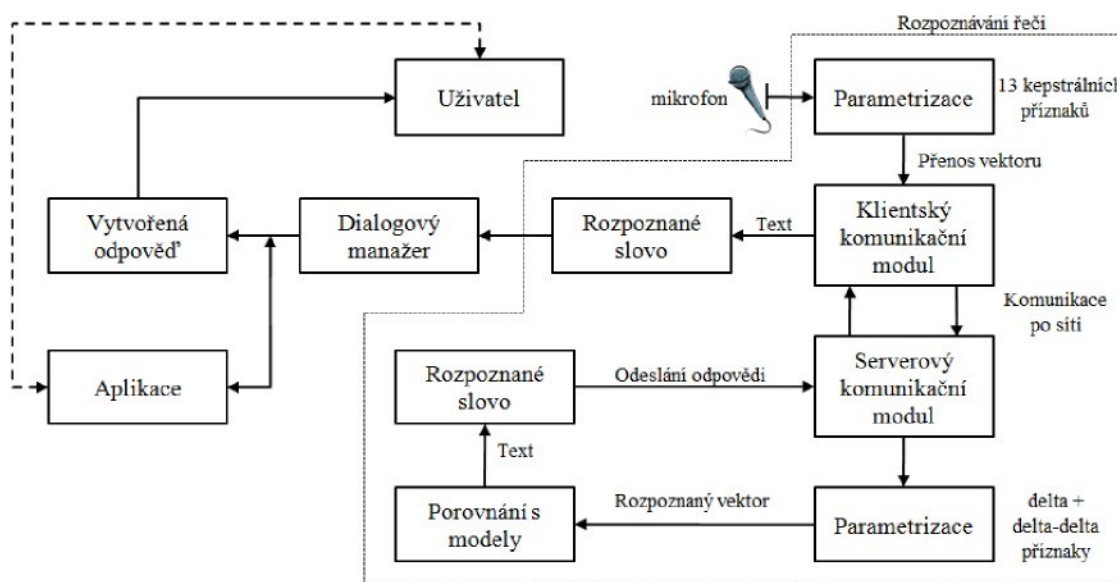
Princip rozpoznání promluvy pomocí distribuovaného hlasového systému je popsán v bodě 3.2. Detailnější přehled struktury distribuovaného rozpoznávacího systému je znázorněn na obrázku 4.1.



Obr. 4.1: Struktura distribuovaného rozpoznávacího systému

Po zpracování a rozpoznání promluvy řečníka je potřeba vhodným způsobem vytvořit odpověď pro uživatele. Existují dvě možnosti – v podobě textové nebo zvukové. V aplikaci byl ponechán navržený způsob odpovědi v podobě zobrazení textu na monitoru počítače. U některých promluv dojde ke spuštění aplikace, internetové stránky či stisku klávesové zkratky.

Kompletní blokové schéma hlasového systému je vidět na obrázku 4.2.



Obr. 4.2: Bloková struktura hlasového systému

4.2 Výchozí nastavení klientského zařízení

Jak již bylo zmíněno, diplomová práce vychází z autorovy bakalářské práce [1], ovšem s využitím pouze funkčních a prospěšných částí zdrojových kódů, které nevykazovaly chyby při běhu programu. Ovladatelnost a přehlednost aplikace byla podle ohlasu uživatelů dobrá a nebylo třeba výrazným způsobem měnit design aplikace.

Přehled jednotlivých uživatelských bloků interaktivního hlasového rozhraní mezi počítačem a uživatelem:

1) hlavní okno Komunikátoru – obsahuje následující funkce nebo možnosti

- zahájit komunikaci nebo konec komunikace
- načtení nového slovníku
- editace slovníku
- přehled slovníků
- výběr prohlížeče internetových stránek

2) ovládací ikona – obsahuje následující funkce nebo možnosti

- možnost spuštění nahrávání rozpoznávacího systému
- posun ikony po ploše, když není zobrazené jiné okno aplikace
- přepnutí do hlavního okna Komunikátoru, editace slovníku nebo nápovědy
- nastavení velikosti nebo průhlednosti

3) editace slovníku – slouží k vytvoření správné struktury slovníku pro distribuovaný rozpoznávací systém

- přidat slovo, přidat internetovou stránku, přidat aplikaci, přidat aplikaci prohlížeč s odpovídajícím souborem
- editace výrazu, úplná změna výrazu
- tvorba slovníku s telefonním seznamem pracovníků TUL
- záměna slov u promluvy

4) přehled slovníků

- hlavní funkcí bylo informovat uživatele o dostupných slovnících

4.3 Realizace klientského zařízení

Využitá aplikace neumožňovala instalaci, což působilo jako nedokončený software a také to vedlo k problémům se spuštěním. Z tohoto důvodu byla vytvořena instalace, která ověří zda je na počítači nainstalovaná požadovaná verze .NET Frameworku, která je potřebná pro chod klienta. Následně dojde k vytvoření adresáře programu, kde jsou soubory aplikace a uživatelského adresáře, ve kterém jsou uloženy slovníky, inicializační soubor a soubor s přehledem slovníků.

Při prvním spuštění proběhne nastavení inicializačního souboru, kde jsou ukládány jednotlivé parametry, které si uživatel v průběhu práce se softwarem mění a tím se aplikace stává více uživatelsky příjemnější. Zároveň je kontrolováno, zda je Komunikátor Sluha spuštěn, aby nedošlo ke spuštění více klientů v jeden moment. Po spuštění aplikace dojde nejprve k otestování, zda již je aplikace spuštěna. Do proměnné existenceProgramu jsou načteny všechny spuštěné procesy na počítači. Následně dochází k ověřování, zda-li je nějaký proces pojmenován Komunikátor Sluha, čímž by nedošlo ke spuštění nového procesu s dalším klientem. Realizace je ukázána na následujícím zdrojovém kódu.

```
array<Process>^ existenceProgramu = Process::GetProcesses();
for (int i = 0; i < existenceProgramu->Length; i++)
{
    if (existenceProgramu[i]->ProcessName == "Komunikátor Sluha")
    {
        goto konec;          //ukončí spouštění nového procesu
    }
}
```

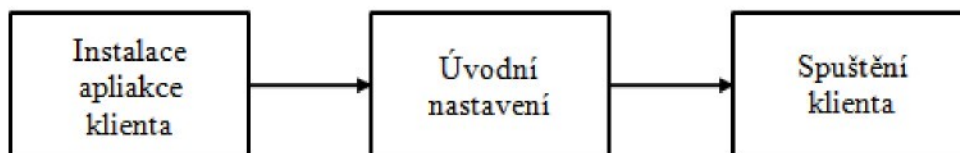
Klient doznal určité grafické úpravy, ale jelikož ovladatelnost a přehlednost byla vnímána pozitivně, bylo provedeno jen minimum změn. I přesto aplikace nabízí možnost náhledu do nápovědy, kde jsou jednotlivé funkce klienta popsány. Nápověda by měla být součástí každého komplexního programu, protože může nastat situace, kdy se uživatel potřebuje v aplikaci více zorientovat. Tím může být dosaženo rychlejšího a úspěšnějšího vyřešení úlohy.

Grafické úpravy spočívaly v podobě vytvoření ikony aplikace a také ikony pro jednotlivá zobrazovaná okna aplikace, což opět software posouvá k pozitivnějšímu vnímání ze strany poskytnutí komplexní aplikace.

Následovaly změny a úpravy na jednotlivých částech aplikace, které vykazovaly chyby, aby došlo k zabránění neočekávaného chodu programu. Z využití aplikace klientského zařízení byla využita převážně myšlenka a vzhled. Také došlo k vytvoření nových žádaných funkcí, které se stanou prospěšně pro klientské zařízení.

Realizace klienta je vytvořena ve dvou variantách. První varianta nenabízí možnost pro nastavování adresy rozpoznávacího systému. Automaticky je nastavena adresa `dundis.ite.tul.cz`. Nutností je připojení k internetu, aby mohlo probíhat rozpoznávání. Tato varianta je převážně určena pro běžné uživatele hlasových systémů. Druhou variantou je aplikace, kde již je možné zadat vlastní adresu rozpoznávacího systému. Je tak umožněno rozpoznávání promluvy i bez využití internetového připojení. V takovém případě je zapotřebí mít zavedeny služby serveru a slavu v operačním systému. Tato varianta je určena již pro zkušenější uživatele.

Přehled postupu tvorby klientského zařízení je zachyceno na obrázku 4.3.



Obr. 4.3: Postup tvorby klientského zařízení

4.3.1 Úvodní nastavení

K dosažení většího pohodlí pro uživatele byly vytvořeny nové funkce nastavení aplikace, které umožňují uživateli přizpůsobení aplikace vlastním požadavkům.

Jednotlivé vlastnosti se následně zaznamenávají do inicializačního souboru *Nastaveni.ini*, který slouží k předávání informací při opětovném spuštění.

Úvodní nastavení a inicializace byla tvořena stejně jako celý program v prostředí Microsoft Visual Studio C++ .NET verze 2008. K vytváření inicializačního souboru nebyly využity žádné podpůrné moduly a vše bylo čistě naprogramováno. Výhodou takovéto strategie je možnost jednoduchého přidávání nových parametrů, kontroly stávajících a v případě nutnosti i pohodlné ladění kódu.

Potřebné parametry inicializačního souboru:

- **adresář programu** – slouží k předávání informace o cestě, která se nastaví při otevírání dialogových oken k vyhledání libovolného spustitelného souboru
- **cesta hlavního adresáře slovníku** – slouží k předání informace, kde má uživatel aplikace uloženy slovníky, které využívá jako prioritní; také slouží k předávání informace o nastavení cesty pro dialogová okna při ukládání nebo otevírání slovníků v Editoru slovníku; dále se z tohoto adresáře v úvodu nastavují další parametry jako cesta zaváděcího slovníku, cesta slovníku s WWW stránkami a cesta k souboru s telefonním seznamem
- **cesta k souboru s přehledem slovníků** – tato informace slouží pro Přehled slovníků, kde jsou zobrazeny všechny používané či použité slovníky; v souboru jsou uloženy informace ke všem slovníkům, které obsahuje adresář hlavního slovníku, další adresáře slovníků a také slovníky, které si uživatel uloží mimo tyto definované adresáře
- **cesta zaváděcího slovníku** – slovník je určen k nahrání do distribuovaného rozpoznávacího systému po zapnutí aplikace, aby uživatel nemusel po spuštění nejprve slovník vyhledat a nahrát a až poté by mohl využít ovládací ikony ke komunikaci s počítačem
- **cesta slovníku s WWW stránkami** – jedná se o soubor sloužící pro rychlé ukládání internetových stránek, které je umožněno speciální promluvou „přidat stránku“
- **cesta k souboru s telefonním seznamem** – slouží pro Editor slovníků a konkrétně pro funkci automatické tvorby telefonního seznamu pracovníků TUL; soubor musí mít danou strukturu, aby bylo možné slovník vytvořit
- **adresa rozpoznávače** – pomocí nastavení této adresy se klientské zařízení připojuje k distribuovanému rozpoznávacímu systému

Více informací a detailnější popis promluvy „Přidat stránku“ nebo tvorby slovníku s telefonním seznamem je v příloze A.

Takto nastavené parametry slouží k dosažení větší stability aplikace, proto nedochází k neočekávaným chybám a padání softwaru. V případě chybějícího parametru nedojde ke spuštění aplikace. Je zapotřebí inicializační soubor smazat a provést jeho nastavení opětovně.

Další parametry inicializačního souboru:

- **cesta dalšího adresáře slovníků** – tento parametr slouží pro soubor s přehledem slovníků, aby v přehledu byly obsaženy veškeré slovníky, které jsou uživatelem využívány
- **barva při nahrávání** – zde je uložena informace jaké orámování ovládací ikony si uživatel přeje, pokud je ikona v režimu nahrávání
- **barva bez nahrávání** – opět informace o orámování ovládací ikony, když je režim bez nahrávání
- **barva odpovědi** – zde je nastavena barva podkladu odpovědi
- **barva textu odpovědi** – zde je určena barva písma při zobrazení odpovědi
- **doba zobrazení odpovědi** – z důvodu urychlení nebo naopak delšího zobrazení odpovědi je možné nastavení doby zobrazení textové odpovědi na monitoru
- **neprůhlednost** – jelikož ovládací ikona je stále na povrchu, tak je možnost nastavení tohoto parametru, aby uživatele neustále nerušila; volba parametru je z pěti přednastavených možností
- **velikost** – nastavení velikosti ovládací ikony, která je na povrchu; je možné volit mezi třemi možnostmi
- **pozice** – obsahuje vždy aktuální pozice ovládací ikony na ploše
- **výběr prohlížeče** – tento parametr slouží k získání informace o internetovém prohlížeči, ve kterém si uživatel přeje otevírat webové stránky a nepřeje si být neustále dotazován
- **uložení slovníku** - má-li parametr hodnotu nezobrazovat, tak v okamžiku rychlého přidání internetové stránky nedojde k zobrazení informace, do jakého slovníku je webová stránka ukládána

V případě chybějícího jakékoliv nedůležitého parametru vyjma výběru prohlížeče a uložení slovníku dojde k nastavení výchozí hodnoty.

4.3.2 Odolnost proti výpadkům spojení

Hlasový distribuovaný systém pro úspěšné vyřešení úlohy používá klientské zařízení, ve kterém je provedena jednodušší část výpočtu rozpoznání. Obtížnější výpočet se následně předá pomocí síťového spojení straně serveru. Jelikož informace jsou zasílány pomocí síťového připojení, je nutné zabezpečit aplikaci proti výpadkům spojení.

Mohou nastat tři situace – připojení existuje, připojení neexistuje a připojení existovalo a došlo k výpadku připojení. Při spuštění aplikace probíhá ověření přítomnosti síťového připojení. V případě existence připojení je předána zpráva, která umožní zaslání uživatelského slovníku straně serveru a zobrazí se ovládací (nahrávací ikona). Systém je propojen a je možné hlasem zadávat promluvy, které následně budou zpracovány.

V případě neexistujícího připojení je předána zpráva, která neprovede žádný úkon směrem k serveru a spustí se čekání v pravidelných intervalech, kdy se kontroluje, zda dojde k připojení počítače k síti. V případě, že se tak stane, dojde k automatickému zaslání zaváděcího slovníku straně serveru a zobrazení ovládací ikony. Tudiž není potřeba se znovu připojovat k distribuovanému systému a vše je provedeno automaticky.

V případě existence síťového připojení je klient spojen se serverem a funguje rozpoznávání promluv řečníka. V okamžiku, kdy dojde k výpadku spojení, je přerušeno nahrávání a opět se spustí čekání na zprávu o přítomnosti spojení. V případě příchodu takové zprávy dojde k obnovení spojení se zasláním používaného slovníku serveru.

Klient se také může připojit k rozpoznávacímu systému pomocí IP adresy, a tudíž není zapotřebí mít přístup k internetu. Distribuovaný rozpoznávací systém může mít zavedeny služby mastera a slavu na nějakém výkonném počítači v lokální síti nebo dokonce na vlastním počítači. V takovém případě stačí mít přístup do lokální sítě nebo se využívá vlastní lokální adresy.

4.3.3 Hlavní panel Komunikátor Sluha

Tato část aplikace nevykazovala příliš mnoho chyb. Bylo proto zapotřebí zabezpečit pouze bezchybné otevírání dalších oken z menu hlavního panelu. Jakýkoliv neexistující soubor využívaný k načtení dat pro nově otevírané okno zapříčinil spadnutí

aplikace. Toto bylo realizováno dotazem na existenci souboru pomocí vyhodnocení daného parametru zapsaného v inicializačním souboru. V případě neexistence souboru je uživatel informován vyskočením tabulky se zprávou o chybě.

Menu obsahovalo dvě téměř stejné položky, které byly spojeny do jednoho tlačítka, což umožní jednodušší a rychlejší ukončení komunikace a načtení nového slovníku do distribuovaného rozpoznávacího systému.

Novými funkcemi v hlavním okně Komunikátoru Sluha jsou aktualizace slovníku, nastavování adresy rozpoznávacího systému a možnost nastavení doby zobrazení odpovědi na monitoru. Tlačítko aktualizace slovníku slouží pro zaslání nově editovaného slovníku (do slovníku byly přidány nové promluvy), který se již využívá k rozpoznávání slov a je tedy nahrán v rozpoznávacím systému. V případě, že je editován jiný slovník než zavedený, nedochází k načtení a zavedení tohoto slovníku do rozpoznávacího systému. Funkce pracuje pouze s aktuálním slovníkem.

Vybráním možnosti nastavení doby zobrazení odpovědi je možné nastavit dobu, po kterou bude odpověď k dispozici na monitoru. Je možné zadat hodnotu v rozmezí od jedné sekundy až po dobu půl minuty.

Nutnou funkcí se stává nastavení adresy rozpoznávacího systému. Klientské zařízení se může připojit k rozpoznávacímu serveru buď pomocí webové adresy, nebo IP adresy. V případě využívání rozpoznávače, který je k dispozici na webové adrese dundis.ite.tul.cz, je potřeba zvolit možnost webová adresa. Druhou možností je zvolení IP adresy, na které je zavedena služba rozpoznávacího systému (masteru). IP adresa zadávaná u klienta se musí shodovat s adresou, která určuje umístění serveru (mastera).

Vzhled hlavního panelu byl upraven přidáním ikony do rámečku a také zobrazováním aktuálního dne a konkrétního data do dolní lišty. Jelikož aplikace pro rozpoznávání řeči je stále spuštěna, není žádoucí, aby v liště (hlavní panel operačního systému) byla zobrazena informace o spuštěné aplikaci. Okno je minimalizováno do oznamovací oblasti, kde dojde k zobrazení ikony S.

Vzhled hlavního panelu Komunikátor Sluha je zobrazen na obrázku 4.4.



Obr. 4.4: Hlavní panel Komunikátor Sluha

4.3.4 Ovládací ikona – nahrávací ikona

Tato část je hlavním prvkem pro spuštění nahrávání rozpoznávacího systému. K zahájení nahrávání je potřeba najet myši na ikonu a ponechat minimálně půl sekundy. Poté probíhá režim nahrávání promluvy po dobu tří sekund. Po skončení této doby se nahrávání automaticky vypne. Ke spuštění opětovného nahrávání je opět potřeba na ovládací ikonu najet.

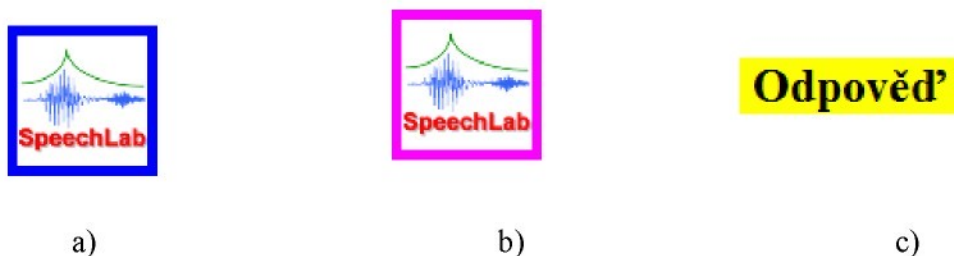
Režim nahrávání je v úvodu znázorněn světle zeleným orámováním nahrávací ikony. Uživatel má možnost si orámování nastavit podle vlastních představ výběrem z palety barev nebo nastavením odpovídající hodnoty RGB. Při režimu bez nahrávání je v úvodu nastavena barva červená. I zde je možnost nastavení vlastní požadované barvy.

V okamžiku, kdy je spuštěno nahrávání a vyřčená promluva, dojde ke zpracování pomocí distribuovaného systému, který odešle výsledek rozpoznání slova zpět klientskému zařízení. Tento výsledek je vždy zobrazen v textové podobě na monitoru a barva podkladu je v úvodu nastavena na tyrkysovou a barva textu je černá. Tyto barvy je také možno nastavit podle vlastních požadavků. Obsahuje-li rozpoznané slovo vytvořenou předponu, dojde i ke grafické odpovědi v podobě spuštění aplikace, otevření internetové stránky atd.

Menu dále obsahuje tlačítko aktualizace slovníku, které slouží stejně jako tlačítko v hlavním panelu aplikace. Tudiž pro zaslání nově editovaného slovníku (do slovníku byly přidány nové promluvy), který se již využívá k rozpoznávání slov a je tedy nahrán v rozpoznávacím systému. Tato funkce byla zvolena z důvodu většího pohodlí pro uživatele, jelikož pomocí ovládací ikony je možnost dostat se do Editoru slovníku a přidat promluvy do slovníku. Pro následnou aktualizaci by bylo třeba zobrazit hlavní panel a poté aktualizovat slovník, takže funkce slouží k urychlení práce s distribuovaným rozpoznávacím systémem.

Dále byla vytvořena možnost přesouvání ovládací ikony i v okamžiku, že je zobrazeno jiné okno spojené se softwarem klientského zařízení. V takovém okamžiku, ovšem nejsou dostupné žádné jiné funkce ať už hlavního panelu nebo menu ovládací ikony, dokud nedojde k uzavření daného okna. Dostupné jsou pouze funkce aktivního okna. V případě využití promluvy „zobrazení slovníku“ dojde k otevření nového okna se zobrazením daného slovníku. Pouze v tomto případě je možné využívat funkci ovládací ikony pro nahrávání.

Jelikož je aplikace pro rozpoznávání řeči stále spuštěna, tak ani okno nahrávací ikony by nemělo být zobrazeno v liště. Z tohoto důvodu je ikona prioritně nastavena, aby byla vždy navrchu. Vzhled ovládací ikony a odpovědi je vidět na obrázku 4.5.



Obr. 4.5: Vzhled ovládací ikony a odpovědi

a) režim při nahrávání **b)** režim bez nahrávání **c)** odpověď

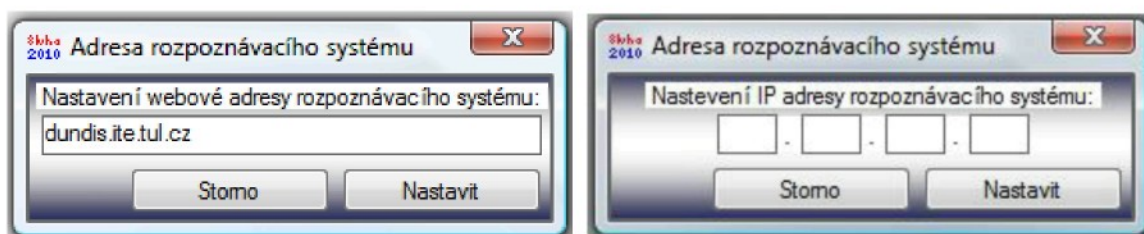
4.3.5 Konfigurace připojení klienta a rozpoznávacího systému

Pro komplexní řešení zadání je potřeba vytvořit možnost připojení klientského zařízení k serveru (masteru) distribuovaného hlasového systému. Je možné využít dvou variant – připojení k systému na adrese `dundis.ite.tul.cz` nebo pomocí IP adresy, kde je zavedena služba serveru.

V případě využívání rozpoznávače, který je k dispozici na webové adrese `dundis.ite.tul.cz`, je potřeba do zobrazeného okna pro nastavení adresy zadat odpovídající adresu. Jelikož se v této aplikaci využívá systém pro rozpoznávání řeči na konkrétní adrese, je automaticky přednastavená v okně pro nastavení potřebné adresy.

Druhou možností je zvolení IP adresy, na které je zavedena služba rozpoznávacího systému (masteru). Rozsah nastavení adresy je v rozsahu 0 až 255, aby uživatel nemohl zadat omylem chybnou IP adresu, která neexistuje. Pro využívání rozpoznávacího systému na IP adrese je potřeba mít zavedeny služby master a slave na počítači, který svou výkonností zvládne zpracovat rozpoznání slova. Nutností je, aby spolu tyto dvě služby byly správně propojeny a věděly o sobě. IP adresa zadávaná u klienta se musí shodovat s nastavenou adresou, která určuje umístění serveru (mastera). V takovém případě není zapotřebí mít připojení k internetu. Stačí mít přístup do lokální sítě, případně veškerý výpočet vyřešit na vlastním počítači.

Na obrázku 4.6 jsou vidět jednotlivá okna pro nastavení adresy rozpoznávacího systému.



a)

b)

Obr. 4.6: Vzhled nastavení adresy rozpoznávacího systému

a) webová adresa b) IP adresa

4.3.6 Editor slovníku

Jedná se o jednu z nejdůležitějších částí klientské části, protože je důležité mít správnou strukturu slovníku, který je zasílán serveru, aby bylo možné promluvu rozpoznat pomocí uživatelského slovníku. Uživatelé ulehčují správnou tvorbu slovníku a hlavní snahou je opět příjemné prostředí a snadná ovladatelnost.

Editor slovníku poskytuje několik základních funkcí jako nový (slovník), otevřít slovník, připojení slovníku, uložit, uložit jako a také konec. Zde bylo potřeba zajistit kvalitnější a propracovanější způsob otevírání, ukládání, či přepisování tvořených slovníků. Hlavní myšlenkou bylo, aby uživatel nemusel opakovaně zadávat jméno či cestu pro uložení slovníku. V případě otevření již existujícího slovníku je název souboru známý. Když dojde k doplnění výrazů nebo odebrání a následnému uložení, není potřeba uživatele obtěžovat dotazy, jak se slovník bude jmenovat a kam se uloží. Tato realizace byla přenesena i do situace, kdy dojde k otevření slovníku, přidání či odebrání výrazu a následně k zvolení nového slovníku nebo konce. Objeví se zpráva o neuloženém slovníku. Při zvolení možnosti uložení slovníku opět nedochází k vyzjištění již známých informací.

Následovala realizace vkládání nově uložených slovníků do přehledu slovníku tak, aby nedocházelo k duplikaci. Výhoda tohoto zajištění se pozitivně projevila na správném zobrazení jednotlivých slovníků v přehledu.

Do editoru slovníku byla přidána funkce na mazání výrazu pomocí připraveného tlačítka v aplikaci nebo stiskem klávesy delete. Vždy je zapotřebí mít označen řádek, který má být smazán. Při úplné změně výrazu je možné editovat jak „Co se má psát“,

tak „Co se má říkat“ a zpětné vložení do slovníku proběhne i v případě, že proběhne změna pouze u odpovědi systému.

Funkce záměny dvou slov u promluvy už podle názvu napovídá, že slouží k výměně první a druhé pozice, například „dobrý den“ je změněno na „den dobrý“. V případě, že je výraz složen pouze z jednoho slova, žádná akce se neprovede.

Funkce přidání internetové stránky zaznamenává správný tvar pro rozpoznávací systém. Tvar se skládá z předpony www a následuje stránka, která může být uvedena v plném tvaru `http://www.domena.cz` nebo ve tvaru `www.domena.cz` a vždy dojde k bezproblémovému spuštění stránky.

Následovala realizace přidávání klávesových zkratk do slovníku. Uživatel má možnost vybrat si pomocí tří zaškrťovacích tlačítek, zda se zkratka bude skládat z tlačítek `control`, `alt` nebo `shift` a k tomu odpovídající klávesa. Po potvrzení je potřeba zadat promluvu, která klávesovou zkratku vyvolá. Při rozpoznání promluvy, kde je odpovědí stisk zkratky, se akce provede vždy v aktivním okně. Jedná se o univerzální algoritmus a je možné pomocí klávesových zkratk ovládat jakékoliv okno, kde je umožněno využívat zkratk. Lišta pro vkládání klávesových zkratk je zobrazena na obrázku 4.7.



Obr. 4.7: Vložení klávesové zkratky do slovníku

Realizace algoritmu, která slouží k vyvolání klávesové zkratky s jedním písmenem, je ukázána následujícím zdrojovým kódem.

```
if (pocetPlus == 0) //kláv.zkratka bez speciální klávesy (CTRL, ALT, Shift) – pouze písmeno
{
    vkladanyText = gnew String(textKlavZkr[4], 1); //odpověď na dispej
    const int VK_KLAV = (int)textKlavZkr[4]; //nastavení hodnoty odpovídající klávesy
    //provedení stisku a uvolnění klávesy
    keybd_event(VK_KLAV, 0, 0, 0); keybd_event(VK_KLAV, 0, KEYEVENTF_KEYUP, 0);
}
if (pocetPlus == 1) //stisk jedné speciální klávesy + písmene
{
    if (pozicePlus == 4) // stisku klávesy ALT + písmeno {
        vkladanyText = "ALT + " + textKlavZkr[pozicePlus + 4];
        const int VK_KLAV = (int)textKlavZkr[pozicePlus + 4];
        //provedení stisku ALT, písmene a následné uvolnění písmena a ALT
        keybd_event(VK_MENU, 0, 0, 0); keybd_event(VK_KLAV, 0, 0, 0);
        keybd_event(VK_KLAV, 0, KEYEVENTF_KEYUP, 0);
        keybd_event(VK_MENU, 0, KEYEVENTF_KEYUP, 0);
    }
}
```

```

    if (pozicePlus == 5) //stisk CTRL + písmeno
    {
        //kód stejný jen stisk CTRL
    }
    if (pozicePlus == 6) //stisk Shift + písmeno
    {
        // kód stejný jen stisk Shift
    }
}
if (pocetPlus == 2) //stisk dvou speciálních kláves + písmeno
{
    if (pozicePrvnihoPlus == 5) //zaručen stisk klávesy CTRL
    {
        if (poziceDruhehoPlus == 4) //druhou klávesou je ATL
        {
            //stisk CTRL + ALT + písmeno
        }
        if (poziceDruhehoPlus == 6) //druhou klávesou je Shift
        {
            //stisk CTRL + Shift + písmeno
        }
    }
    if (pozicePrvnihoPlus == 4) // zaručen stisk ALT + Shift + písmeno
    {
    }
}
}
if (pocetPlus == 3) // stisk CTRL + ALT + Shift + písmeno
{
}
}

// následuje ještě zobrazení odpovědi v podobě textového výpisu klávesové zkratky

```

Editor umožňuje tvorbu slovníku z databáze. Jednoduchým způsobem lze vytvořit spojení libovolných sloupců do „Co se má psát“ a „Co se má říkat“. Tato funkce je prospěšná v okamžiku větší databáze s informacemi, které mají sloupcovou strukturu a je potřeba rychlého vyhledání libovolné informace. Soubor pro tvorbu slovníku musí mít danou strukturu. Mezi jednotlivými sloupci musí být vytvořená mezera pomocí tabulátoru. Na tabulce 4.1 je sloupcová struktura souboru.

Tab.4.1: Struktura souboru pro tvorbu slovníku (soubor s hlavními městy)

1.sloupec	2.sloupec
Česká republika	Praha
Slovensko	Bratislava
Německo	Berlín

Následuje volba parametrů potřebných pro vytvoření slovníku. Prvním je, kolik slovník obsahuje sloupců. Maximální hranice počtu sloupců byla stanovena na deset. Druhým je, jaké sloupce se spojí do odpovědi. Je možno provést spojení až deseti sloupců.

Poslední parametr určuje, jaké sloupce se spojí do promluvy. Ukázka tvorby slovníku s hlavními městy s daty z Tab.4.1, kdy do odpovědi bude zadán druhý sloupec s hlavními městy a promluva bude obsahovat stát, je ukázáno na obrázku 4.8.

a)

b)

c)

Obr. 4.8: Tvorba slovníku hlavních měst

a) volba počtu sloupců v databázi

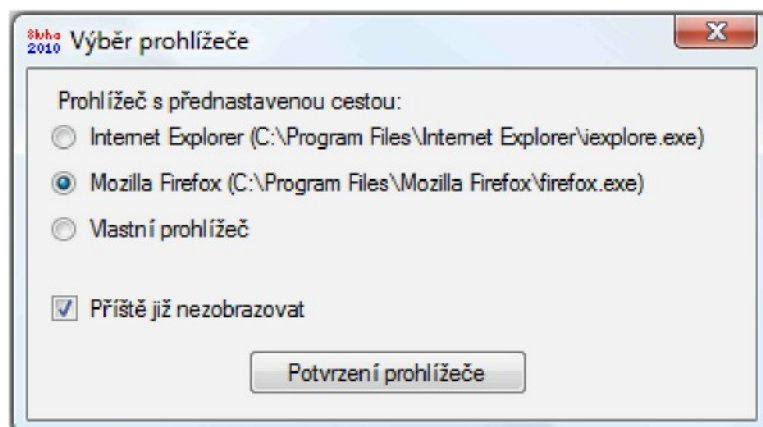
b) spojení sloupců do odpovědi c) spojení sloupců do promluvy

4.3.7 Výběr prohlížeče internetových stránek

Při vložení internetové adresy do slovníku je potřeba zvolit, v jakém internetovém prohlížeči se stránka spustí. Byla zvolena koncepce dvou přednastavených webových prohlížečů – Internet Explorer, který je dodáván k operačním systémům Windows a druhým je Mozilla Firefox, který vyšel z průzkumu jako nejvyužívanější. V případě využívání jiného prohlížeče má uživatel možnost zvolit si vlastní, který nejčastěji používá. Označením možnosti vlastní prohlížeč je otevřen dialog, kde uživatel vybere soubor, kterým se prohlížeč spouští.

V aplikaci je testováno a kontrolováno, zda vybraný prohlížeč existuje a je tedy možné stránku spustit. V případě, kdy není možné najít spustitelný soubor na uvedené adrese, dojde k zobrazení zprávy o neexistujícím souboru a otevření okna s výběrem prohlížeče, aby uživatel zadal vhodnou a správnou cestu. Novou promluvou, která vede ke spuštění webové adresy, dojde k otevření zvoleného prohlížeče a odpovídající internetové adresy.

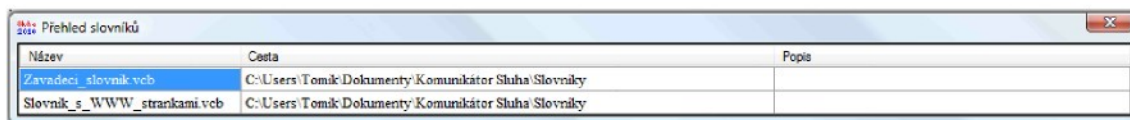
Na obrázku 4.8 je zobrazeno okno výběru prohlížeče.



Obr. 4.8: Výběr prohlížeče internetových stránek

4.3.8 Přehled používaných slovníků

Toto okno nemá charakter pouze obecného náhledu na slovníky, ale umožňuje detailní a propracovaný nahlížení na všechny slovníky. Jedná se o ty, které jsou v hlavním adresáři slovníku a dalších adresářích slovníků. Navíc také umožňuje zobrazení slovníků, které uživatel uloží do jiných adresářů. Pomocí dvojkliku na název slovníku dojde k zobrazení slov (odpovědi a promluv). V případě dvojkliku na neexistující slovník je uživatel dotázán, zda má dojít ke smazání slovníku. Pro detailnější popis slovníků slouží kolonka popisu, kterou si vytváří uživatel sám. V případě potřeby vymazání slovníku ze seznamu přehledu a také smazání slovníku z disku počítače stačí, kliknou na daný řádek a pomocí stisku klávesy delete dojde k vymazání slovníku z přehledu i počítače. Na obrázku 4.9 je zobrazeno okno přehledu slovníků.



Obr. 4.9: Přehled slovníků

4.4 Instalace klientského zařízení

Pro vyšší komplexnost klientského zařízení byla vytvořena instalace, která umožní jednoduché a pohodlné zavedení klienta na počítač uživatele. Jelikož je software navržen pro uživatele hovořící česky, bylo nutností, aby i instalace byla

v českém jazyce. Aplikace byla vytvářena v prostředí Microsoft Visual Studio 2008, proto bylo snahou provést vytvoření instalace také v tomto prostředí. Daný software ovšem nenabízel možnost pro instalaci v českém jazyce. I přes tento nedostatek nebylo uvažováno o využití jiného softwaru, kde by existovala takováto možnost. Tudiž došlo k vytvoření české lokalizace pro tvorbu instalace v prostředí Microsoft Visual Studio 2008.

Strategie návrhu zobrazování jednotlivých oken při instalaci byla navržena s ohledem na standardně používané instalátory jiných aplikací, aby se uživatel lépe orientoval. Snahou také bylo, aby program byl nainstalován do složky Program Files, kam se standardně programy ukládají. Ovšem v průběhu testování aplikace vznikl problém s editováním souborů jako například inicializačního souboru *Nastaveni.ini* či dalších v případě, že soubory byly uloženy právě v adresáři Program Files. V takovém případě bylo zapotřebí otevírat soubory s právy administrátora, což pro danou aplikaci nebylo žádoucí, jelikož se nejedná o aplikaci, kde by bylo těchto práv zapotřebí. Proto proběhla změna pro ukládání souborů do jednotlivých složek.

Hlavní adresář, který obsahuje spouštěcí soubor Komunikátor Sluha.exe a soubor pro odinstalování aplikace, je umístěn právě v adresáři Program Files. Soubory s poskytovanými slovníky, inicializační soubor a další jsou umístěny v uživatelském adresáři, který je umístěn v Documents and settings pod jménem uživatele používajícího v současné době počítač. Takto zvolená realizace přináší ještě další výhodu a to možnost vlastního nastavení klienta pro libovolného uživatele a nedochází ke vzájemnému ovlivňování nastavených parametrů klienta.

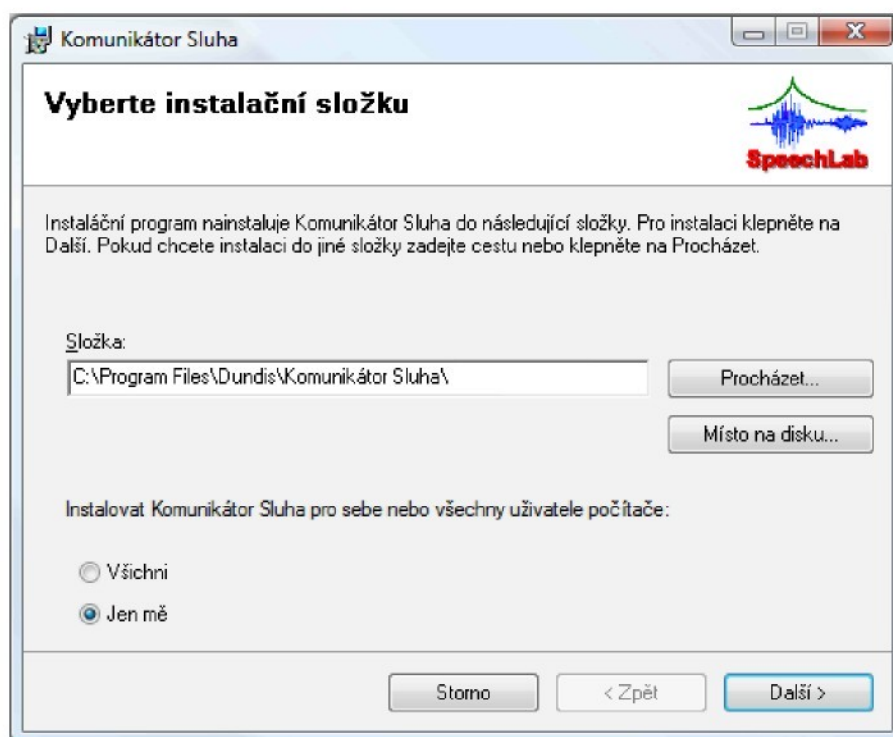
V úvodu proběhne ověření nutných programů potřebných pro spuštění aplikace. Jedná se o program .NET Framework určité verze, závislé na daném operačním systému, a Windows Installer verze 3.1. V případě chybějícího programu .NET Framework dojde ke spuštění instalace přímo z disku a není potřeba program stahovat z internetu. Výhodou této možnosti je využití distribuovaného rozpoznávacího systému i v případě, kdy není dostupné připojení. Pomocí zavedení služeb rozpoznávacího systému na daný počítač, umožní rozpoznávání řeči i bez připojení k internetu.

Po ověření potřebných součástí dochází ke spuštění samotné instalace klientského zařízení na počítač. Úvodní okno obsahuje logo Laboratoře počítačového zpracování řeči, která spadá pod Ústav Informačních technologií a elektroniky, jelikož

se podílela na vývoji této diplomové práce poskytnutím distribuovaného rozpoznávacího systému.

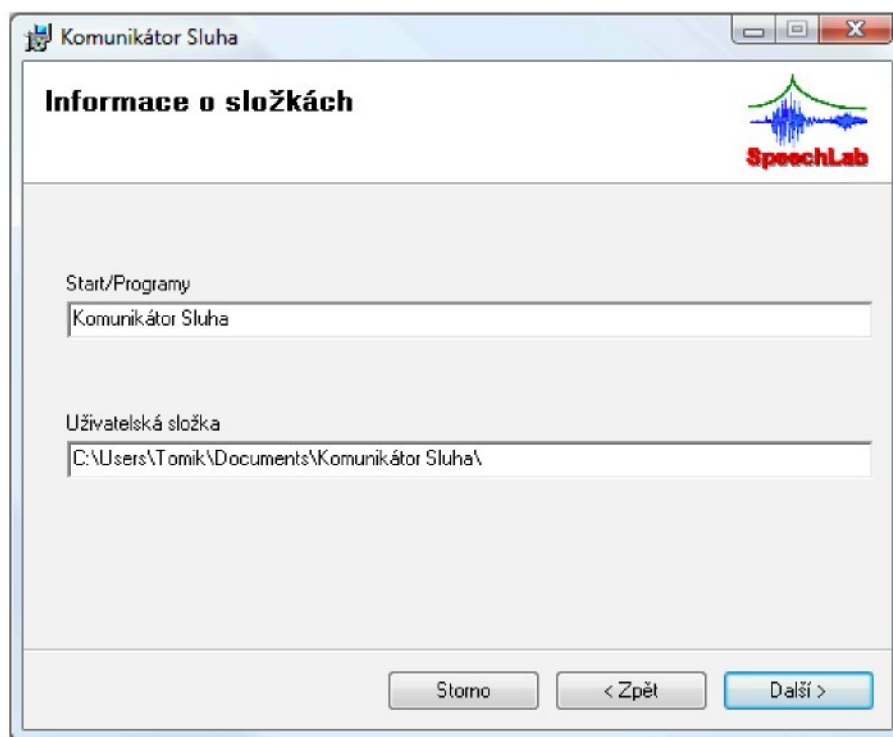
Následuje zobrazení uvítacího okna k instalaci klientského zařízení Komunikátor Sluha. Poté je zobrazena licenční smlouva, kde jsou uvedeny informace, aby nedocházelo k volnému pořizování kopií za účelem dalšího prodeje či vymezení odpovědnosti za chybu.

Instalace má nastavenou cestu C:\Program Files\Dundis\Komunikátor Sluha pro hlavní adresář, jak je vidět z obrázku 4.10, kde je možné eventuálně zadat jinou cestu. Pro standardně ukládané aplikace je zvolen adresář Program Files, proto i pro tuto aplikaci byl přednastaven tento adresář. Dále byl přidán adresář Dundis, kam se bude instalovat, jak klientské zařízení Komunikátor Sluha, tak i Nastavení rozpoznávače.

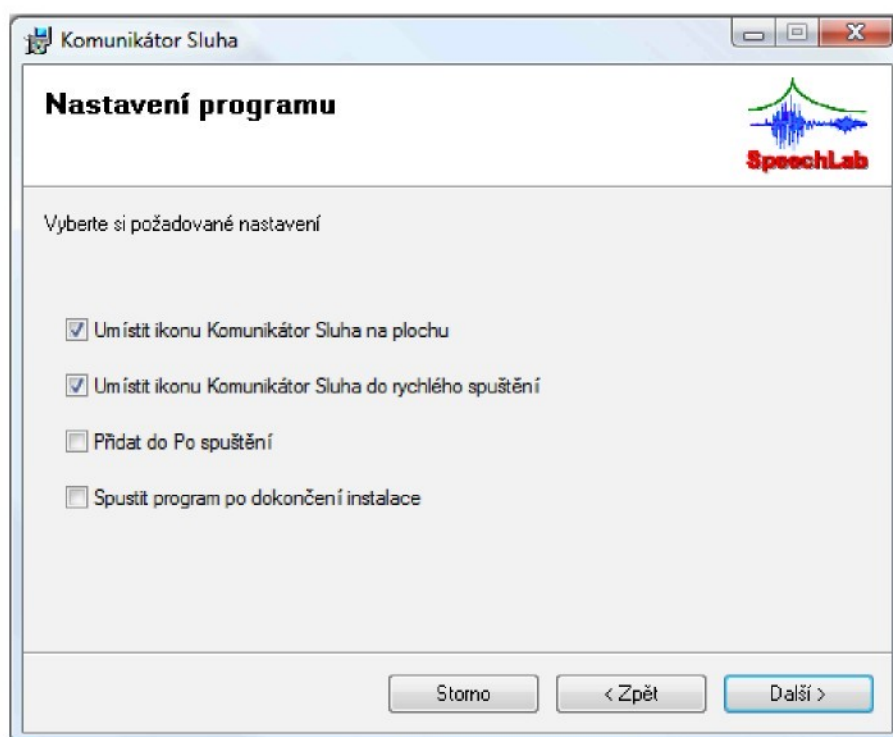


Obr. 4.10: Výběr instalační složky Komunikátoru Sluha

Následuje zobrazení okna s informacemi, kde je vytvářena uživatelská složka a také že dojde k vytvoření adresáře do Programů v nabídce Start. Okno je znázorněno na obrázku 4.11. Obrázek 4.12 ukazuje jednotlivé možnosti uživatelského nastavení, které vytváří větší ucelenost dané aplikace a uživateli poskytuje pohodlí při instalaci klientského zařízení. Toto okno se mírně liší pro jednotlivé verze operačních systémů. U verze Windows 7 není k dispozici možnost Umístit ikonu Komunikátor Sluha do rychlého spuštění.



Obr 4.11: Okno s informacemi o složkách Komunikátoru Sluha



Obr. 4.12: Možnost uživatelského nastavení Komunikátoru Sluha

Následuje průběh instalace a okno pro potvrzení správného nainstalování aplikace. V případě zvolení možnosti Spustit program po dokončení instalace dojde k prvnímu spuštění klienta.

5 Konfigurace rozpoznávacího systému

Data zpracovaná v klientském zařízení jsou zaslána serveru, který z přijatých dat provede detekci začátku a konce slova (hledání ticha v signálu) a předá data slave počítačům na další zpracování. Proběhne další parametrizace a jsou vytvořeny delta a delta-delta příznaky. Následuje porovnání s modely a vyhodnocení rozpoznávaného slova, které je předáno serveru, který data předá zpět klientovi.

Pro vytvoření expediční kompletace distribuovaného systému bylo potřeba vytvořit aplikaci (konfigurátor), která umožní zavedení služeb (master, slave) rozpoznávacího systému. Díky této aplikaci bude možné systém využívat dvěma způsoby. Využíváním rozpoznávače z webové adresy dundis.ite.tul.cz nebo zavedením rozpoznávacího systému na výkonný počítač v lokální síti nebo na vlastní PC. Tím je umožněno rozpoznávání promluv i bez nutnosti internetového připojení.

Dané konfigurační aplikace byly již v dřívější době vytvořeny pomocí prostředí, které v dnešní době není využíváno. Proto došlo k tvorbě nových aplikací pro nastavování rozpoznávacího systému v používaném prostředí Microsoft Visual Studio C++ .NET. Samozřejmě také byla dodržena strategie pro využívání aplikace u česky mluvících lidí.

Opět byla vytvořena instalace, která ověří zda je na počítači nainstalovaná potřebná verze .NET Frameworku. Pomocí této instalace dojde ke správnému vytvoření adresáře, který poskytuje potřebné soubory pro nastavení služeb systému Windows. Zároveň instalace dává punc uživatelsky příjemnější aplikace.

5.1 Master - server

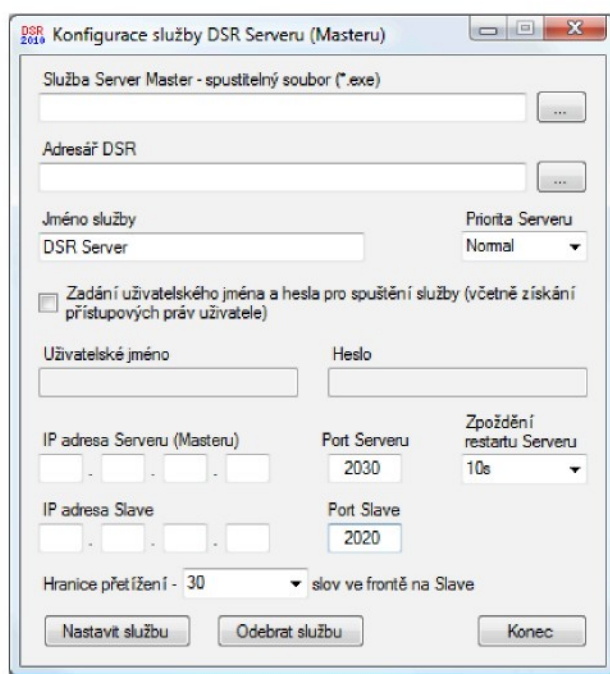
Pomocí vytvořené aplikace Konfigurace služby DSR Serveru (Masteru) je možné zavést první potřebnou službu rozpoznávacího systému. Jedná se o zavedení služby serveru do operačního systému. U této služby je potřeba nastavit důležité parametry, aby došlo ke správnému propojení klienta se serverem a také serveru se slavem. Dále je třeba zadat cestu k dané službě serveru, která se má zavést do operačního systému a také jméno zaváděné služby. Posledním důležitým parametrem je zadání adresáře se soubory rozpoznávače. V tomto adresáři je potřeba mít minimálně tři soubory potřebné pro synchronizaci slave počítačů. Jedná se o soubory *default.vcb*

(přípona vcb označuje soubor slovníku), *default.mod* (přípona mod označuje soubor hláskových modelů) a *default.abc* (přípona abc označuje soubor abecedy)

Propojení klienta a serveru se provede nastavením IP adresy a odpovídající hodnoty portu. Hodnota portu pro spojení s klientem je zadána na 2030. Propojení serveru a slave se provede nastavením IP adresy a zadáním hodnoty portu. V tomto případě je nastavena hodnota portu na 2020.

Zavedení služby serveru do operačního systému je možné pomocí tlačítka *Nastavit službu*. Podmínkou pro zavedení je mít administrátorská práva. Dojde k vytvoření dvou souborů *DSRServer.ini* a *DSRServer.log*. Inicializační soubor je serverem automaticky načten při spuštění služby. Log soubor poskytuje informace o spojení klienta se serverem, serveru se slavem nebo o případných vzniklých chybách. Odstranění služby ze systému je možné provést tlačítkem *Odebrat službu* a zadáním jména služby serveru. Opět je nutností mít administrátorská práva.

Hodnota zpoždění restartu serveru slouží k opakování pokusu o spojení v pravidelných intervalech. Pokud je zaškrtnuta možnost *Zadání uživatelského jména a hesla*, je možnost nastavit účet, pod kterým má služba běžet, včetně získání uživatelských práv daného uživatele. V opačném případě běží služba pod standardním systémovým účtem. Vzhled zavedení služby server do operačního systému je vidět na obrázku 5.1.



Obr. 5.1: Konfigurace služby serveru (master)

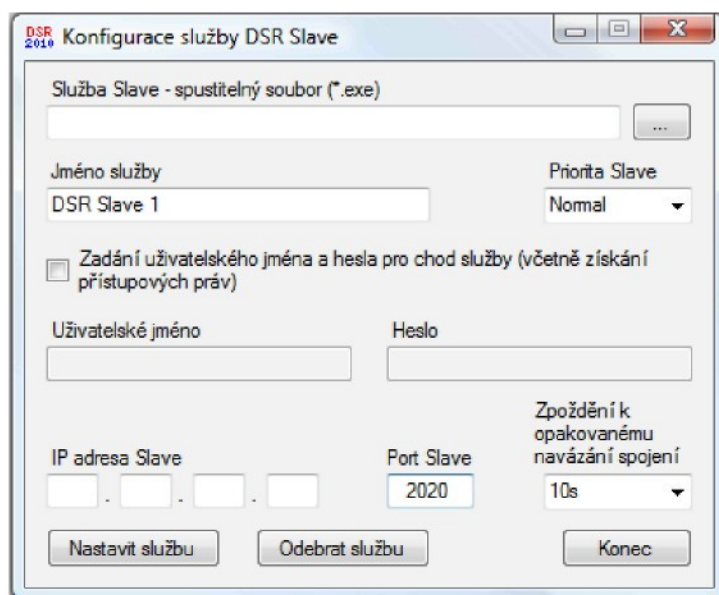
5.2 Slave

Pomocí vytvořené aplikace Konfigurace služby DSR Slave je možné zavést druhou službu rozpoznávacího systému. Dochází k zavedení služby slave do operačního systému. V aplikaci pro nastavení služby je potřeba zadat nutné parametry, aby bylo umožněno spojení slave s masterem. Propojení slave a master se provede zadáním IP adresy a portu, která byla nastavena při zavádění služby serveru. Hodnota portu je přednastavena na 2020. Dále je třeba zadat cestu k dané službě slave, která se má zavést do operačního systému a také jméno zaváděné služby.

Zavedení služby slave do operačního systému se provádí stejným způsobem jako služba serveru. Je potřeba mít administrátorská práva a stisknout tlačítko *Nastavit službu*. Opět dojde k vytvoření dvou souborů DSRSlave.ini a DSRSlave.log. Inicializační soubor pracuje stejně jako u serveru. Log soubor poskytuje informace o spojení slavu se serverem nebo o případných vzniklých chybách. Odstranění služby ze systému může být provedeno v případě administrátorských práv, zadáním jména a stiskem tlačítka *Odebrat službu*.

Hodnota zpoždění k opakovanému navázání spojení slouží k opakování pokusu o spojení v pravidelných intervalech v případě ztráty spojení. Vlastnost *Zadání uživatelského jména a hesla* funguje stejně jako u služby serveru.

Vzhled zavedení služby slave do operačního systému je vidět na obrázku 5.2.



Obr. 5.2: Konfigurace služby slave

5.3 Instalace

Stejně jako u klientské části byla pro větší pohodlí vytvořena instalace nastavení rozpoznávacího systému, která umožní jednoduchou a pohodlnou tvorbu souborů potřebných pro zavádění služeb do operačního systému. Jelikož je software navržen pro uživatele hovořící česky, byla vytvořena instalace v českém jazyce. Instalace byla vytvářena v prostředí Microsoft Visual Studio 2008, kde již byla česká lokalizace pro tvorbu instalace.

Opět byla zachována strategie návrhu zobrazování jednotlivých oken s ohledem na standardně používané instalátory. Program je instalován do složky Program Files, kam se standardně programy ukládají.

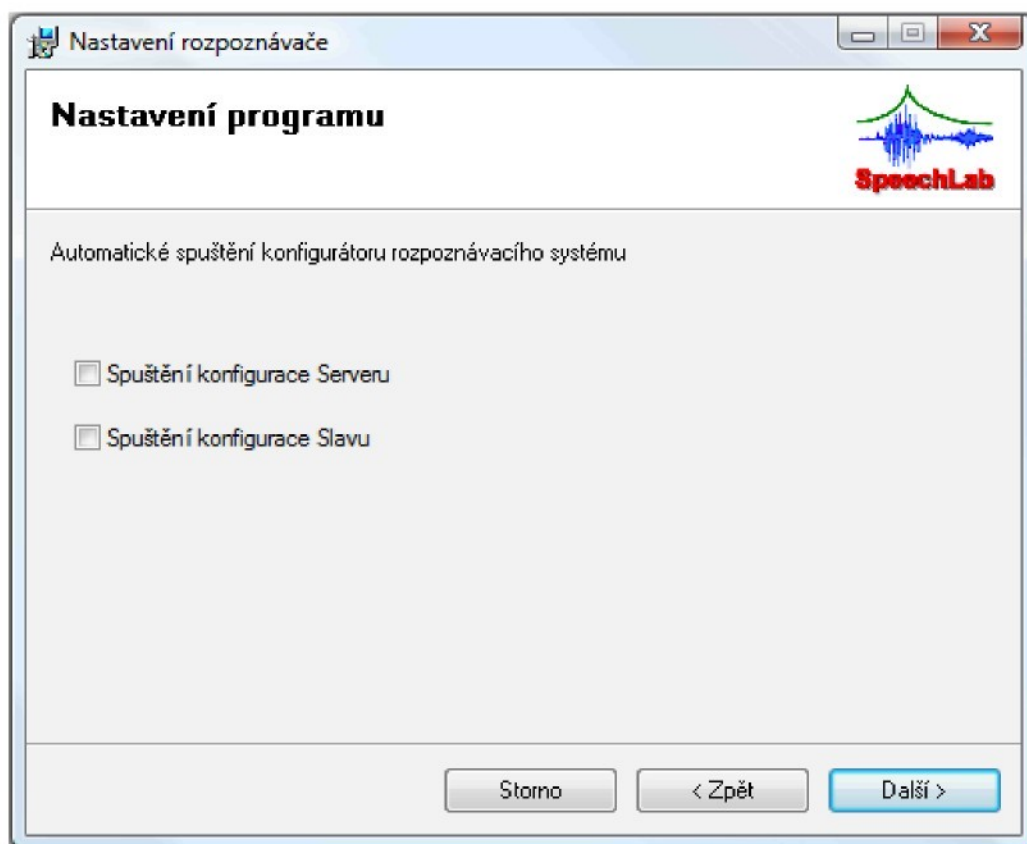
V úvodu proběhne ověření nutných programů potřebných pro spuštění aplikace. Konkrétně jde o program .NET Framework určité verze, závislé na daném operačním systému, a Windows Installer verze 3.1. V případě chybějícího programu .NET Framework opět dojde ke spuštění instalace z disku. Po ověření potřebných programů dochází ke spuštění instalace nastavení rozpoznávacího systému.

Instalace začíná zobrazením úvodního okna, kde je opět zobrazeno logo Laboratoře počítačového zpracování řeči. Následuje zobrazení uvítacího okna k instalaci nastavení rozpoznávacího systému. Poté je zobrazena licenční smlouva. Obsahuje vymezení odpovědnosti za škodu, omezení užívání ve smyslu, aby nedocházelo k volnému pořizování kopií za účelem prodeje.

Instalace má přednastavenou cestu pro hlavní adresář. Danou cestu je možné změnit v okně nastavení instalační složky pro aplikaci. Zvolena byla cesta C:\Program Files\Dundis\Nastavení rozpoznávače. Do adresáře Dundis je ukládán klient Komunikátor Sluha, takže z adresáře je možné provést spuštění klientského zařízení nebo nastavení rozpoznávacího systému.

Obrázek 5.3 ukazuje možnost uživatelského nastavení. Je možno nastavit dva parametry, zda má dojít v průběhu instalace ke spuštění aplikace konfigurace služby serveru nebo ke spuštění aplikace konfigurace služby slave.

Následuje průběh instalace a v případě zvolení možností spuštění konfigurací je nejprve potřeba odpovídajícím způsobem nastavit služby server a slave zaváděné do operačního systému a až poté je instalace dokončena.



Obr. 5.3: Uživatelská možnost nastavení spuštění konfigurace při instalaci

6 Webové prostředí

Pro možnost dostat se k aplikacím distribuovaného rozpoznávacího systému jiným způsobem, než jenom pomocí přenosného média, bylo vytvořeno webové prostředí, kde je možné potřebný software stáhnout. Tím dojde k dostupnějšímu šíření rozpoznávacího systému mezi nové uživatele.

Hlavní myšlenkou tvorby webu bylo rozdělení na jednotlivé sekce. Jednou sekci je úvodní stránka, další je stránka s přehledem novinek na dostupných verzích. Dále je vytvořena sekce s informacemi o programu Komunikátor Sluha a Konfigurátor rozpoznávacího systému a sekce stahování jednotlivých částí distribuovaného systému.

Sekce úvodní stránky obsahuje obecné informace o poskytovaných aplikacích na webu a také poděkování vedoucímu diplomové práce a ostatním kolegům z Laboratoře počítačového zpracování řeči, kteří se podíleli na vývoji distribuovaného rozpoznávacího systému použitého v této diplomové práci.

V sekci *Novinky* jsou zaznamenány nově vytvořené verze jednotlivých programů – Komunikátor Sluha nebo Konfigurátor rozpoznávacího systému. U nových verzí je vždy popsáno vylepšení oproti předešlé verzi a také odkaz do sekce stahuj s novou verzí aplikace.

Sekce *O programu* je rozdělena na dvě části. První popisuje aplikaci Komunikátoru Sluha a druhý obsahuje popis Konfigurátoru nastavení rozpoznávacího systému. Kromě popisu jednotlivých aplikací ohledně možnosti využití, jsou v této sekci ukázány i některé funkce, které aplikace nabízejí.

V poslední sekci *Stahuj* je možné stáhnout jednotlivé verze pro odpovídající operační systémy. Klient je vytvořen ve dvou variantách. První je aplikace, kde dojde k automatickému nastavení adresy pro rozpoznávací systém. Jedná se o adresu `dundis.ite.tul.cz` a tudíž je potřeba připojení k internetu, aby mohlo probíhat rozpoznávání. Druhou je aplikace, kde je možné zadat vlastní adresu rozpoznávacího systému. Konfigurace rozpoznávacího systému má vytvořenou pouze jednu variantu.

Tyto aplikace nabízejí ještě možnost rozdělení podle operačního systému. Aplikace Komunikátoru Sluha nabízí tři varianty – pro Windows 7 nebo pro Windows Vista a Windows XP se Service Packem 2 a poslední je pro Windows XP bez Service Packu a Windows 2000. Aplikace pro konfiguraci rozpoznávacího systému nabízí dvě

varianty – pro Windows 7, Vista a XP se Service Packem 2 nebo pro Windows XP bez Service Packu a Windows 2000.

Vytvoření webových stránek bylo realizováno v prostředí Microsoft Expression Web 3. V současné době jsou zprovozněny zkušební stránky na provizorním serveru. Adresa současných stránek je <http://dundis.ite.tul.cz:1111/klucho/>.

7 Testování klientského zařízení

Testování aplikace Komunikátor Sluha se účastnila skupina 20 lidí, kteří testovali funkčnost tohoto softwaru. Předávali informace o konfiguracích počítačových sestav, na kterých byla aplikace testovaná. A také zodpovídali na 4 otázky ohledně ovladatelnosti aplikace a využití nápovědy.

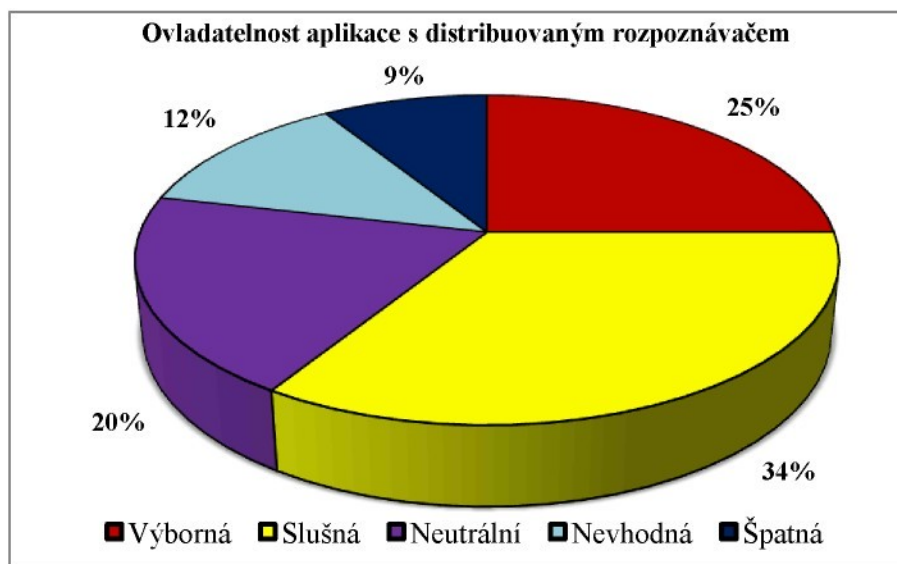
Počítačové sestavy zahrnovaly různé konfigurace. Asi nejslabší testovanou konfigurací byl počítač s procesorem Intel Celeron M 370 1,5GHz, operační paměť 2x512MB DDR RAM a operačním systémem Windows XP Service Pack 2.

Testování aplikace Konfigurace rozpoznávače byla provedena na počítači následující konfigurace - procesor Intel Core 2 Duo E8400 – 3,00GHz, operační paměť 2x2GB DDR2 RAM a hlavním operačním systémem Windows 7 64bit. Ostatní operační systémy byly simulovány pomocí vytvoření virtuálního operačního systému.

7.1 Vyhodnocení výsledků

Uživatelé, kteří testovali aplikaci klientského zařízení, odpovídali i na čtyři otázky ohledně ovladatelnosti Komunikátoru Sluha. Ze získaných informací je vidět pozitivní ohlas na ovladatelnost. Celkové výsledky jsou znázorněny na grafu 7.1.

Editor slovníku byl také pozitivně hodnocen, jelikož umožňuje jednoduchou a rychlou tvorbu uživatelských slovníků. Především dodržení správné struktury pro distribuovaný rozpoznávací systém a také vkládání jednotlivých předpon pro spouštění internetových stránek či aplikací.



Graf 7.1: Vyhodnocení ovladatelnosti aplikace s distribuovaným rozpoznávačem

7.2 Podporované verze operačního systému

Aplikace byly tvořeny pro operační systém Microsoft Windows. Testování aplikací probíhalo na všech systémech od verze Windows 2000 a vyšší. Test proběhl i na verzi Windows 7 64bit. Jednotlivým verzím operačních systémů odpovídají určité verze programu.

Přehled podporovaných systémů pro jednotlivé aplikace je vidět v tabulce 7.1, 7.2 a 7.3.

Tab. 7.1: Komunikátor Sluha s pevně danou adresou rozpoznávacího systému

Název	Operační systém
Komunikátor Sluha (Win7_no)	Windows 7
Komunikátor Sluha (Vista_XPSP2_no)	Vista, Windows XP Service Pack 2
Komunikátor Sluha (XP_2000_no)	Windows XP bez Service Packu, Windows 2000

Tab. 7.2: Možnost nastavení adresy rozpoznávacího systému u Komunikátoru Sluha

Název	Operační systém
Komunikátor Sluha (Win7)	Windows 7
Komunikátor Sluha (Vista_XPSP2)	Vista, Windows XP Service Pack 2
Komunikátor Sluha (XP_2000)	Windows XP bez Service Packu, Windows 2000

Tab. 7.3: Možnost nastavení adresy rozpoznávacího systému u Komunikátoru Sluha

Název	Operační systém
Nastavení rozpoznávače (Win7_Vista_XPSP2)	Windows 7, Windows Vista, Windows XP Service Pack 2
Nastavení rozpoznávače (XP_2000)	Windows XP bez Service Packu, Windows 2000

7.3 Shrnutí

Realizace expediční kompletace distribuovaného hlasového rozpoznávacího systému byla v prostředí Microsoft Visual Studio C++ .NET, který nabízí nové struktury a lepší programovatelné vlastnosti než starší jazyky.

V Laboratoři počítačového zpracování řeči je vyvinut software MyVoice. Tato aplikace umožňuje komplexní ovládání počítače od spouštění aplikací, přes pohyb myši na ploše až po vkládání písmen do libovolného textového editoru či zadávání internetových adres. Jedná se o software určený převážně handicapovaným osobám, kterým by bez tohoto programu nebylo umožněno pracovat s počítačem. Jelikož je

možné ovládat téměř cokoliv, tak i slovník obsahuje velké množství slov, které si uživatel musí pamatovat, aby dosáhl úspěšného cíle.

Oproti tomu aplikace Komunikátor Sluha není prioritně určena pro využití osobám s handicapovým postižením, jelikož je potřeba spouštět nahrávání najetím myši na ovládací ikonu. Výhodou tohoto kompletního systému je možnost ovládání chodu počítače – spouštění internetových stránek, aplikací a ovládání libovolného aktivního okna pomocí klávesových zkratk. Dále je systém využitelný pro vyhledávání informací z databází. To vše umožňuje uživatelský slovník, který je uživatelem tvořen a volí si sám vlastní promluvy k jednotlivým odpovědím. Výhodou tedy je i velikost slovníku, kde nejsou žádné nepotřebné výrazy.

Hlavní cíle diplomové práce byly splněny a došlo k vytvoření kompletace distribuovaného rozpoznávacího systému, který je možné dále šířit mezi nové uživatele, kteří by rádi využívali možnost ovládání počítače hlasem.

Aplikace klientského zařízení má široké uplatnění, jelikož je funkční na operačních systémech Windows od verze 2000. Nainstalování aplikace se provádí pomocí jednoduchého a intuitivního instalátoru, kde je i možnost uživatelského nastavení určitých parametrů. Aplikace Komunikátoru Sluha se pomocí úvodního nastavení a využívání inicializačního souboru stává více uživatelsky příjemnou a s dobrou ovladatelností a přehledností se software stává komplexní aplikací, která by měla uživatele oslovit. Pro úplné doplnění aplikace je vytvořena nápověda, která uživateli pomůže v případě potřeby vyhledat potřebnou funkci aplikace.

Aplikace nastavení rozpoznávacího systému umožňuje vytvořit si vlastní rozpoznávací systém na vlastním počítači či v lokální síti a tudíž není potřeba být trvale připojen k internetu.

Rozpoznávací systém kvalitním způsobem rozpoznává promluvy izolovaných slov či krátkých spojení a umožňuje zaslání uživatelského slovníku s kapacitou až dva tisíce slov.

Závěr

U aplikace klientského zařízení došlo k vytvoření instalace, úvodního nastavení pro možnost uživatelem modifikovat design aplikace a nových funkcí pro možnost ovládat počítač hlasem. Jednalo se například o možnost jednoduchého načtení nově editovaného slovníku, nastavení adresy pro rozpoznávací systém, vkládání klávesových zkratk do slovníku se správnou strukturou. Tyto funkce dávají aplikaci větší punc uživatelsky příjemnějšího prostředí.

Aplikace nastavení rozpoznávacího systému je opět vytvořena pomocí instalace, která zaručí správné vytvoření potřebných programů k zavádění služeb. Pomocí tohoto softwaru je možné zavést služby rozpoznávacího systému (server, slave) na počítač a vytvořit si tak vlastní rozpoznávací systém. Nespornou výhodou je možnost používat aplikaci na rozpoznávání promluvy i v situacích, kdy uživatel nemá k dispozici internetové připojení a potřebuje vyhledat určité informace z databáze.

Vytvořené aplikace je možné využít na operačních systémech Windows 2000 a vyšší v případě dostatečné konfigurace daného počítače. Z tohoto důvodu se dá předpokládat, že se aplikace stane široce uplatnitelná.

Využití aplikace pro zdravotně handicapované by mohlo nastat v případě nastavení neustálého režimu nahrávání a vytvořením odpovídajícího uživatelského slovníku, který by obsahoval všechny výrazy, které by uživateli zaručily možnost ovládat počítač.

Z práce vyplývá, že problematika v oblasti dialogových systémů je neobyčejně rozsáhlá, složitá a nabízí mnoho prostoru pro další vývoj nových nebo vylepšování stávajících dialogových systémů.

Seznam použité literatury

- [1] Klucho T.: *Návrh interaktivního hlasového rozhraní mezi člověkem a počítačem*. Liberec, 2008. Bakalářská práce na Fakultě mechatroniky a mezipředmětových inženýrských studií TUL. Vedoucí bakalářské práce Ing. Miroslav Holada, Ph.D.
- [2] Psutka J.: *Mluvíme s počítačem česky*. Academia, Praha, 2006, ISBN 80-200-1309-1
- [3] Nouza J.: *Počítačové zpracování řeči*. Sborník článků, Liberec, 2001, ISBN 80-7083-551-6
- [4] Huang X., Acero A., Hon H.W.: *Spoken language Processing, A Guide to Theory, Algorithm and System development*. Prentice Hall PTR. Upper Saddle River, New Jersey, 2001. ISBN 0-13-022616-5
- [5] Pelc M.: *Víceuživatelský hlasový rozpoznávací server s distribuovanými výpočty*. Liberec, 2006. Diplomová práce na Fakultě mechatroniky a mezipředmětových inženýrských studií TUL. Vedoucí bakalářské práce Ing. Miroslav Holada, Ph.D.
- [6] Psutka J.: *Komunikace s počítačem mluvenou řečí*. Academia, Praha, 1995, ISBN 80-200-0203-0
- [7] The Microsoft Developer Network (MSDN) Library, URL: <http://msdn.microsoft.com/library>

Příloha A – manuál aplikace klientské zařízení

Úvodní nastavení Komunikátoru Sluha

Při spuštění aplikace dojde k vytvoření inicializačního souboru *Nastaveni.ini*. Soubor je uložen v adresáři *C:\Documents and Settings\Username\Dokumenty\Komunikátor Sluha*. Soubor obsahuje parametry k nastavení Komunikátoru Sluha.

Potřebné parametry, které musí být v souboru zaznamenány:

- adresář programu
- cesta hlavního slovníku (adresář s prioritními slovníky)
- cesta k souboru s přehledem slovníků
- cesta zaváděcího slovníku
- cesta slovníku s WWW stránkami
- cesta k souboru s telefonním seznamem
- adresa rozpoznávacího systému

Další parametry jsou v úvodu a v případě chybějící informace nastaveny na:

- barva při nahrávání – Lime
- barva bez nahrávání – Red
- barva odpovědi – Cyan
- barva textu odpovědi – Black
- doba zobrazení odpovědi – 3s
- neprůhlednost – 1.0
- velikost – malý
- pozice – 50 x 50

Tyto parametry nemusí být nastaveny:

- prohlížeč internetových stránek
- uložení slovníku

Automatické načtení slovníku

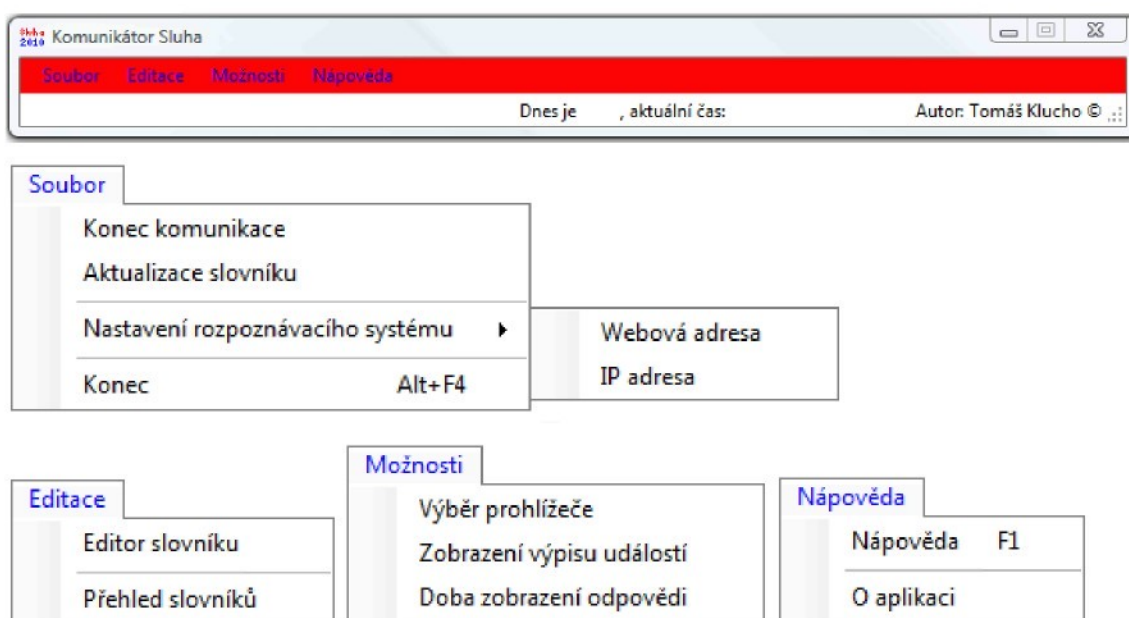
Po úvodním nastavení aplikace dojde k zaslání zaváděcího slovníku do rozpoznávacího systému. V inicializačním souboru je tento slovník u parametru *cesta zaváděcího slovníku*. Výchozí hodnota je nastavena na *C:\Documents and Settings\Username*

name\Dokumenty\Komunikátor Sluha\Slovníky\Zavadeći_slovník.vcb. Po načtení dojde k zobrazení ikony pro nahrávání a nápis „*Slovník načten*“.

V případě, že zaváděcí slovník, který je u daného parametru, neexistuje, dojde k otevření okna pro výběr slovníku. Slovník musí mít příponu *.vcb. Po správném načtení slovníku se zobrazí ikona pro nahrávání a nápis „*Slovník načten*“.

Hlavní okno Komunikátoru Sluha

Zobrazení hlavního panelu a jednotlivých funkcí je na obrázku A.1.



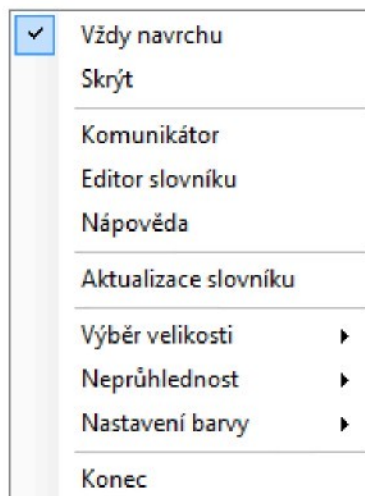
Obr.A.1: Hlavní panel a zobrazení jednotlivých funkcí menu

Nahrávání rozpoznávacího systému + nahrávací ikona

Po načtení slovníku do rozpoznávacího systému je zobrazena nahrávací ikona. V úvodu je orámovaná červeně - režim nenahrávání. Po najetí myši na tuto ikonu a ponechání kurzoru alespoň 0,5s se spustí nahrávání, které je znázorněno zeleným orámováním - režim nahrávání. Nahrávání trvá po dobu 3s a poté je zastaveno. Pro nové nahrávání je nutné znovu na ikonu najet. Pouhé přejetí ikony (pokud není delší jak 0,5s) nahrávání nespustí.

Ikonu můžete pomocí stisku levého tlačítka myši přesouvat po ploše a pomocí pravého tlačítka si zobrazíte menu, které slouží k nastavení nahrávací ikony nebo umožňuje přepnutí do jiného okna. Menu ikony je vidět na obrázku A.2. V obou uvedených případech je nahrávání zastaveno.

Vzhled ovládací ikony a odpovědi je vidět na obrázku A.3.



Obr.A.2: Menu ovládací ikony



a)



b)

Odpověď

c)

Obr. A.3: Vzhled ovládací ikony a odpovědi

a) režim při nahrávání b) režim bez nahrávání c) odpověď

Načtení slovníku / Konec komunikace

V hlavním panelu kliknout na Soubor → Načtení slovníku/Konec komunikace. Tlačítko realizuje dvě funkce. V případě, že je tlačítko v menu označeno textem Načtení slovníku, je možné načíst nový slovník do rozpoznávacího systému. Po kliknutí se zobrazí okno Výběr slovníku pro rozpoznávací systém, kde je potřeba vybrat požadovaný slovník, který musí mít příponu *.vcb. Po načtení slovníku do rozpoznávacího systému se zobrazí ikona pro nahrávání a nápis „Slovník načten“.

V případě, že je tlačítko v menu označeno textem Konec komunikace, dojde k ukončení připojení klienta k rozpoznávacímu systému a dále není možné rozpoznávat promluvu. Pro opětovné spuštění rozpoznávacího systému je potřeba načíst nový slovník pomocí Soubor → Načtení slovníku.

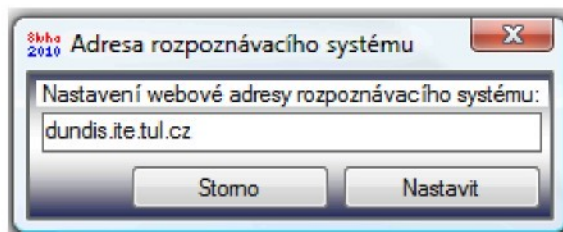
Aktualizace slovníku

V hlavním panelu kliknout na Soubor → Aktualizace slovníku. Dojde k aktualizaci zavedeného slovníku v rozpoznávacím systému. Vhodné použít po doplnění nových výrazů do právě využívaného slovníku.

Nastavení adresy rozpoznávacího systému

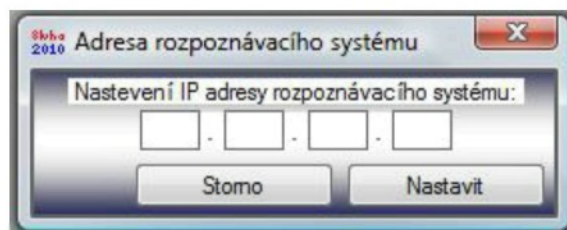
Je možné nastavení dvou adres pro rozpoznávací systém:

1. **Nastavení webové adresy** – v hlavním panelu kliknout na Soubor → Nastavení rozpoznávacího systému → Webová adresa. Nyní je potřeba zadat webovou adresu, kde je nahrán distribuovaný rozpoznávací systém. V našem případě se jedná o adresu dundis.ite.tul.cz a je nastavena při zobrazení daného okna. Nastavení adresy je vidět na obrázku A.4.



Obr.A.4: Nastavení webové adresy pro rozpoznávací systém

2. **Nastavení IP adresy** – v hlavním panelu kliknout na Soubor → Nastavení rozpoznávacího systému → IP adresa zadání. Nyní je třeba zadat IP adresy, kde běží rozpoznávací systém. IP adresa má umožněn rozsah 0 až 255 hodnot, aby nedošlo ke špatnému nastavení adresy. Okno pro nastavení IP adresy je vidět obrázku A.5.



Obr.A.5: Nastavení IP adresy pro rozpoznávací systém

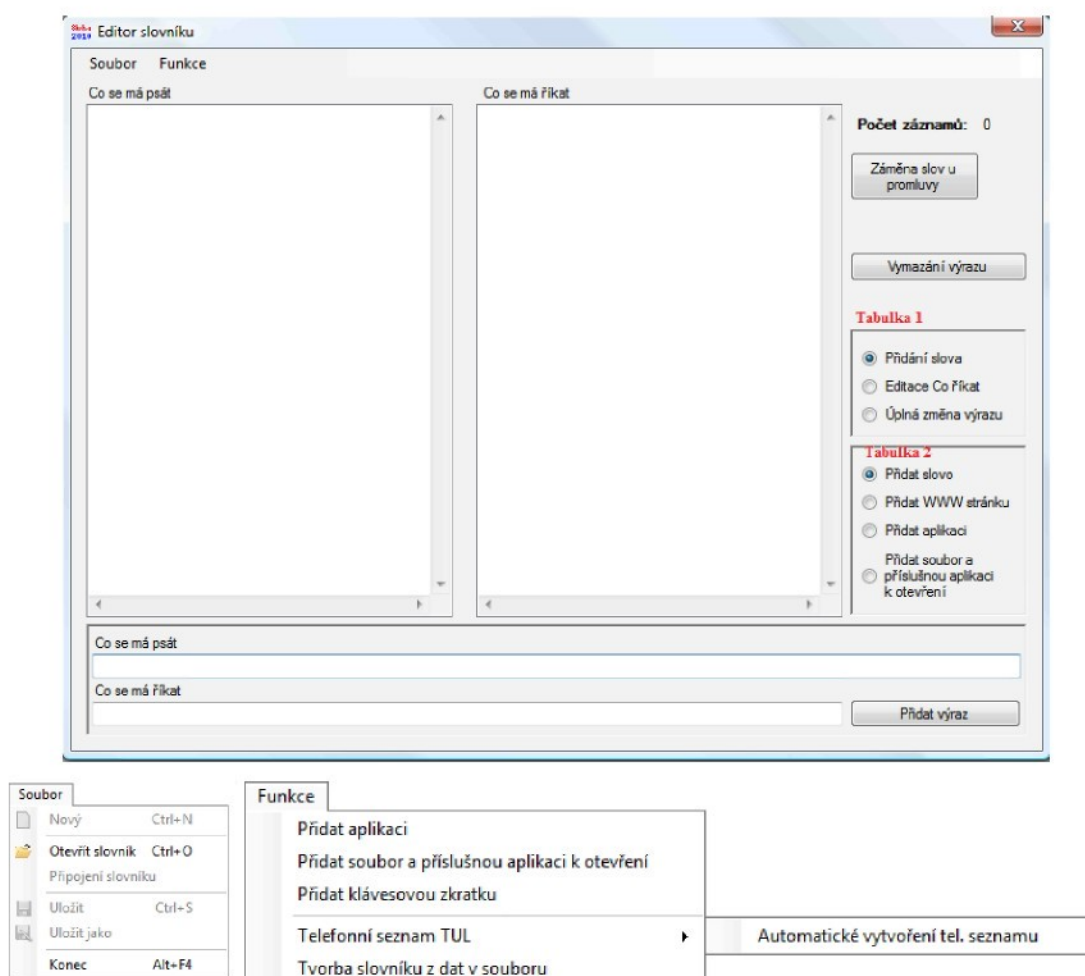
Editor slovníku

Editor slovníku slouží k vytváření nebo editování slovníku se zachováním správné struktury pro rozpoznávací systém. Pro spuštění okna je potřeba v hlavním panelu kliknout na Editace → Editor slovníku.

Dostupné funkce editoru slovníku:

- nový slovník, otevření slovníku, připojení dalšího slovníku k již vytvořenému, uložení a uložení jako
- přidat výraz, WWW stránku, aplikaci, soubor s příslušnou aplikací k otevření, klávesovou zkratku
- automatické vytvoření telefonního seznamu Technické univerzity v Liberci
- tvorba slovníku z dat v souboru
- záměna slov u promluvy
- editace Co se má říkat, úplná změna výrazu, vymazání výrazu

Na obrázku A.6 je zobrazeno okno Editoru slovníku a jednotlivých funkcí menu.



Obr. A.6: Editor slovníku a zobrazení jednotlivých funkcí menu

Přidání výrazu do aktuálního slovníku

Přidání výrazu (povelu) a odpovídající promluvy do slovníku se provede kliknutím na Přidání slova v tabulce 1 → v tabulce 2 vyberte možnost Přidat slovo →

do volných políček vložte promluvu a odpověď → klikněte na tlačítko Přidat výraz pro přidání do slovníku.

Přidání WWW stránky do aktuálního slovníku

Přidání internetové stránky a odpovídající promluvy do slovníku se provede kliknutím na Přidání slova v tabulce 1 → v tabulce 2 vyberte možnost Přidat WWW stránku → do volných políček zadejte promluvu a internetovou stránku ke spuštění (tvar stránky www.domena.cz nebo <http://www.domena.cz>) → klikněte na tlačítko Přidat WWW adresu pro přidání do slovníku.

Přidání aplikace do aktuálního slovníku

Přidání aplikace a odpovídající promluvy do slovníku se provede kliknutím na Přidání slova v tabulce 1 → v tabulce 2 vyberte možnost Přidat aplikaci nebo v Editoru slovníku v menu kliknout na Funkce → Přidat aplikaci, zobrazí se okno Vyberte aplikaci, vyberte spustitelný soubor (*.exe) a cesta k danému souboru se automaticky vloží do Co se má psát → nastavte promluvu, která spustí danou aplikaci → kliknutím na tlačítko Přidat aplikaci vložíte výraz do slovníku.

Přidání souboru a příslušné aplikace do aktuálního slovníku

Přidání souboru s příslušnou aplikací a odpovídající promluvy se provede kliknutím na Přidání slova v tabulce 1 → v tabulce 2 vyberte možnost Přidat soubor a příslušnou aplikaci k otevření nebo v Editoru slovníku v menu kliknout na Funkce → Přidat soubor a příslušnou aplikaci k otevření, zobrazí se okno Vyberte aplikaci, vyberte příslušný spustitelný soubor (*.exe), který umí otevřít vkládaný soubor. Poté se zobrazí okno Vyberte soubor k otevření ve vybrané aplikaci (*.*), vyberte soubor, který chcete spouštět, cesty se automaticky vloží do Co se má psát → nastavte promluvu, která spustí danou aplikaci → kliknutím na tlačítko Přidat aplikaci a soubor vložíte výraz do slovníku.

Přidání klávesové zkratky do aktuálního slovníku

Přidání klávesové zkratky a odpovídající promluvy se provede v Editoru slovníku v menu kliknout na Funkce → Přidat klávesovou zkratku. Zobrazí se lišta, zobrazena na obrázku A.7, pro vložení požadované zkratky → pomocí zaškrtnutí políček si vyberte zkratky CTRL, ALT či Shift a z rolovacího menu příslušnou klávesu,

výběr potvrdíte tlačítkem OK → klávesová zkratka se zapíše do pole Co se má psát → nastavte promluvu pro danou zkratku → do slovníku se výraz vloží tlačítkem Přidat výraz.

Klávesová zkratka se vždy vykonává v aktivním okně. Ovládat je možné libovolné okno.



Obr.A.7: Vkládání klávesové zkratky do slovníku

Automatické vytvoření telefonního seznamu

V Editoru slovníku v menu klikněte na Funkce → Telefonní seznam TUL → Automatické vytvoření tel. seznamu. Dojde k automatickému vytvoření slovníku s telefonním seznamem pracovníků Technické univerzity v Liberci.

Pro vytvoření tohoto slovníku je potřeba mít přesnou strukturu souboru. První sloupec je pro tituly, druhý sloupec je příjmení a jméno včetně vědeckých titulů, třetí sloupec je telefonní číslo a poslední sloupec určuje pracoviště pracovníka. Struktura je vidět v tabulce A.1. V případě, že pracovník titul nemá je sloupec vynechán pomocí tabulátoru. Mezery mezi sloupci jsou tvořeny tabulátorem.

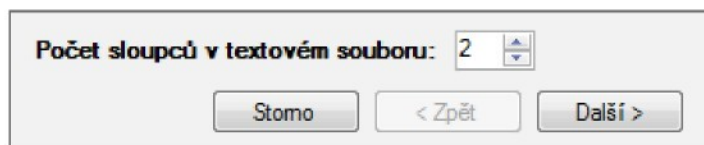
Tab.A.1: Struktura textového sloupce pro automatickou tvorbu slovníku

1.sloupec	2.sloupec	3.sloupec	4.sloupec
tituly	příjmení a jméno	tel. číslo	pracoviště
Ing.	Holada Miroslav, Ph.D.	3080	Ústav ITE
Ing.	Chaloupka Josef, Ph.D.	3099	Ústav ITE
prof. Ing.	Konopa Vojtěch, CSc.	3483	Ústav MTI
	Pánková Věra	3429	Děkanát fakulty mechatroniky

Tvorba slovníku z dat v textovém souboru

V Editoru slovníku v menu klikněte na Funkce → Tvorba slovníku z dat v souboru. Pro vytvoření tohoto slovníku je potřeba mít přesnou strukturu souboru. Počet sloupců musí být minimálně jeden a maximálně deset. V případě, že v nějakém sloupci není žádný text, tak je potřeba vložit prázdný znak (mezeru). Mezery mezi sloupci jsou tvořeny tabulátorem.

Nejprve se zobrazí okno, kde se zadává počet sloupců v textovém souboru. Okno je zobrazeno na obrázku A.8.

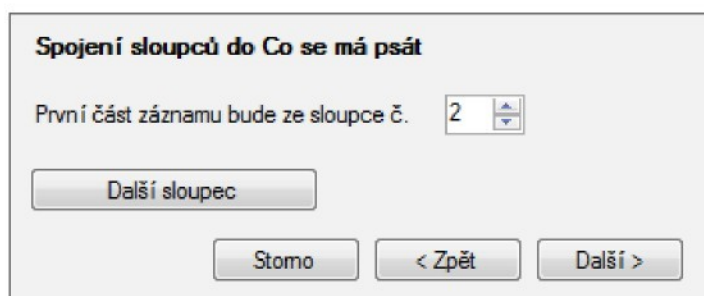


Počet sloupců v textovém souboru: 2

Storno < Zpět Další >

Obr.A.8: Počet sloupců v textovém souboru sloužící pro tvorbu slovníku

Následuje okno, kde se zadávají pozice sloupců, které se mají spojit do Co se má psát. Okno je zobrazeno na obrázku A.9.



Spojení sloupců do Co se má psát

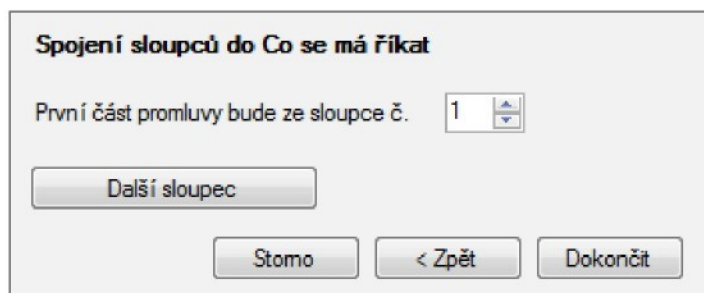
První část záznamu bude ze sloupce č. 2

Další sloupec

Storno < Zpět Další >

Obr.A.9: Okno pro spojení odpovědi z daných sloupců v textovém souboru

V posledním panelu se zadávají pozice sloupců, které vytvoří promluvu Co se má říkat. Okna jsou zobrazeno na obrázku A.10.



Spojení sloupců do Co se má říkat

První část promluvy bude ze sloupce č. 1

Další sloupec

Storno < Zpět Dokončit

Obr.A.10: Okno pro spojení promluvy za daných sloupců v textovém souboru

Počet spojovaných sloupců v Co se má psát nebo Co se má říkat může být také maximálně deset.

Editace Co se má říkat

Pro změnu promluvy je potřeba ze seznamu slovníku vybrat výraz → v tabulce 1 vyberte Editace Co říkat → zadejte novou promluvu v daném políčku → editovaný výraz vložíte tlačítkem Editace výrazu.

Úplná změna výrazu

Pro úplnou změnu výrazu je potřeba ze seznamu slovníku vybrat výraz → v tabulce 1 vyberte Úplná změna výrazu → zadejte nové informace daného výrazu → změněný výraz vložíte tlačítkem Úplná změna výrazu.

Aby nebyla porušena správná struktura slovníku, je vhodnější výraz smazat a zadat opětovně.

Vymazání výrazu ze slovníku

Pokud chcete smazat výraz, musíte ze seznamu slovníku vybrat výraz → klikněte na tlačítko Vymazání výrazu nebo stiskněte Delete.

Záměna slov u promluvy

Pokud chcete u promluvy, která obsahuje dvě slova, zaměnit pořadí, musíte ze seznamu slovníku vybrat výraz → klikněte na tlačítko Záměna slov u promluvy. Následně dojde k vytvoření nového záznamu s promluvou složenou z přehozených slov.

Připojení slovníku

Pro spojení dvou slovníků nebo vytvořených výrazů a již vytvořeného slovníku můžete provést pomocí menu Editoru slovníku, vyberte Soubor → Připojení slovníku. Zobrazí se okno a vyberte slovník, který chcete připojit.

Skryté funkce Komunikátoru Sluha

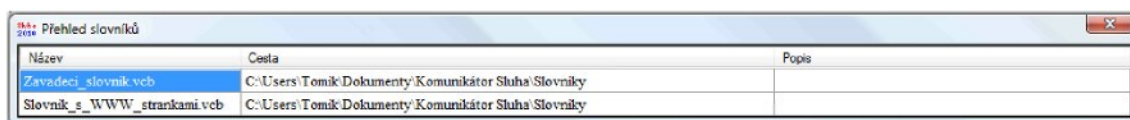
Jednotlivé skryté funkce se přidávají automaticky do všech slovníků vytvořených pomocí Editoru slovníku.

- 1. přidání WWW stránky** – můžete rychle přidávat internetové stránky bez použití Editoru. Internetovou adresu načtete do clipboardu (CTRL + C) → využijte promluvy Přidat stránku → zobrazí se okno Přidat WWW stránku do slovníku, kde je stránka vložena → do pole Co se má říkat stačí zadat promluvu k otevření stránky. V případě chybně načtené stránky existuje možnost stornování akce tlačítkem Storno. V opačném případě klikněte na Přidej WWW stránku a stránka se vloží do slovníku, kam se automaticky přidávané stránky ukládají. V úvodu se jedná o slovník Slovník_s_WWW_strankami.vcb. Soubor je v adresáři C:\Documents and Settings\User name\Dokumenty\Komunikátor Sluha\Slovníky.

2. **zobrazení aktuálního slovníku** - pro zobrazení slovníku zadejte promluvu Zobrazení slovníku.
3. **editace aktuálního slovníku** – pro spuštění Editoru slovníku zadejte promluvu Editace slovníku.

Přehled vytvořených slovníků

V hlavním menu vyberte Editace → Přehled slovníků. Zobrazí se okno, kde jsou uvedeny všechny vámi používané slovníky. Jedná se o slovníky z dostupných adresářů uvedených v inicializačním souboru a také slovníky vytvořené pomocí Editoru slovníku, které jsou uloženy do jiných adresářů. Přehled slovníků je zobrazen na obrázku A.11.



Název	Cesta	Popis
Zavadeči_slovník.vcb	C:\Users\Tomik\Documents\Komunikátor Sluha\Slovníky	
Slovník_s_WWW_strankami.vcb	C:\Users\Tomik\Documents\Komunikátor Sluha\Slovníky	

Obr.A.11: Přehled používaných slovníků

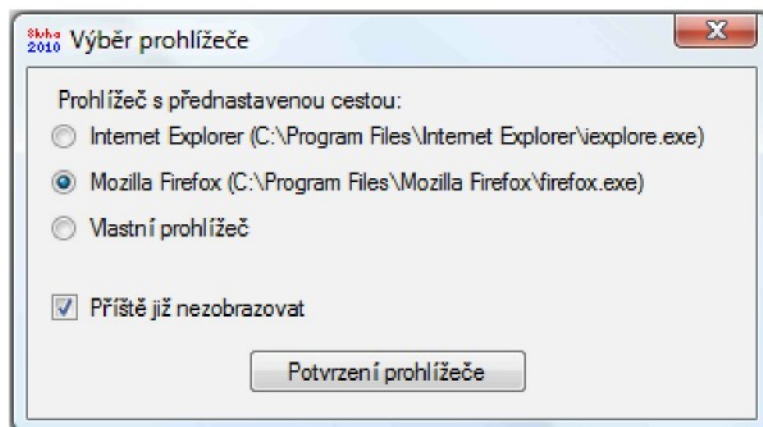
U každého slovníku můžete zadat detailnější popis pro větší orientaci mezi slovníky. Popis nastavíte pomocí třetího sloupce. Po napsání popisku stisknete tlačítko Enter pro zaznamenání popisku. Zobrazení výrazů uložených ve slovníku můžete provést pomocí dvojkliku na název slovníku (v 1.sloupci). K odstranění slovníku z Přehledu slovníků a také i počítače můžete použít tlačítko Delete. Vždy je vymazán slovník, který je označen podle řádku nebo daného políčka.

Výběr prohlížeče

Pro zobrazení okna Výběr prohlížeče v hlavním menu vyberte Možnosti → Výběr prohlížeče. Pro zobrazení internetové adresy můžete použít jeden ze dvou přednastavených prohlížečů - Internet Explorer nebo Mozilla Firefox či vybrat vlastní prohlížeč pomocí dialogu, kde vyberete cestu k danému prohlížeči.

Zaškrtnutí políčka Příště již nezobrazovat nastavíte jeden z prohlížečů jako hlavní a otevírané stránky se budou automaticky otevírat v tomto prohlížeči.

Okno výběru prohlížeče je zobrazeno na obrázku A.12.



Obr.A.12: Výběr prohlížeče internetových stránek

Doba zobrazení odpovědi

V případě potřeby delší či kratší doby zobrazení odpovědi můžete nastavit danou hodnotu. Vyberte v menu hlavního panelu Možnosti → Doba zobrazení odpovědi. Dojde k zobrazení možnosti nastavit hodnotu od jedné sekundy do půl minuty. Možnost nastavení je vidět na obrázku A.13.



Obr.A.13: Nastavení doby zobrazení odpovědi

Nastavení barvy

1. **barva při nahrávání** - v úvodu je nastavena barva zelená, pomocí menu ovládací ikony můžete nastavit požadovanou barvu výběrem z palety barev.
2. **barva bez nahrávání** - v úvodu je nastavena barva červená, pomocí menu ovládací ikony můžete nastavit požadovanou barvu výběrem z palety barev.
3. **barva odpovědi** - v úvodu je nastavena barva tyrkysová, pomocí menu ovládací ikony můžete nastavit požadovanou barvu výběrem z palety barev.
4. **barva odpovědi** - v úvodu je nastavena barva černá, pomocí menu ovládací ikony můžete nastavit požadovanou barvu výběrem z palety barev.

Příloha B – obsah přiloženého DVD

- Text diplomové práce ve formátu PDF
- Instalace klienta
- Instalace konfigurace rozpoznávacího systému
- Webové stránky
- Využívané datové soubory – slovníky, telefonní seznam osob pracujících na Technické univerzitě v Liberci, soubor s databázovým zaznamenáním informací









Fakulta mechatroniky,
informatiky
a mezinárodních studií,
Technická univerzita v Liberci



Laborator počítačového
zpracování řeči



Ústav Informatických
technologií a elektroniky



Fakulta mechatroniky,
informatiky
a mezinárodních studií,
Technická univerzita v Liberci

Laborator počítačového
zpracování řeči

Ústav Informačních
technologií a elektroniky





Aktualizace slovníku

Kliknutím na toto tlačítko dojde k aktualizaci slovníku v rozpoznávacím systému. Nutné použít v případě zadání nových výrazů a povelů do aktuálního slovníku, který je načten v rozpoznávacím systému. Dojde k aktualizaci slovníku a bude možné nové promluvy již využívat.

Automatické načtení slovníku

Při spuštění aplikace se automaticky načte slovník, který je určen v inicializačním souboru *Nastaveni.ini*. Do rozpoznávacího systému je v úvodu automaticky načten slovník z adresáře *C:\Documents and Settings\User name\Dokumenty\Komunikátor Sluha\Slovníky\Zavadeći_slovník.vcb* a dojde k zobrazení ikony pro nahrávání a nápisu „*Slovník načten*“.

V případě, že zaváděcí slovník není k dispozici, dojde k otevření okna pro výběr slovníku, který se načte do rozpoznávacího systému. Slovník musí mít příponu *.vcb. Po správném načtení slovníku se zobrazí ikona pro nahrávání a nápis „*Slovník načten*“.

Co je Komunikátor Sluha

Komunikátor Sluha je výkonný nástroj, který umožňuje uživateli komunikovat a ovládat počítač pomocí hlasu.

Komunikátor Sluha poskytuje:

- ovládání počítače pomocí Vámi vytvořeného slovníku
- možnost hlasem spustit jakoukoliv internetovou adresu
- možnost hlasem spustit jakoukoliv aplikaci
- možnost hlasem spustit jakýkoliv soubor s odpovídajícím programem
- hlasem ovládat jakékoliv aktivní okno pomocí klávesových zkratk
- možnost hlasem vyzjistit jakoukoliv informaci z databáze
- možnost hlasem vyzjistit jakékoliv telefonní číslo pracovníka Technické univerzity v Liberci.

Doba zobrazení odpovědi

V případě potřeby delší či kratší doby zobrazení odpovědi můžete nastavit danou hodnotu. Vyberte v menu Komunikátoru Sluha Možnosti → Doba zobrazení odpovědi. Dojde k zobrazení možnosti nastavit hodnotu od jedné sekundy do půl minuty.

Editace Co se má říkat

Pokud je vložený výraz chybný, je možné provést změnu promluvy - Co se má říkat. Daný výraz je zapotřebí označit v Editoru a následně musí být vybrána možnost Editace Co říkat. V tomto případě je možné měnit pouze promluvu, aby nedošlo k nesprávnému tvaru jednotlivých funkcí a tím i špatnému tvaru pro rozpoznávací systém. Do slovníku změněný výraz vložíme tlačítkem Editace výrazu.

Editor slovníku

Možnost vytvářet a editovat slovníky.

Dostupné funkce:

- nový slovník, otevření slovníku, připojení dalšího slovníku k již vytvořenému, uložení a uložení jako
- přidat výraz, WWW stránku, aplikaci, soubor s příslušnou aplikací k otevření, klávesovou zkratku
- automatické vytvoření telefonního seznamu Technické univerzity v Liberci
- tvorba slovníku z dat v souboru
- záměna slov u promluvy
- editace Co se má říkat, úplná změna výrazu, vymazání výrazu

Nahrávání rozpoznávacího systému + nahrávací ikona

Po načtení slovníku do rozpoznávacího systému viz. Automatický start nebo Načtení slovníku / Konec komunikace je zobrazena nahrávací ikona v úvodu orámovaná červeně - režim nenahrávání (změna viz. Nastavení barvy). Při najetí myši na tuto ikonu a ponechání myši alespoň 0,5s se spustí nahrávání, které je v úvodu znázorněno zeleným orámováním - režim nahrávání (změna viz. Nastavení barvy). Nahrávání trvá po dobu 3s a poté je nahrávání opět zastaveno a je nutné znovu na ikonu najet. Pouhé přejetí ikony (pokud není delší jak 0,5s) nahrávání nespustí.

Ikona lze pomocí levého tlačítka myši přesouvat po ploše a pomocí pravého tlačítka se zobrazí menu, které slouží k nastavení nahrávací ikony případně přepnutí do jiného okna. V obou uvedených případech je nahrávání zastaveno.

Načtení slovníku / Konec komunikace

V případě, že tlačítko je označeno textem Načtení slovníku, je možné načíst libovolný slovník do rozpoznávacího systému. Po kliknutí se zobrazí okno Výběr slovníku pro rozpoznávací systém, kde je potřeba vybrat požadovaný slovník. Slovník musí mít příponu *.vcb a také správnou strukturu jednotlivých řádků, na které bude rozpoznávací systém reagovat. Pro správnou strukturu slovníku je vhodné použít Editor slovníku (viz. Tvorba nebo editace slovníku). Po načtení slovníku do rozpoznávacího systému se zobrazí ikona pro nahrávání a nápis „*Slovník načten*“.

V případě, že je tlačítko označeno textem Konec komunikace, dojde k zastavení možnosti komunikovat s počítačem a dojde k odstranění slovníku z rozpoznávacího systému. Pro načtení nového slovníku je nutné kliknout na tlačítko Načtení slovníku (viz. výše).

Nastavení barvy

1. **barva při nahrávání** - v úvodu je nastavena barva zelená, je možné barvu pomocí palety barev nastavit na vlastní
2. **barva bez nahrávání** - v úvodu je nastavena barva červená, je možné barvu pomocí palety barev nastavit na vlastní
3. **barva odpovědi** - v úvodu je nastavena barva tyrkysová, je možné barvu pomocí palety barev nastavit na vlastní
4. **barva textu odpovědi** - v úvodu je nastavena barva černá, je možné barvu pomocí palety barev nastavit na vlastní

Neprůhlednost nahrávací ikony

- 100%
- 80%
- 60%
- 40%
- 20%

Poděkování

Děkuji svému vedoucímu diplomové práce Ing. Miroslavu Holadovi, Ph.D. za užitečné konzultace, rady, připomínky k této práci a také za jeho trpělivost a především pak za poskytnutí již hotového programu k rozpoznávacímu systému Dundis.

Přehled vytvořených slovníků

Pro zobrazení slovníků z dostupných adresářů, které jsou uvedeny v inicializačním souboru `Nastaveni.ini` plus slovníky vytvořené pomocí Editoru slovníků (viz. Tvorba nebo editace slovníku).

U každého slovníku je možný detailnější popis, aby bylo ihned jasné, o jaký slovník se jedná. Popis je možné nastavit v okně **Přehled slovníků** ve třetím sloupci. Po napsání popisku a stisknutí tlačítka **Enter** je popisek zaznamenán.

Zobrazení výrazů ve slovníku je možné učinit pomocí dvojkliku na název slovníku (v 1.sloupci).

K odstranění slovníku z Přehledu slovníků a také i počítače je možné použít tlačítko **Delete**. Vždy je vymazán slovník, který je označen podle řádku nebo daného políčka.

Záměna u promluvy

Daný výraz je zapotřebí označit v Editoru a v případě, že v poli Co se má říkat jsou dvě slova, pomocí daného tlačítka je možné části slova zaměnit. Dojde k vytvoření nového záznamu s promluvou složenou z přehozených slov.

Přidání WWW stránky do aktuálního slovníku

Přidání internetové stránky a odpovídající promluvy se provede tak, že musí být vybrána možnost Přidání slova a Přidat WWW stránku. Do políčka Co se má psát se vloží odpovídající internetová stránka a do Co se má říkat promluva na zobrazení stránky a do slovníku se výraz vloží tlačítkem Přidat WWW adresu.

Od prvního přidání výrazu jsou aktivní možnosti Nový, Připojení slovníku, Uložit a Uložit jako.

Přidání aplikace do aktuálního slovníku

Přidání aplikace a odpovídající promluvy se provede tak, že se klikne na Přidat aplikaci a zobrazí se okno Vyberte aplikaci, kterou chcete přidat do slovníku. Vybere se spustitelný soubor (*.exe) a cesta k danému souboru se automaticky vloží do Co se má psát. Nyní se nastaví promluva Co se má říkat, která spustí danou aplikaci. Do slovníku se výraz vloží tlačítkem Přidat aplikaci.

Od prvního přidání výrazu jsou aktivní možnosti Nový, Připojení slovníku, Uložit a Uložit jako.

Přidání klávesové zkratky do aktuálního slovníku

Přidání klávesové zkratky a odpovídající promluvy se provede kliknutím na Přidat klávesovou zkratku. Pomocí zaškrťovacích políček se vyberou zkratky CTRL, ALT či Shift a z rolovacího menu se vybere příslušná klávesa. Pomocí tlačítka OK se klávesová zkratka zapíše do pole Co se má psát. Nyní se nastaví promluva Co se má říkat, která vyvolá danou klávesovou zkratku. Do slovníku se výraz vloží tlačítkem Přidat výraz.

Klávesová zkratka se provede vždy v aktivním okně.

Od prvního přidání výrazu jsou aktivní možnosti Nový, Připojení slovníku, Uložit a Uložit jako.

Přidání souboru a příslušné aplikace do aktuálního slovníku

Přidání souboru s příslušnou aplikací a odpovídající promluvy se provede tak, že se klikne na Přidat soubor a příslušnou aplikaci k otevření. Zobrazí se okno Vyberte aplikaci, která bude sloužit k otevření souboru. Vybere se příslušný spustitelný soubor (*.exe). Poté se zobrazí okno Vyberte soubor k otevření ve vybrané aplikaci (*.*) a dané cesty se automaticky vloží do Co se má psát. Nyní se nastaví promluva Co se má říkat, která spustí daný soubor. Do slovníku se výraz vloží tlačítkem Přidat aplikaci a soubor.

Od prvního přidání výrazu jsou aktivní možnosti Nový, Připojení slovníku, Uložit a Uložit jako.

Přidání výrazu do aktuálního slovníku

Přidání výrazu (povelu) a odpovídající promluvy se provede vybráním možnosti Přidání slova a Přidat slovo. Do volných políček se napíše - Co se má psát a Co se má říkat a do slovníku se výraz vloží tlačítkem Přidat výraz.

Od prvního přidání výrazu jsou aktivní možnosti Nový, Připojení slovníku, Uložit a Uložit jako.

Připojení slovníku

Tato možnost umožňuje spojení více slovníků v jeden nebo také možnost vytvořit si vlastní slovník pomocí editoru a připojit k němu již vytvořený slovník, který nemá všechny potřebné výrazy.

Skryté funkce Komunikátoru Sluha

Jednotlivé skryté funkce se přidávají automaticky do všech slovníků vytvořených pomocí Editoru, tudíž je možnost jich využít pokaždé, kdy je rozpoznávací systém spuštěn.

1. **přidání WWW stránky** - rychlé přidání internetové stránky bez použití Editoru. Je nutné požadovanou internetovou adresu načíst do clipboardu (CTRL + C) a pomocí nahrávání rozpoznávacího systému a promluvy Přidat stránku se zobrazí okno Přidat WWW stránku do slovníku, kde je stránka vložena. Do pole Co se má říkat stačí zadat promluvu k otevření stránky. V případě chybně načtené stránky existuje možnost stornování akce tlačítkem Storno. V opačném případě se klikne na Přidej WWW stránku a stránka se vloží do slovníku, kam se automaticky přidávané stránky ukládají. V úvodu se jedná o slovník Slovník_s_WWW_strankami.vcb. Soubor je v adresáři C:\Documents and Settings\User name\Dokumenty\Komunikátor Sluha\Slovníky.
2. **zobrazení aktuálního slovníku** - pro zobrazení slovníku v Prohlížeči slovníku a možnosti projít si možné promluvy z aktuálního slovníku načteného do rozpoznávacího systému. Potřeba je nahrávání rozpoznávacího systému a promluva Zobrazení slovníku.
3. **editace aktuálního slovníku** - pro editaci aktuálního slovníku načteného do rozpoznávacího systému. Potřeba je nahrávání rozpoznávacího systému a promluva Editace slovníku.

Tvorba slovníku z dat v textovém souboru

Pro vytvoření tohoto slovníku je potřeba mít přesnou strukturu souboru. Počet sloupců musí být minimálně jeden a maximálně deset. V případě, že v nějakém sloupci není žádný text, tak je potřeba vložit prázdný znak (mezeru). Mezery mezi sloupci jsou tvořeny tabulátorem.

Nejprve se zobrazí okno, kde se zadává počet sloupců v textovém souboru. Následuje okno, kde se zadávají pozice sloupců, které se mají spojit do Co se má psát a v posledním panelu se zadávají pozice sloupců, které vytvoří promluvu Co se má říkat.

Počet spojovaných sloupců v Co se má psát nebo Co se má říkat může být také maximálně deset.

Úplná změna výrazu

Pokud je vložený výraz chybný, tak je možné provést úplnou změnu promluvy. Daný výraz je zapotřebí označit v Editoru a následně musí být vybrána možnost Úplná změna výrazu. Nyní je možné měnit, jak Co se má psát, tak i Co se má říkat. Ovšem v takém případě je jednodušší a jistější výraz smazat (viz. Vymazání výrazu) a zadat znovu, aby omylem nedošlo k nesprávnému tvaru nějaké funkce a tím i špatnému tvaru pro rozpoznávací systém.

Velikost nahrávací ikony

- malá
- střední
- velká

Vymazání výrazu ze slovníku

Pokud je vložený výraz chybný, je možné jej celý smazat. Daný výraz je zapotřebí označit v Editoru a následně buď pomocí tlačítka Vymazání výrazu, který je přímo v Editoru, nebo pomocí tlačítka Delete vymazat.

Automatické vytvoření telefonního seznamu

Dojde k automatickému vytvoření slovníku s telefonním seznamem pracovníků Technické univerzity v Liberci. Do Co se má psát je vložen nejprve titul, jméno pracovníka a telefonní číslo a do Co se má říkat se vloží příjmení a jméno pracovníka.

Pro vytvoření tohoto slovníku je potřeba mít přesnou strukturu souboru. První sloupec je pro tituly, druhý sloupec je příjmení a jméno včetně vědeckých titulů, třetí sloupec je telefonní číslo a poslední sloupec určuje pracoviště pracovníka.

V případě, že pracovník titul nemá je sloupec vynechán pomocí tabulátoru. Mezery mezi sloupci jsou tvořeny tabulátorem.

Výběr prohlížeče

Pro zobrazení internetové adresy je možné použít jeden ze dvou přednastavených prohlížečů - Internet Explorer nebo Mozilla Firefox či vybrat vlastní prohlížeč pomocí dialogu a vybrání spouštěcího souboru nového prohlížeče internetových stránek.

Pomocí zaškrtnutí políčka Příště již nezobrazovat je možné nastavit jeden z prohlížečů jako hlavní a otevírané stránky se budou automaticky otevírat ve vybraném prohlížeči.

Zobrazení výpisu událostí

V tomto okně je zobrazen výpis všech úkonů, které byly vykonány a provedeny - např. načtení slovníku, spuštění aplikací, internetových stránek nebo vypsání promluvy, která se zobrazila na obrazovce.

Úvodní nastavení Komunikátoru Sluha

Při spuštění aplikace dojde k vytvoření inicializačního souboru *Nastaveni.ini*. Soubor je uložen v adresáři *C:\Documents and Settings\User name\Dokumenty\Komunikátor Sluha*. Soubor obsahuje nutné informace k nastavení Komunikátoru Sluha.

Potřebné parametry, které musí být v souboru zaznamenány:

- adresář programu
- cesta hlavního slovníku (adresář s prioritními slovníky)
- cesta k souboru s přehledem slovníků
- cesta zaváděcího slovníku
- cesta slovníku s WWW stránkami
- cesta k souboru s telefonním seznamem
- adresa rozpoznávacího systému

Další parametry jsou v úvodu a v případě chybějící informace nastaveny na:

- barva při nahrávání – Lime
- barva bez nahrávání – Red
- barva odpovědi – Cyan
- barva textu odpovědi – Black
- doba zobrazení odpovědi – 3s
- neprůhlednost – 1.0
- velikost – malý
- pozice – 50 x 50

Tyto parametry nemusí být nastaveny:

- prohlížeč internetových stránek
- uložení slovníku