

**Západočeská univerzita v Plzni**

**Fakulta filozofická**

**Diplomová práce**

**Komparativní lingvistická analýza paralelních  
komunikátů vytvářených stínovým řečníkem**

**Jaroslava Nováková**

Plzeň 2014

**Západočeská univerzita v Plzni**

**Fakulta filozofická**

Katedra filozofie

**Studijní program Humanitní studia**

**Studijní obor Teorie a filozofie komunikace**

**Diplomová práce**

**Komparativní lingvistická analýza paralelních  
komunikátů vytvářených stínovým řečníkem**

**Jaroslava Nováková**

*Vedoucí práce:*

PhDr. Martin Prošek, Ph.D.

Katedra českého jazyka a literatury

Fakulta pedagogická Západočeské univerzity v Plzni

Plzeň 2014

Prohlašuji, že jsem práci zpracovala samostatně a použila jen uvedených pramenů a literatury.

*Plzeň, duben 2014*

.....

Děkuji PhDr. Martinu Proškovi, Ph.D. za odborné rady a vedení mé diplomové práce. Mé poděkování patří i Ing. Aleši Pražákovi, Ph.D. za poskytnutí dat nutných pro lingvistickou analýzu. V neposlední řadě patří poděkování i mému kolegovi Bc. Václavu Jindřichovi, bez jehož podnětných rad by nevznikla jedna z kapitol této diplomové práce.

## OBSAH

<b>1</b>	<b>ÚVOD .....</b>	<b>1</b>
<b>2</b>	<b>STÍNOVÝ MLUVČÍ .....</b>	<b>3</b>
<b>3</b>	<b>ZPŘÍSTUPŇOVÁNÍ POŘADŮ NESLYŠÍCÍM .....</b>	<b>4</b>
<b>3.1</b>	<b>Definice pojmu „titulky“ a jejich klasifikace.....</b>	<b>4</b>
3.1.1	Způsoby zobrazování skrytých titulků.....	6
<b>3.2</b>	<b>Mediální zákon v České republice .....</b>	<b>7</b>
<b>3.3</b>	<b>Uživatelé skrytých titulků a jejich požadavky.....</b>	<b>7</b>
3.3.1	Neslyšící osoby.....	8
3.3.2	Osoby ohluchlé po osvojení si mluveného jazyka .....	9
3.3.3	Osoby nedoslýchavé .....	9
3.3.4	Osoby slyšící .....	10
<b>3.4</b>	<b>Zpřístupňování pořadů neslyšícím v České televizi .....</b>	<b>10</b>
3.4.1	Skryté titulky .....	10
3.4.1.1	Přímé titulkování .....	11
3.4.1.2	Stínové přemlouvání .....	13
3.4.2	Pořady přemlouvané stínovými řečníky.....	14
3.4.2.1	Otázky Václava Moravce .....	15
3.4.2.2	Hyde Park .....	15
3.4.2.3	Ekonomika ČT24 .....	16
3.4.2.4	Pořad 168 hodin.....	16
3.4.2.5	Interview Daniely Drtinové .....	17
3.4.2.6	Další pořady .....	17
3.4.3	Statistiky .....	18
<b>4</b>	<b>SYSTÉM AUTOMATICKÉHO ROZPOZNÁVÁNÍ ŘEČI.....</b>	<b>20</b>

<b>4.1 Akustický model.....</b>	<b>22</b>
<b>4.2 Jazykový model.....</b>	<b>23</b>
4.2.1 Trénování jazykového modelu.....	24
4.2.1.1 Zdroje dat pro tvorbu jazykových modelů .....	24
<b>4.3 Dekodér.....</b>	<b>26</b>
<b>4.4 Rozpoznávací komplet .....</b>	<b>27</b>
<b>5 TRÉNINK STÍNOVÝCH MLUVČÍCH .....</b>	<b>28</b>
<b>5.1 Úkony prováděné stínovým mluvčím během diktování .....</b>	<b>28</b>
<b>5.2 Způsoby a specifika stínového přemlouvání.....</b>	<b>30</b>
<b>5.3 Charakteristika komunikátu vytvářeného stínovým mluvčím</b>	<b>32</b>
<b>5.4 Trenažér stínových mluvčích.....</b>	<b>36</b>
<b>5.5 První fáze tréninku .....</b>	<b>38</b>
<b>5.6 Druhá fáze tréninku.....</b>	<b>39</b>
<b>5.7 Třetí fáze tréninku .....</b>	<b>41</b>
<b>5.8 Čtvrtá fáze tréninku.....</b>	<b>44</b>
<b>5.9 Stínové přemlouvání živého vysílání.....</b>	<b>45</b>
<b>6 HODNOCENÍ VÝSLEDNÝCH TITULKŮ.....</b>	<b>48</b>
<b>6.1 Hodnocení rozpoznávání řeči .....</b>	<b>48</b>
<b>6.2 Hodnocení stínového přemlouvání na KKY ZČU .....</b>	<b>48</b>
6.2.1 První úroveň hodnocení.....	49
6.2.2 Druhá úroveň hodnocení .....	50
6.2.3 Třetí úroveň hodnocení .....	50

<b>7 KOMPATIVNÍ LINGVISTICKÁ ANALÝZA KOMUNIKÁTŮ VYTVÁŘENÝCH STÍNOVÝMI MLUVČÍMI.....</b>	<b>52</b>
<b>7.1 Metodika výzkumu .....</b>	<b>52</b>
<b>7.2 Analýza komunikátů vytvořených ve druhé fázi tréninku .....</b>	<b>54</b>
7.2.1 Komunikáty vytvořené stínovou mluvčí DT.....	54
7.2.1.1 Nahrávky zpomalené na 70 % rychlosti původního vysílání .....	55
7.2.1.2 Nahrávka přemluvená v plné rychlosti původního vysílání .....	59
7.2.2 Komunikáty vytvořené stínovou mluvčí DZ.....	62
7.2.2.1 Nahrávka zpomalená na 70 % rychlosti původního vysílání .....	63
7.2.2.2 Nahrávky přemluvené v plné rychlosti původního vysílání .....	64
7.2.3 Komunikáty vytvořené stínovým mluvčím PP.....	68
7.2.3.1 Nahrávky zpomalené na 70 % rychlosti původního vysílání .....	68
7.2.3.2 Nahrávky přemluvené v plné rychlosti původního vysílání .....	74
<b>7.3 Shrnutí chyb zjištěných z analýzy komunikátů vyhotovených ve druhé fázi tréninku .....</b>	<b>76</b>
7.3.1 Nadindividuální chyby vnášené do systému stínovými mluvčími .....	76
<b>7.4 Analýza komunikátů vytvořených ve čtvrté fázi tréninku.....</b>	<b>80</b>
7.4.1 Komunikát vytvořený stínovou mluvčí DT.....	81
7.4.2 Komunikát vytvořený stínovou mluvčí DZ.....	87
7.4.3 Komunikát vytvořený stínovým mluvčím PP.....	89
<b>7.5 Shrnutí nadindividuálních chyb zjištěných v komunikátech vyhotovených ve čtvrté fázi tréninku.....</b>	<b>94</b>

<b>7.6 Komunikáty vytvořené v živém vysílání.....</b>	<b>95</b>
7.6.1 Komunikát vytvořený stínovou mluvčí DT.....	95
7.6.2 Komunikát vytvořený stínovou mluvčí DZ.....	101
7.6.3 Komunikát vytvořený stínovým mluvčím PP.....	105
<b>7.7 Shrnutí nadindividuálních chyb zjištěných v komunikátech vyhotovených v živém vysílání.....</b>	<b>113</b>
<b>7.8 Chyby systému ovlivňující analyzované komunikáty .....</b>	<b>113</b>
<b>7.9 Klasifikace zjištěných chyb.....</b>	<b>116</b>
7.9.1 Závažné chyby.....	116
7.9.2 Standardní chyby.....	117
7.9.3 Chyby téměř neovlivňující pochopení výsledného titulku .	117
<b>8 ZÁVĚR .....</b>	<b>118</b>
<b>9 SEZNAM POUŽITÉ LITERATURY A ZDROJŮ .....</b>	<b>122</b>
<b>10 RESUMÉ .....</b>	<b>125</b>
<b>11 PŘÍLOHY.....</b>	<b>126</b>



## 1 ÚVOD

Tato diplomová práce si klade za cíl představit koncept stínového přemlouvání jako metody opatřování živě vysílaných pořadů skrytými titulky. Jedná se o nový přístup tvorby skrytých titulků za pomoci systému automatického rozpoznávání řeči, který je vyvíjen v posledních několika letech na mnoha odborných pracovištích po celém světě. Systém rozpoznávání řeči je během procesu ovládán odborným pracovníkem, tzv. stínovým mluvčím, který je zodpovědný za výslednou formu skrytého titulku zobrazeného jeho uživateli, především divákům s různými formami poruch sluchu. Cílem této práce je pak charakterizovat komunikát vytvořený stínovým mluvčím během procesu přemlouvání a za pomoci komparativní analýzy zjistit nedostatky vnášené do procesu rozpoznávání stínovými mluvčími a systémem rozpoznávání řeči.

Analyzovány budou komunikáty vytvořené třemi stínovými mluvčími, kteří prošli všemi fázemi tréninku a v současné době opatřují skrytými titulky vybrané živě vysílané pořady České televize. Data pro analýzu byla poskytnuta pracovníky Katedry kybernetiky Západočeské univerzity. Analýza bude prováděna komparativní metodou za pomoci počítačového programu MegaWord. Zjišťovány budou chyby vnášené do procesu rozpoznávání stínovými mluvčími a samotným systémem rozpoznávání řeči. Jednotlivé chyby budou demonstrovány na příkladech a rozděleny do tří kategorií: chyby individuální, chyby nadindividuální a chyby systému rozpoznávání řeči. Všechny tři výše zmíněné kategorie chyb budou následně rozděleny do kategorií podle závažnosti ovlivnění pochopení skrytého titulku jeho uživateli.

Skryté titulky vytvářené za pomoci systému automatického rozpoznávání řeči jsou využívány ve vysílání České televize, která je jako veřejnoprávní subjekt ze zákona povinna opatřit alespoň 70 % svého vysílání skrytými

nebo otevřenými titulky. Aby byla tato zákonná kvóta splněna, využívá Česká televize služeb Katedry kybernetiky Západočeské univerzity, která se aktivně podílí na vývoji systému rozpoznávání řeči a zajišťuje také výcvik kompetentních stínových mluvčích.

Dále tato diplomová práce přibližuje nejčastější způsoby zobrazování skrytých nebo otevřených titulků v televizním vysílání. Protože stínové přemlouvání je metodou užívanou pro tvorbu skrytých titulků, jsou v diplomové práci představeny jednotlivé skupiny uživatelů skrytých titulků včetně jejich požadavků na podobu skrytých titulků, které ale vzhledem k jejich různorodosti není možné vždy naplnit.

V diplomové práci je také představen systém rozpoznávání řeči, který zajišťuje přepis mluveného slova do psaného textu. Tato metoda je velice efektivní, protože výrazně urychluje proces tvorby a zobrazení výsledného titulku jeho uživatelům. Diplomová práce si ovšem neklade za cíl představit systém rozpoznávání řeči po technické stránce.

Protože činnost stínového mluvčího je velice náročná, zejména po psychické stránce, a zahrnuje velký počet činností vykonávaných paralelně v jeden okamžik, byl na Katedře kybernetiky vyvinut Trenažér stínových mluvčích, který umožňuje výcvik adeptů na pozici stínového mluvčího. Trénink je rozdělen do čtyř etap, ve kterých se kandidáti na stínové mluvčí učí postupně celý proces vytváření skrytých titulků od nejjednodušších činností až po fázi poslední, která simuluje tvorbu skrytých titulků v živém vysílání.

Komunikát, který vzniká během stínového přemlouvání, se vyznačuje mnoha specifiky. Jeho výsledná podoba záleží zejména na způsobu stínového přemlouvání, který je specifický pro každého ze sedmi stínových mluvčích, kteří v současné době opatřují skrytými titulky živé vysílání České televize.

Na základě analýzy výsledných komunikátů bude navržena funkčně orientovaná taxonomie zásad tvorby komunikátů, která umožní zefektivnění výkonu stávajících stínových mluvčích i výcvik nových kandidátů.

## **2 STÍNOVÝ MLUVČÍ**

Pro účely této diplomové práce musí být nejdříve vymezen pojem „stínový mluvčí“<sup>1</sup>.

Stínový mluvčí je simultánní tlumočnick z daného jazyka do téhož jazyka, který poslouchá zvukovou stopu živého televizního vysílání a podle potřeby ji zjednodušuje či opravuje takovým způsobem, aby jeho výsledná promluva byla vhodná pro rozpoznávací systém, který stínovým řečníkem vyslovenou promluvu automaticky převede na psané titulky určené primárně pro neslyšící diváky.

V české literatuře není k dispozici obecně uznávaná ucelená definice tohoto pojmu, proto byla výše uvedená definice vytvořena pro účely této diplomové práce.

Na stínového řečníka je během jeho činnosti kladeno množství nároků, které budou zmíněny v další části diplomové práce.

---

<sup>1</sup> Lze užít také označení stínový řečník. V angličtině se užívá pojmů re-speaker nebo shadow-speaker.

### 3 ZPŘÍSTUPŇOVÁNÍ POŘADŮ NESLYŠÍCÍM

V současné době musí televizní stanice ze zákona zpřístupňovat určité procento svého vysílání osobám se zrakovým a sluchovým postižením. Následující kapitola se věnuje pojmu „titulky“, jejich zobrazování divákovi, mediálnímu zákonu, požadavkům neslyšících a přístupu České televize k divákům se specifickými potřebami.

#### 3.1 Definice pojmu „titulky“ a jejich klasifikace

Obecně se termínem „titulky“ označuje text, který je přepisem nebo překladem zvukové stopy pořadu. V televizní a filmové tvorbě lze klasifikovat dva druhy titulků.

Prvním typem jsou tzv. „otevřené titulky“<sup>2</sup>, které jsou určeny všem druhům diváků. Jsou umístěny přímo ve vysílaném obraze bez možnosti vypnutí. Nejčastější formou tohoto typu titulků jsou úvodní a závěrečné titulky<sup>3</sup> pořadu, které obsahují informace o jeho tvůrcích. Závěrečné titulky jsou nejčastěji uváděny na ploše celé obrazovky.

Příkladem otevřených titulků jsou tzv. „titulky překladové“, které slouží zejména jako překlady cizojazyčných dialogů nebo komentáře u dokumentů a pořadů, které jsou uváděny v původním znění. Překladové titulky jsou umístěny přímo ve vysílaném obraze a mění se průběžně spolu s probíhajícím dějem měnícím se obrazem a postupující zvukovou stopou.

---

<sup>2</sup> Také označovány pojmem pevné titulky.

<sup>3</sup> Označovány také jako začáteční a koncové.

Tento druh titulků je primárně určen slyšícím divákům, protože neobsahuje další dodatečné informace nutné pro neslyšícího diváka – například přepis ruchů (skřípání schodů) nebo barevné rozlišení řečníků pro lepší orientaci.

Dále lze za otevřené titulky považovat informativní texty, které oznamují jméno a funkci osoby hovořící na obrazovce nebo místo, kde se odehrává aktuální dění pořadu.

Druhým typem titulků jsou tzv. „skryté titulky“, které jsou určeny zejména pro neslyšící diváky. Zvuková stopa pořadu je převedena do písemné podoby stejně jako u otevřených titulků, současně ale výsledné titulky obsahují i další specifické náležitosti požadované neslyšícími. Sem spadá zejména barevné odlišení promluv jednotlivých mluvčích ve výsledných titulcích, což výrazně usnadňuje orientaci v ději. U předtočených pořadů se skryté titulky vytvářejí po dokončení pořadů. Ve zpravodajství a dalších živě vysílaných pořadech jsou vytvářeny přímo během vysílání.<sup>4</sup> Zásadní rozdíl u titulků vytvářených po dokončení pořadu a u živého vysílání spočívá v načasování výsledných titulků. Zatímco titulky vytvořené po dokončení pořadu jsou načasovány tak, aby přesně korespondovaly s promluvou na obrazovce, titulky vytvářené živě jsou zobrazovány s určitou časovou prodlevou způsobenou přenosem dat a jinými faktory.

---

<sup>4</sup> KOPLÍK, STRNADOVÁ, *Problematika tvorby skrytých titulků k předtočeným pořadům*, s. 15–19.

Skryté titulky nejsou pevnou součástí obrazu. V České republice je lze vyvolat na teletextu televizních stanic na stránkách 888 nebo lze využít i titulky DVB<sup>5</sup>.

### 3.1.1 Způsoby zobrazování skrytých titulků

V současné době existují tři způsoby zobrazování skrytých titulků. Prvním z nich je formát titulků tzv. „rolující po slovech“, který užívá pro zobrazování britská BBC. Titulky se v tomto případě zobrazují slovo po slově, dokud se nezaplní řádka. Jakmile se řádka naplní, nejvyšší řádka zmizí a současně se objeví řádka nová, ve které se titulek objevuje opět slovo po slově. Tyto titulky umožňují vidět předchozí kontext, nicméně pro čtenáře může být obtížné sledovat s titulky zároveň i obraz pořadu.

Eventualitou tohoto způsobu zobrazování jsou pak titulky, které rolují po řádcích. Jejich zobrazování funguje na stejném principu jako u titulků rolujících po slovech, namísto jednotlivých slov se ale objevují celé věty nebo jejich části.<sup>6</sup> Stejně jako titulky rolující po slovech, může i tento způsob zobrazení způsobovat divákovi obtíže sledovat dění na obrazovce a zároveň číst titulky.

Česká televize používá pro zobrazení skrytých titulků formát tzv. blokový, jedno- či dvouřádkový<sup>7</sup>. Tento způsob sice neumožňuje zobrazení

---

<sup>5</sup> DVB titulky jsou obdobou skrytých titulků na teletextových stránkách 888. Jedná se o digitální vysílání skrytých titulků přímo, bez vazby na teletext.

<sup>6</sup> Záleží na délce věty. Jeden řádek vysílaný teletextovou technikou pojme cca 35 znaků.

<sup>7</sup> Jednořádkový titulek by měl na obrazovce setrvat cca tři sekundy, dvouřádkový pak až šest sekund, aby jej divák stihl přečíst.

předchozího kontextu, nicméně ulehčuje čtenáři titulků sledovat zároveň vizuální stránku pořadu.<sup>8</sup>

### 3.2 Mediální zákon v České republice

Sbírka zákonů, zákon č. 231/2001 Sb., o provozování rozhlasového a televizního vysílání, v § 32 uvádí základní povinnosti provozovatelů vysílání a provozovatelů převzatého vysílání. V bodě 2 se uvádí: *Provozovatel celoplošného televizního vysílání s licencí je povinen opatřit minimálně 15 % vysílaných pořadů skrytými nebo otevřenými titulky pro sluchově postižené a provozovatel celoplošného televizního vysílání ze zákona je povinen opatřit 70 % vysílaných pořadů skrytými nebo otevřenými titulky pro sluchově postižené, pokud zvláštní zákon nestanoví jinak.*“

### 3.3 Uživatelé skrytých titulků a jejich požadavky

Cílová skupina uživatelů skrytých titulků je velice rozmanitá, je tedy logické, že každá z těchto skupin má na skryté titulky jiné požadavky. Centrálním bodem diskuze o skrytých titulcích je spor o to, zda by obsah titulků měl být doslovný či nějakým způsobem upravovaný, například zkracovaný či parafrázovaný. Technologie přímého rozpoznávání řeči umožňuje tvořit titulky doslovné, vyžaduje ovšem určitá specifika, za kterých může bezchybně pracovat. Způsob vytváření titulků stínovým

---

<sup>8</sup> KOPLÍK, STRNADOVÁ, *Problematika tvorby skrytých titulků k předtočeným pořadům*, s. 29.

řečníkem závisí na mnoha faktorech, které budou zmíněny v další části práce.<sup>9</sup>

Uživatele skrytých titulků lze obecně rozdělit do pěti skupin.

### 3.3.1 Neslyšící osoby<sup>10</sup>

Jako neslyšící označujeme osoby, které se jako neslyšící narodily nebo u nich došlo ke ztrátě sluchu před vytvořením schopnosti užívat mluvený jazyk. Neslyšící osoby mají malou slovní zásobu a nedostatečnou znalost gramatiky mluveného jazyka. Z tohoto důvodu u nich nemohlo dojít k osvojení většinového jazyka spontánně poslechem do takové míry, aby neslyšící porozuměli běžnému písemnému textu. Pro plné chápání diskutovaného tématu potřebují tlumočníky znakového jazyka.

Pro tyto osoby je z výše zmíněných důvodů text skrytých titulků nesrozumitelný, nicméně jejich čtení jim pomáhá rozšiřovat pasivní slovní zásobu a umožňuje seznámení se s větnou skladbou daného jazyka. Skryté titulky mohou také vytvářet novou motivaci ke čtení a učení se většinovému jazyku. Pokud by cílovou skupinu uživatelů skrytých titulků měly tvořit pouze neslyšící osoby, musely by být titulky výrazně zjednodušeny, zejména po formální stránce. Z důvodu omezených dispozic neslyšících osob by na obrazovce musely také setrvat déle, aby byla informace plně pochopena.<sup>11</sup>

---

<sup>9</sup> Viz. kapitoly 3.4.2 Přímé titulkování a 5.2 Specifika stínového přemlouvání

<sup>10</sup> Užívá se také termínu prelingválně neslyšící.

<sup>11</sup> KOPLÍK STRNADOVÁ, *Problematika tvorby skrytých titulků k předtočeným pořadům*, s. 207.



### 3.3.2 Osoby ohluchlé po osvojení si mluveného jazyka<sup>12</sup>

Pod pojmem „osoba ohluchlá po osvojení si mluveného jazyka“ chápeme člověka, u kterého došlo ke ztrátě sluchu poté, kdy si již plně osvojil mluvenou řeč a mateřský jazyk. Jejich znalosti mateřského jazyka i schopnost čtení jsou srovnatelné se slyšící populací. Ohluchlí uživatelé skrytých titulků požadují jejich plnou verzi bez zkracování a bez záměn slov. Pro plné porozumění potřebují navíc dodatečné informace, například určení osob na obrazovce a přepis neřečových zvuků.<sup>13</sup>

### 3.3.3 Osoby nedoslýchavé

Početně největší skupinu uživatelů skrytých titulků tvoří osoby nedoslýchavé. Tyto osoby využívají skryté titulky zejména z důvodu podbarvení mluveného projevu různými zvuky (například hudba). Za takovýchto podmínek rozpoznávají nedoslýchaví mluvenou řeč jen obtížně, proto většina z nich využívá skrytých titulků. Protože je jejich sluch postižen jen částečně, není nutná indikace rozlišování mluvčích a označování neřečových složek děje. Nedoslýchaví lidé užívají skryté titulky pouze pro doplnění částí promluv, které nezachytí sluchem. Proto vyžadují přesné načasování titulků a doslovný přepis mluveného slova. Doslovný přepis ale není možný z hlediska čitelnosti titulku a způsob tvorby titulků u živého vysílání nezajišťuje ani přesné načasování.

---

<sup>12</sup> Lze užít také termínu postlingválně neslyšící.

<sup>13</sup> KOPLÍK STRNADOVÁ, Problematika tvorby skrytých titulků k předloženým pořadům, s. 208.

Pokud by ale titulky respektovaly požadavky pouze této cílové skupiny, osoby neslyšící a ohluchlé by nedostávaly vzhledem ke svým potřebám plnohodnotné titulky.<sup>14</sup>

### **3.3.4 Osoby slyšící**

Početnou skupinou uživatelů skrytých titulků jsou slyšící diváci, kteří si z různých důvodů nepřejí nebo nemohou zapnout zvuk vysílání. Skryté titulky jsou také často využívány cizinci, pro které je mluvená řeč obtížně pochopitelná, ale ve spojení s písemnou formou zvukové stopy dokáží informaci plnohodnotně pochopit.<sup>15</sup>

## **3.4 Zpřístupňování pořadů neslyšícím v České televizi**

Česká televize jako provozovatel celoplošného vysílání má za povinnost opatřit 70 % celého vysílání otevřenými nebo skrytými titulky pro sluchově postižené. V této části práce budou uvedeny způsoby, jakými ČT tento zákonný požadavek naplňuje. Důraz bude kladen na skryté titulky vytvářené během živého vysílání. V závěru budou zmíněny statistiky vysílání pořadů opatřených skrytými titulky za rok 2013.

### **3.4.1 Skryté titulky**

Jak již bylo uvedeno v předchozích kapitolách, skryté titulky jsou vytvářeny buď po dokončení pořadu, nebo přímo v živém vysílání. Protože skryté titulky vyráběné po dokončení pořadu nejsou předmětem

---

<sup>14</sup> KOPLÍK STRNADOVÁ, *Problematika tvorby skrytých titulků k předtočeným pořadům*, s. 207.

<sup>15</sup> TAMTÉŽ, s. 209

této diplomové práce, zaměřuje se následující kapitola na způsoby tvorby skrytých titulků během živého vysílání.

### 3.4.1.1 Přímé titulkování

Přímé titulkování využívá aplikací systému automatického rozpoznávání řeči (ASR – Automatic Speech Recognition), který umožňuje tvorbu skrytých titulků v reálném čase. Poprvé byl podobný systém představen v roce 2003 ve Velké Británii.

V České republice se začal tento systém vyvíjet v roce 2006 na Katedře kybernetiky Západočeské univerzity v Plzni (KKY ZČU) v rámci projektu ELJABR.<sup>16</sup> Primárním účelem projektu byl vývoj nového systému automatického rozpoznávání řeči při tvorbě skrytých titulků u živého vysílání. V první fázi bylo cílem vytvořit systém vytvářející skryté titulky přímo ze zvukové stopy pořadu, ve druhé fázi pak tvorba titulků za pomoci tzv. stínového řečníka.

První test přímého titulkování se uskutečnil v roce 2008, od roku 2010 Západočeská univerzita zajišťuje přímé titulkování pro ČT jako stálou službu v úhrnu s více než šesti sty otitulkovánými hodinami vysílání.<sup>17</sup>

Přímé titulkování se používá v případě živých přenosů z Poslanecké sněmovny a Senátu Parlamentu České republiky. Možnost využití systému přímého rozpoznání z originální zvukové stopy pořadu je umožněna díky tomu, že účastníci jednání jsou školení mluvčí a pravidla jednání neumožňují mluvit několika řečníkům najednou. Důležitými faktory

---

<sup>16</sup> ELIMINACE Jazykových Bariér handicapovaných diváků České televize.

<sup>17</sup> MÜLLER, *Automatické titulkování živých pořadů České televize – současný stav a výhled do budoucna*, s. 1–2.

pro korektní automatické rozpoznávání jsou stabilní akustika v jednacím sále a kvalitní audiosystém přenášející zvukový signál.<sup>18</sup>

Protože systémy automatického rozpoznávání řeči obvykle nepracují bezchybně, je nutné ohodnotit, jak dobře systém vyřčená slova rozpoznává. Ohodnocení správného rozpoznání slova systémem se vyjadřuje tzv. „mírou důvěry v rozpoznávání“ vyjádřenou v procentech. Čím je míra důvěry vyšší, tím větší je šance, že slovo bylo rozpoznáno správně. V běžných aplikacích přesahuje procento špatně rozpoznané promluvy zhruba dvacet procent. Nejčastějším faktorem, který má vliv na míru rozpoznávání je výskyt slov stojících mimo rozpoznávaný slovník (Ang. Out-of-Vocabulary, zkr. OOV). Taková slova systémem automatického rozpoznávání řeči nedokáže nikdy správně rozpoznat.<sup>19</sup> Problém slov OOV je eliminován, pokud je systém automatického rozpoznávání řeči ovládán stínovým řečníkem, který vkládá neznámé výrazy do slovníku systému i během procesu tvorby titulků.

Přesnost rozpoznávání u přímého titulkování je různá, záleží na způsobu a tempu mluvy rozpoznávaných řečníků. V průměru se přesnost pohybuje mezi 85–88 %. Pokud je průběh zasedání striktně řízen, například v případě hlasování, pohybuje se přesnost rozpoznání až okolo 90 %. Většina rozpoznávaných chyb je způsobena chybějícími, přebytečnými nebo na špatné místo umístěnými předloškami – toto ale nepůsobí problémy v pochopení obsahu titulků. Tyto chyby lze navíc opravit při případném dalším zpracování výsledných titulků při repríze pořadu.<sup>20</sup>

---

<sup>18</sup> TRMAL, *Online TV captioning of Czech Parliamentary Sessions*, s. 417.

<sup>19</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 449–450.

<sup>20</sup> TRMAL, *Online TV captioning of Czech Parliamentary Sessions*, s. 418.

Při tvorbě skrytých titulků pomocí přímého titulkování dochází k časové prodlevě dvou až čtyř vteřin mezi proneseným slovem a zobrazením výsledného titulku na obrazovce. Toto zpoždění je způsobeno zpracováním akustického signálu systémem automatického rozpoznávání řeči a formátováním výsledných titulků. K finálnímu zpracování titulků dochází v Plzni, a následně jsou odesílány po telefonní lince do Prahy.<sup>21</sup>

#### **3.4.1.2 Stínové přemlouvání**

Protože výše zmíněné přímé titulkování živého vysílání není vhodné pro všechny druhy pořadů, byl na Katedře kybernetiky Západočeské univerzity v Plzni vyvinut systém automatického rozpoznávání řeči ovládaný stínovým řečníkem. Tento způsob tvorby skrytých titulků se používá zejména u pořadů, kde je zvuková stopa hůře rozpoznatelná, řečníci si skáčou do řeči, mluví spontánně nebo nespisovně. Dále se jedná o pořady s hlukem na pozadí, nebo pořady, v nichž zvuk není přenášen pomocí mikrofonu.

V případě těchto pořadů pak není rozpoznávána přímo akustická stopa pořadu, ale zvuková stopa, kterou vytváří stínový řečník. Stínový řečník poslouchá originální zvukovou stopu pořadu ve sluchátkách a přemlouvá ji do mikrofonu systému automatického rozpoznávání mluvené řeči, který diktovaný text zapíše v reálném čase do výsledných titulků. Přesnost rozpoznávání se pohybuje v průměru okolo 95 %.

Systém automatického rozpoznávání řeči je nainstalován v přenosných počítačích, které mají stínoví řečníci k dispozici. Počítač je propojen přes internet se zařízením zajišťujícím přenos vyhotovených titulků, které je umístěno v prostorách Katedry kybernetiky Západočeské univerzity.

---

<sup>21</sup> TAMTÉŽ, s. 419.

Stínový řečník může pracovat odkudkoliv, pokud má k dispozici vhodné prostředí pro diktování do systému a stabilní připojení k internetu.

Tento způsob tvorby skrytých titulků funguje v České televizi od roku 2012. Stejně jako u přímého titulkování jsou titulky k dispozici na teletextové stránce 888.<sup>22</sup> Titulky vyhotovené stínovým řečníkem se pak dále upravují, zejména v načasování, a znovu používají při reprízách pořadů.

Katedra kybernetiky Západočeské univerzity v současné době disponuje sedmi stínovými mluvčími, kteří vytvářejí skryté titulky v živém vysílání a dvěma stínovými řečníky v závěrečné fázi tréninku. Trénovaný stínový řečník zvládne bez pauzy přemluvit až dvě hodiny vysílání.

Seznam aktuálně přemlouvaných pořadů je k dispozici na webové stránce [www.zivetitulky.cz](http://www.zivetitulky.cz).

### **3.4.2 Pořady přemlouvané stínovými řečníky**

V této podkapitole budou blíže představeny pořady, které stínoví mluvčí opatřují skrytými titulky. Do diplomové práce byla tato podkapitola zařazena zejména z toho důvodu, že většina z níže uvedených pořadů zaměřených na politická témata bude dále analyzována v praktické části diplomové práce.

---

<sup>22</sup> MÜLLER, *Automatické titulkování živých pořadů České televize – současný stav a výhled do budoucna*, s. 2–3.

### 3.4.2.1 Otázky Václava Moravce

Otázky Václava Moravce jsou diskusní pořad zabývající se především politickými a celospolečenskými tématy uváděný moderátorem Václavem Moravcem. Vysílán je pravidelně každou neděli. Stopáž pořadu činí zhruba dvě hodiny, přičemž vysílání je rozděleno mezi dvě stanice České televize. První hodina Otázek se vysílá na programech ČT1 i ČT24, druhá hodina pak pouze na programu ČT24. Obě části jsou opatřovány skrytými titulky, které v celé délce pořadu vytváří jeden stínový řečník. Přestávku mezi oběma částmi pořadu tvoří krátké zprávy, během kterých má stínový řečník také pauzu, protože skryté titulky ke zprávám zajišťuje Česká televize sama. Pořad Otázky Václava Moravce je specifický svými ustálenými frázemi. Jedná se zejména o věty „*Dobrý den a hezkou neděli Vám všem*“, kterou obvykle začíná první hodina vysílání. Druhou hodinu pak otevírá fráze „*A vstoupili jsme do druhé hodiny dnešních otázek*“. Mottem pořadu je pak otázka „*A o jakých tématech se po dnešních Otázkách začne mluvit?*“

Diskuze je řízena jedním moderátorem, počet hostů ve studiu se pak liší. V závislosti na diskutovaném tématu a aktivitě hostů se diskuze často vyznačuje vzájemným vstupováním do projevu a významově nesouvislými celky. Proto stínové přemlouvání Otázek Václava Moravce vyžaduje mnoho hodin tréninku, aby si stínový řečník zvykl na specifika tohoto pořadu, jakým je například rychlost mluveného projevu moderátora Václava Moravce.

### 3.4.2.2 Hyde Park

Hyde Park je další z živě vysílaných pořadů, které stínoví mluvčí pravidelně opatřují skrytými titulky. Do vysílání je zařazován každý všední večer na stanici ČT24. Doba vysílání se pohybuje okolo jedné hodiny. Hyde Park již není v takové míře zaměřen na politická témata, jako je

tomu u Otázek Václava Moravce. Objevují se zde diskuze na témata z rozličných oblastí, jako je například zdravotnictví, armáda, sport a další. V případě tohoto pořadu stínový mluvčí přemlouvá nejčastěji diskuzi dvou lidí, hosta a moderátora. Ve výjimečných případech se ve vysílání objevuje i více hostů.

Specifikem tohoto diskusního pořadu je způsob pokládání otázek, na které odpovídá host pořadu. Otázky pokládají diváci pořadu prostřednictvím telefonu, textových zpráv a zejména sociálních sítí. Dotazy položené prostřednictvím sociální sítě se objeví na obrazovce ve studiu, uživatel skrytých titulků si ji tedy může přečíst, stínový mluvčí ji nepřemlouvá, pokud ano, tak pouze ve zkrácené formě. Časté jsou také vstupy do vysílání, kdy diváci pokládají otázky přímo v přímém přenosu prostřednictvím kamer České televize.

Celý pořad je přemlouván jedním stínovým mluvčím.

### **3.4.2.3 Ekonomika ČT24**

Ekonomika ČT24 je pořad zaměřený zejména na ekonomická témata, který se objevuje na obrazovce stanice ČT24 vždy po odvysílání pořadu Hyde Park. Stopáž pořadu se pohybuje okolo třiceti minut. Obsah pořadu sestává z ekonomických aktualit a rozhovorů s odborníky.

Vzhledem k tomu, že Ekonomika ČT24 bezprostředně navazuje na vysílání pořadu Hyde Park, je přemlouvána stejným stínovým mluvčím.

### **3.4.2.4 Pořad 168 hodin**

168 hodin je zpravodajský pořad, který vysílá stanice ČT1 vždy v neděli večer. Jedná se o půlhodinový souhrn hlavních událostí, které se odehrály



v předchozím týdnu. Specifikem tohoto pořadu je, že se nejedná o živé vysílání, ale pořad je předem předtočen. Protože se ale jedná o zpravodajství, které má zachytit ty nejaktuálnější informace, je finální podoba pořadu vytvořena pouze několik hodin před odvysíláním. Česká televize proto není schopna zajistit vytvoření skrytých titulků předem. Z tohoto důvodu je pořad, i přes to, že není vysílán živě, opatřován skrytými titulky prostřednictvím stínového mluvčího. Jelikož se jedná o předtočený pořad, je jazykový projev moderátorky i jednotlivých redaktorů velice rychlý. Stínové přemlouvání tohoto pořadu tak vyžaduje stínového mluvčího, který si dokáže promluvu dopředu poslechnout a co nejvíce ji zjednodušit, aby ji stihl nadiktovat do systému rozpoznávání řeči, a to samozřejmě při zachování plného sémantického obsahu.

#### **3.4.2.5 Interview Daniely Drtinové**

Jedná se o pravidelnou půlhodinovou debatu na stanici ČT24, která je do vysílání zařazena každý všední den v podvečer. Náplní pořadu je aktuální téma, o kterém debatují vždy jeden host a moderátor pořadu. Po určitou dobu byl pořad zpřístupňován divákům se specifickými potřebami prostřednictvím znakového jazyka, v současné době je přemlouván jedním stínovým řečníkem.

#### **3.4.2.6 Další pořady**

Stínoví mluvčí samozřejmě nepřemlouvají pouze výše zmíněné pořady. Dále opatřují skrytými titulky různé druhy zábavných pořadů a sportovních přenosů.

Vrátíme-li se ale ještě k pořadům zaměřeným na politiku, přemlouvány jsou různé speciální politické debaty, zejména v době voleb. V minulých letech opatřovali stínoví mluvčí skrytými titulky také například pořad Máte slovo moderovaný Michaelou Jílkovou.

Co se týče přemlouvání zábavných pořadů, jedná se zejména o přímé přenosy udílení různých cen, titulován je pravidelně například Atlet roku nebo Zlatá hokejka. V posledních dvou letech byly také skrytými titulky opatřovány přímé přenosy pořadu StarDance vysílaný na stanici ČT1.

V posledních letech začal také stoupat objem pořadů opatřovaných skrytými titulky na stanici ČT sport. Pravidelně jsou přemlouvány fotbalové a hokejové zápasy, mezinárodní i tuzemské. Dále jsou skrytými titulky opatřovány tenis a klasické lyžování. V budoucnu by mělo dojít ještě k dalšímu zvýšení počtu přímých sportovních přenosů, u kterých budou skryté titulky vyhotovovány.

Sportovní i zábavné pořady přemlouvají jeden nebo dva stínoví mluvčí v závislosti na délce pořadu.

### **3.4.3 Statistiky**

Česká televize každoročně uveřejňuje na svých webových stránkách statistiky opatřování vysílaných pořadů skrytými a otevřenými titulky nebo znakovým. Statistika ukazuje data na pořadech čtyř stanic České televize – ČT1, ČT2, ČT24 a ČT sport v jednotlivých čtvrtletích roku.

Česká televize v souhrnu všech čtyř stanic splnila zákonný požadavek opatřit alespoň 70 % svého vysílání skrytými nebo otevřenými titulky, hranice byla dokonce překročena, celkový objem pořadů opatřených skrytými či otevřenými titulky činil 71,8 %. Sedmdesátiprocentní požadavek na opatření pořadů skrytými nebo otevřenými titulky plní pouze televizní stanice ČT1 (90,4 %), ČT2 (86,7 %) a ČT24 (72,1 %). Za

zákonnou kvótou výrazně zaostává program ČT sport, kde bylo skrytými nebo otevřenými titulky opatřeno pouze 35,4 % vysílaných pořadů.<sup>23</sup>

---

<sup>23</sup> 2013 – *Pořady s titulky a/nebo se znakovým jazykem ve vysílání České televize* [online]. [cit. 2014-01-31]. Dostupné z: [http://img.ceskatelevize.cz/boss/image/contents/sledovanost/cisla-o-vysilani/2013/4q2013\\_4\\_neslysi.pdf](http://img.ceskatelevize.cz/boss/image/contents/sledovanost/cisla-o-vysilani/2013/4q2013_4_neslysi.pdf)

## 4 SYSTÉM AUTOMATICKÉHO ROZPOZNÁVÁNÍ ŘEČI

V následující kapitole bude blíže představen systém rozpoznávání řeči a jeho jednotlivé součásti – akustický model, jazykový model a jeho trénování a dekodér systému.

Počítačovým rozpoznáváním řeči se výzkumné laboratoře zajímají více než padesát let. Úkolem systému automatického rozpoznávání řeči je transformovat posloupnost slov pronesenou řečníkem a převést ji do psaného textu.

I když byly na poli tohoto výzkumu zaznamenány mnohé úspěchy, je konstrukce zařízení, které by rozpoznalo a do textu přepsalo mluvu jakéhokoliv řečníka užívajícího libovolná slova, zatím stále ve fázi vývoje. Jako nejobtížnější faktory se při vývoji takového systému projevují variabilita řečníků a zvukového prostředí, stejně jako tvorba dostatečně obsáhlého slovníku rozpoznatelných slov.

Odlišné parametry hlasového ústrojí jednotlivých lidí způsobují různost lidských hlasů a také artikulace. Toto má za následek, že každá osoba má jinou barvu hlasu, jinou intonaci, tempo řeči atd. Z tohoto důvodu se systémy rozpoznávání řeči dělí na dva typy. Prvním jsou tzv. „systémy na řečníku nezávislé“, které jsou natrénovány na hlasech stovek i tisíců různých mluvčích. Druhým typem jsou pak „systémy na řečníku závislé“, které jsou natrénovány na hlas konkrétního řečníka.<sup>24</sup> Systém závislý na řečníku vykazuje vyšší míru přesnosti rozpoznávání, proto každý ze stínových mluvčích užívá akustický model natrénovaný přímo na svůj hlas.<sup>25</sup>

---

<sup>24</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 195.

<sup>25</sup> Viz. kapitola 4.1. Akustický model

Systém rozpoznávání řeči funguje na statistických metodách. Slova a celé promluvy jsou modelovány pomocí tzv. „skrytých Markovových modelů“<sup>26</sup>. „Princip metody modelování řeči Markovovými modely vychází z představy o vytváření řeči. Při generování řeči člověkem si lze představit, že hlasové ústrojí je během krátkého časového intervalu (např. mikrosegmentu) v jednom konečném počtu stavů artikulačních konfigurací (je nastaveno pro vyslovení určité hlásky). V uvažovaném mikrosegmentu je pak hlasovým ústrojím produkován krátký signál, který závisí na stavu artikulačního ústrojí a může být popsán určitými spektrálními charakteristikami, jež jsou prezentovány vhodnými příznaky (vektorem příznaků).<sup>27</sup> Skryté Markovovy modely pak v diskretních časových okamžicích generují náhodnou posloupnost pozorování. V jednotlivých časových krocích mění model svůj stav podle předem daných pravděpodobností. Stav, do kterého model přejde, vygeneruje vektor pozorování a to podle rozdělení výstupní pravděpodobnosti příslušné tomuto stavu.<sup>28</sup>

Jedním skrytým Markovovým modelem mohou být modelována jednotlivá slova, častěji jsou ale vytvářeny skryté Markovovy modely subslovních jednotek (fonémů, trifónů apod.) a výsledná promluva je pak konstruována sřetením těchto subslovních jednotek.

Systém rozpoznávání řeči se skládá z akustického procesoru a lingvistického dekodéru. Uživatel a akustický procesor jsou spojeni akustickým kanálem. Akustický procesor mění řečové kmity řečníka na

---

<sup>26</sup> Hidden Markov Model, zkr. HMM.

<sup>27</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 200.

<sup>28</sup> TAMTÉŽ, s. 200–201.

posloupnosti vektorů příznaků. Lingvistický dekodér transformuje řetězce příznaků na řetězce slov.<sup>29</sup>

#### 4.1 Akustický model

Akustický model modeluje fonémy daného jazyka za využití statistických přístupů.

Protože systém automatického rozpoznávání řeči užívaný stínovým mluvčím rozpoznává zvukovou stopu produkovanou stínovým řečníkem, je nutné systém natrénovat přímo na hlas daného stínového řečníka.

V první fázi tréninku stínových mluvčích se používá obecný, na pohlaví závislý akustický model. Tento model se automaticky přizpůsobí charakteristikám hlasu stínového řečníka během tréninku, nejlépe v jeho druhé fázi.

Protože řečový signál vyslovený uživatelem systému je velmi variabilní a tutéž promluvu vysloví každý řečník jinak, je ve druhé fázi, po absolvování tréninku, systém natrénován přímo na charakteristickou mluvu daného řečníka. Pro vytvoření plnohodnotného akustického modelu závislého na konkrétním stínovém řečníkovi je potřeba shromáždit zhruba sto hodin tréninkových nahrávek, ze kterých je následně akustický model natrénován.<sup>30</sup> Z tohoto důvodu by adaptovaný akustický model neměl využívat jiný mluvčí, než kterým byla adaptace příslušného akustického modelu provedena.

---

<sup>29</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 196–200.

<sup>30</sup> PRAŽÁK a kol., *Novel Approach to Live Captioning Through Re-speaking: Tailoring Speech Recognition to Re-speaker's Needs*, s. 1371.

V přesnosti rozpoznávání řečového signálu se pak výrazně projeví také jakákoliv změna prostředí, například jiná akustika místnosti, než ve které byl systém užíván při tréninku, nebo také změna stálosti přenosového kanálu, tzn. např. užití jiného mikrofону nebo výkyvy ve stálosti řeči (příliš hlasitá mluva, šeptání, koktání apod.). Z těchto důvodů tedy akustický model funguje nejlépe za stejných podmínek, za jakých byla pořizována trénovací data, tzn. stejný mikrofón, identická akustika místnosti a úroveň hluku, přízvuk, styl řeči atd.

## 4.2 Jazykový model

Vedle akustického modelu je jazykový model další důležitou částí systému rozpoznávání řeči. Jazykový model určuje pravděpodobnosti pro všechny možné posloupnosti slov. Algoritmus určování pravděpodobností jazykového modelu by měl přispívat k rozpoznávání posloupnosti slov již v průběhu přemlouvání stínovým řečníkem. Systém rozpoznávání řeči vyvinutý na Katedře kybernetiky ZČU pro tento účel využívá tzv. „trigramy“, kdy je pravděpodobnost každého následujícího slova podmíněna dvěma slovy bezprostředně předcházejícími.

Protože každý jazyk má svá specifika, je nutno tato specifika vzít v potaz i při tvorbě jazykových modelů. Jedním z takových specifíků je slovník, který daný jazyk využívá, a také zákonitosti, podle kterých se slova v určitém jazyce řetězí do vyšších celků. Dalším ze specifických kritérií jazyků je také výslovnost jednotlivých výrazů. V některých případech lze k jednomu slovu přiřadit i více výslovnostních variant. Všechna tato specifika jsou v systému rozpoznávání řeči zajišťována právě jazykovým modelem.<sup>31</sup>

---

<sup>31</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 225–226.

### 4.2.1 Trénování jazykového modelu

Pro natrénování jazykového modelu je nutno shromáždit velký počet trénovacích dat odlišného charakteru a pocházejících z různých zdrojů. Potřebné texty musí být upraveny pro potřeby systému rozpoznávání řeči. Tvorba jazykového modelu pro program je stejná důležitá jako samotný vývoj programu automatického rozpoznávání řeči.

Texty jsou získávány z novin, otevřených i skrytých titulků, přepisu televizních pořadů a z internetu. Zejména webové stránky jsou zdrojem článků pro tvorbu velkých textových korpusů potřebných pro tvorbu jazykových modelů. Webové stránky nejsou vybírány náhodně, ale jsou voleny takové servery, které jsou relevantním zdrojem jazykového materiálu. Webové servery publikují články vztahující se k aktuálním tématům a s velkou frekvencí se v nich objevují nová slova. Jako příklad lze uvést jadernou havárii v Japonsku, která do médií přinesla název elektrárny Fukušima. Předností dat získávaných z webových stránek je možnost okamžitého zpracování nově publikovaných článků, což umožňuje aktualizaci stávajících jazykových modelů na denní bázi.<sup>32</sup>

#### 4.2.1.1 Zdroje dat pro tvorbu jazykových modelů

Jazykové modely pro systém rozpoznávání řeči jsou vytvářeny na základě dat získávaných ze zpravodajských webových serverů. Jedná se zejména o weby [ceskenoviny.cz](http://ceskenoviny.cz), [idnes.cz](http://idnes.cz) a [lidovky.cz](http://lidovky.cz). Články získávané z webových serverů Českých novin a Lidovek obsahují zejména zprávy z domácího a zahraničního prostředí, dále pak obchodní, kulturní

---

<sup>32</sup> ŠVEC a kol., *Web Text Data Mining for Building Large Scale Language Modelling Corpus*, s. 356–357.



a sportovní zprávy. Webový server iDnes navíc zveřejňuje také místní zprávy, články o bydlení a technice.

Důležitým zdrojem jsou přepisy televizního zpravodajství a diskusních pořadů České televize, klíčovým pramenem jsou zejména přepisy politického diskusního pořadu Otázky Václava Moravce. Nemalou část dat tvoří také přepisy zpravodajských pořadů rádiových stanic.<sup>33</sup>

Při tvorbě jazykových modelů pro český jazyk naráží vývojáři často na závažný problém týkající se flexe. Čeština, stejně jako ostatní slovanské jazyky, vykazuje vysokou mírou ohebnosti. To znamená, že pro každý základní tvar, tzv. „lemma“, zároveň existuje velké množství jeho morfologických tvarů. U podstatných jmen, přídavných jmen, zájmen a číslovek se tyto tvary odlišují nejčastěji v koncovce vyjadřující mluvnický pád, číslo a rod. U sloves jsou to pak mluvnické číslo, rod a osoba. Často ale dochází také ke změně kmene slova nebo k připojení předpony ke slovesnému tvaru. Tímto vznikají stovky různých morfologických variant u každého ohebného slovního druhu. Flektivnost jazyka z výše uvedených důvodů představuje závažný problém, protože slovo je z hlediska rozpoznávání systémem automatického rozpoznávání řeči definováno svým zápisem. Jazykový model vnímá tedy každý tvar jako nové samostatné slovo. Flektivnost jazyka tak podstatně ovlivňuje velikost slovníku, která roste s narůstajícím objemem korpusu. V tomto případě se ukazuje výhoda tvorby jazykových modelů například v angličtině, která se vykazuje velmi nízkou mírou flexe.

Problém lze vyřešit využitím jazykového modelu založeného na morfémech. Jako příklad lze uvést rozložení slov z trénovacích dat na kmeny a koncovky. Takové morfémy lze pak použít jako základní jednotku

---

<sup>33</sup> ŠVEC a kol., *Web Text Data Mining for Building Large Scale Language Modelling*

jazykového modelu. Tento způsob značně sníží velikost slovníku, protože kmen obvykle není ovlivněn ohebností jazyka. Koncovky pak tvoří jen zanedbatelnou část slovníku, jejich počet se v českém jazyce pohybuje okolo jednoho tisíce.<sup>34</sup>

### 4.3 Dekodér

Dekodér využívá pravděpodobností z jazykového a akustického modelu. Na základě těchto informací pak určuje nejpravděpodobnější posloupnost slov, kterou chtěl řečník vyslovit. Dekodér pracuje na bázi kompromisu mezi přesností a rychlostí rozpoznávání.

Specifickou vlastností všech systémů rozpoznávání řeči je, že nový akustický signál přicházející od uživatele systému mění posledních několik rozpoznávaných slov. V systému automatického rozpoznávání řeči užívaném na Katedře kybernetiky Západočeské univerzity se tato poslední slova zobrazují šedou barvou a stínový řečník má možnost je smazat a informaci znovu přemluvit, pokud vidí, že se diktát rozpoznal špatně. Je to umožněno z toho důvodu, že tato „šedá slova“ systém dosud neodeslal do vysílání a je možno je změnit. Je tedy nezbytné, aby stínový mluvčí šedá slova neustále kontroloval a případně opravoval, což výrazně sníží počet chyb ve výsledném titulku. Mazat a opravovat nesprávně rozpoznaná slova se stínový mluvčí učí ve třetí a čtvrté fázi tréninku pomocí klávesového povelu.

---

*Corpus*, s. 358.

<sup>34</sup> PSUTKA a kol., *Mluvíme s počítačem česky*, s. 267–268.

#### 4.4 Rozpoznávací komplet

Systém automatického rozpoznávání řeči má k dispozici rozpoznávací slovník, který obsahuje okolo milionu slov.<sup>35</sup> Aby bylo dosaženo co nejvyšší přesnosti rozpoznávání, užívá se pro každý pořad titulkovaný stínovým řečníkem navíc specifický rozpoznávací komplet, který obsahuje slovník určený pro danou tematickou oblast. Jiný komplet je tedy užit například pro přemlouvání politických debat a jiný pro titulkování sportovních přenosů. U sportu jsou jazykové komplety navíc doplňovány o soupisky jmen sportovců.

Pro vyšší efektivitu jsou jazykové komplety pro politické debaty ještě vytvořeny přímo pro specifika jednotlivých titulkovaných pořadů. Uvedeme-li příklad, znamená to, že stínový řečník titulkující pořad Otázky Václava Moravce volí jiný jazykový komplet než mluvčí vytvářející skryté titulky pro diskusní pořad Hyde Park. Stejně je tomu i u sportovních přenosů.

Během tréninku mají kandidáti na stínové řečníky k dispozici dvě verze kompletu pro Otázky Václava Moravce. První verze jazykového kompletu neumožňuje vkládat interpunkci, užívá se tedy při přemlouvání v prvních třech etapách tréninku. Druhá verze kompletu určená pro poslední fázi tréninku již disponuje možností vkládat do rozpoznávaného textu interpunkční znaménka.

---

<sup>35</sup> PRAŽÁK a kol., *Novel Approach to Live Captioning Through Re-speaking: Tailoring Speech Recognition to Re-speaker's Needs*, s. 1371.

## **5 TRÉNINK STÍNOVÝCH MLUVČÍCH**

Na stínového řečníka jsou během jeho činnosti kladeny vysoké nároky. Protože musí zvládnout vykonávat několik činností najednou, je nutné, aby každý ze stínových řečníků absolvoval čtyřfázový trénink, který pomáhá postupně rozvíjet jeho schopnosti ovládat systém automatického rozpoznávání řeči a správně do něj svou promluvu diktovat. V následující kapitole budou uvedeny požadavky kladené na stínového řečníka a dále se tato kapitola bude věnovat specifickým stínového přemlouvání. Protože stínový mluvčí vytvářejí během své činnosti slohově specifický komunikát, budou v této kapitole také představeny jeho charakteristiky. V dalších částech kapitoly budou popsány jednotlivé fáze tréninku, který musí být stínovými mluvčími absolvován předtím, než začnou vytvářet skryté titulky přímo v živém vysílání. Proces tréninku zefektivňuje, zkracuje a také zlevňuje výcvik stínových mluvčích.

### **5.1 Úkony prováděné stínovým mluvčím během diktování**

Jak již bylo zmíněno výše, práce stínového řečníka není jednoduchá, proto také není každý z kandidátů pro plnohodnotné vykonávání této funkce vhodný.

Během tvorby skrytých titulků musí být stínový řečník schopen vykonávat několik paralelních činností. Klíčovým faktorem je schopnost poslouchat zvukovou stopu přemlouvaného pořadu a zároveň ji diktovat do systému automatického rozpoznávání řeči. Samozřejmostí je, že stínový řečník stačí svým diktátem tempu řečníka na obrazovce takovým způsobem, aby byl co nejvíce zachován sémantický obsah původní promluvy.

Nutností je také neustálá kontrola rozpoznávaného textu. Pokud systém rozpozná slovo chybně, musí stínový řečník slovo smazat a znovu

zopakovat. V případě, že systém slovo opakovaně nerozpozná, výraz není součástí slovníku, proto musí stínový řečník slovo do slovníku přidat.<sup>36</sup> Učiní tak stiskem příslušné klávesy a napsáním slova na klávesnici počítače. Pokud se jedná o cizí slovo, má možnost přidat návrh výslovnosti nového výrazu.

Systém výsledné titulky odesílá do vysílání automaticky, pokud se ale ve vysílání objeví delší pauza (např. obraz bez komentáře, zvuková upoutávka k pořadu), je úkolem stínového mluvčího odeslat rozpoznaný text pomocí stisku klávesy, aby se zobrazil divákovi v celém znění.

Další činnost, kterou stínový mluvčí vykonává, je rozlišování mluvčích stiskem příslušné klávesy. Systém následně barevně ve skrytých titulcích rozliší až čtyři debatující. Rozlišení řečníků může být obtížné, zejména pokud během debaty mluví více mluvčích najednou. I v takovém případě ale musí být stínový řečník schopen určit řeč hlavního mluvčího.

Další schopností, kterou musí stínový mluvčí disponovat, je převedení projevu debatujících do korektní jazykové formy, kterou systém dokáže rozpoznat. Pokud stínový mluvčí bude do systému diktovat nespisovně, tak jak to například slyší od řečníka ve vysílání, systém diktát s vysokou pravděpodobností nerozpozná, protože je natrénován na textech obsahujících spisovný jazyk.

Stínový mluvčí musí také disponovat znalostmi pravidel interpunkce ve větách, protože do rozpoznávaného textu vkládá během mezislovních pauz stiskem klávesy<sup>37</sup> také interpunkční znaménka.<sup>38</sup> Interpunkce je důležitá

---

<sup>36</sup> Typicky příkladem slov vkládaných do slovníků jsou vlastní jména a názvy.

<sup>37</sup> Systémy automatického rozpoznávání umožňují vkládat interpunkci také slovním povelům. Systém vyvinutý na KKY ZČU slovních povelů pro zadání interpunkce

zejména z pohledu členění přemluveného textu do skrytých titulků. Protože řádka titulku na obrazovce disponuje pouze omezeným počtem znaků, dělí se zobrazovaný titulek do jednotlivých řádků právě podle vložené interpunkce. Jazykový model chápe interpunkční znaménko jako samostatné slovo, které ovlivňuje následné rozpoznávání.<sup>39</sup> Zadá-li tedy stínový mluvčí do systému například tečku, bude první slovo následující věty zapsáno s velkým písmenem.

## 5.2 Způsoby a specifika stínového přemlouvání

Styl diktátu do systému rozpoznávání řeči závisí především na stínovém řečníkovi samotném a také na tom, zda je řeč mluvčích pořadu předem připravená či spontánní. Rozhodující jsou i další faktory, které budou v této podkapitole uvedeny.

Je-li řeč dostatečně pomalá, věty gramaticky správné a významově ucelené, lze při diktování použít doslovný způsob přemlouvání. Jedná se zejména o ty druhy pořadů, kdy je řeč čtena nebo předem důkladně připravena. Příkladem může být pravidelná politická debata Otázky Václava Moravce (OVM). Moderátor má precizně připravený a vytříbený jazykový projev, který lze přemlouvat doslova. Nicméně i v tomto případě může doslovný diktát činit stínovému mluvčímu potíže, protože projev Václava Moravce vykazuje značnou rychlost. Doslovný způsob diktování do systému se ale ve většině případů nedá použít

---

využívá pouze pro vybranou interpunkci (např. dvojtečka), protože slovní povely výrazně zpožďují zobrazení výsledných titulků divákovi.

<sup>38</sup> MÜLLER, *Automatické titulkování živých pořadů České televize – současný stav a výhled do budoucna*, s. 3–4.

<sup>39</sup> PRAŽÁK a kol., *Captioning of Live TV Programs through Speech Recognition and Re-speaking*, s. 514–515.

v tomto pořadu u hostů, kteří si často skáčí do řeči a mluví spontánně. V tom případě je nutné jejich projev parafrázovat a ucelovat významově nezřetelné celky takovým způsobem, aby byl výsledný titulček pro uživatele pochopitelný.

Pokud je promluva v pořadu pro stínového řečníka příliš rychlá, je možno ji zjednodušit. Myšlenku, kterou mluvčí na obrazovce vyjádří v několika větách, si může stínový řečník poslechnout a do systému pak nadiktovat jeho myšlenku ve zkrácené formě. Zjednodušování lze také provést tzv. „prořezáním“ promluvy, kdy stínový řečník do systému neopakuje řečové výplně (například hezitační zvuky), intenzifikátory (například velmi velmi poučné) nebo odbočky nesouvisející s diskutovaným tématem.

Pokud systém rozpoznávání řeči nerozpozná nějaké slovo, nemá ho tzv. „ve slovníku“, má stínový řečník několik možností, jak situaci vyřešit. Buď může slovo do slovníku přidat, což ale výrazně zpožďuje zobrazování titulků divákovi, nebo se chybějícímu slovu vyhnout. Pokud je to možné, může použít příbuzné slovo nebo výraz opsat jinými slovy. S tímto problémem se stínový řečník často může setkat například při přemlouvání pořadu Hyde Park (HP), který se zabývá různými tématy, a ne všechna slova se mohou ve slovníku systému nacházet.

Jak bylo uvedeno v předchozí podkapitole, musí být stínový mluvčí schopen oddělit řeč hlavního mluvčího a ostatních debatujících, což je důležité zejména v případě, pokud si hosté pořadu skáčí do řeči. Neschopnost odlišit od sebe řečníky může výrazným způsobem ovlivnit kvalitu výsledných titulků, nejen z toho důvodu, že divák nebude schopen přiřadit promluvy k jednotlivým debatujícím, vezmeme-li k tomu ještě v úvahu zpoždění titulků způsobené jeho zpracováním. Také se může stát, že ve výsledných titulcích bude chybět nějaká pro diváka podstatná informace.

Obecně platí, že čím více informací si stínový řečník poslechne dopředu, tím větší prostor má pro parafrázování. Jak už ale bylo zmíněno, dochází pak ještě k výraznějšímu zpoždění ve zobrazení titulku divákovi. Nicméně styl a způsob přemlouvání závisí zejména na schopnostech a individuálním přístupu každého stínového řečníka a také na specifických vlastnostech, kterými se vyznačuje přemlouvání pořad. Nejdůležitějším výsledkem celého procesu jsou pak výsledné titulky, které musí být srozumitelné, obsahově úplné, gramaticky správné a správně naformátované. Formátování primárně určuje formát titulků a titulkovací server, který titulky odesílá do České televize. Stínový řečník může ale správným zadáváním interpunkce výrazně zlepšit čitelnost a srozumitelnost výsledných titulků.

Výše zmíněná specifika stínového přemlouvání jsou spíše obecného rázu. Detailnější informace budou uvedeny v praktické části diplomové práce.

### **5.3 Charakteristika komunikátu vytvářeného stínovým mluvčím**

Během vytváření skrytých titulků za pomoci systému automatického rozpoznávání řeči vzniká specifický typ komunikátu. Mluvené slovo je za pomoci programu automatického rozpoznávání řeči převáděno do textu psaného, přičemž výsledná podoba skrytých titulků, které se zobrazí divákovi, záleží zejména na stínovém mluvčím a jeho schopnostech efektivně diktovat do systému automatického rozpoznávání řeči takovým způsobem, aby byla zachována syntaktická a zejména sémantická přesnost výsledných skrytých titulků.

Mezi základní objektivní slovo tvorné faktory komunikátů vytvářených stínovými mluvčími patří skutečnost, že se jedná o mluvené slovo, které je zapisováno do textu. Ve výsledném komunikátu se v určité míře projevují



prvky mluveného projevu, který je zejména po syntaktické stránce velmi charakteristický. Klíčovým znakem mluveného projevu je, že nelze spolehlivě identifikovat větu, a to ani na základě intonační signalizace jejího konce. Další obtíží je často nemožnost rozhodnout, zda je možno bezprostředně po sobě následující syntaktické konstrukce chápat jako souvětné spojení parataktické nebo hypotaktické<sup>40</sup>, anebo jako konstrukce samostatné.<sup>41</sup> Z těchto důvodů může mít stínový mluvčí v mnoha případech problémy se zadáváním interpunkce a může dojít k vytváření příliš dlouhých souvětí. Protože ale řádek pro zobrazení skrytého titulku disponuje pouze omezeným počtem znaků, je vhodnější příliš komplikovaná souvětí, která byla řečena mluvčím ve vysílání, rozdělit do kratších celků. Kratší věty umožňují uživatelům skrytých titulků efektivnější pochopení obsahu. Prvky mluvenosti by měl být stínový mluvčí schopen během procesu tvorby skrytých titulků eliminovat na nejnižší možnou míru. Ve výsledném komunikátu by se tedy neměla objevit ve velké míře např. ukazovací zájmena, jejichž nadužívání je jedním z typických znaků mluveného projevu. Tyto projevy mluvenosti se stínový řečník učí postupně odstraňovat během procesu tréninku. Ne vždy se ale stínovému mluvčímu podaří během přemlouvání rysy mluvenosti redukovat, a to ani během stínového přemlouvání živého vysílání. I přes to má výsledný komunikát vytvořený stínovým mluvčím prostřednictvím systému automatického rozpoznávání řeči svými rysy blíže k textu psanému než mluvenému.

V průběhu stínového přemlouvání musí stínový řečník také neustále vyhodnocovat, zda bude systém schopen diktovaná slova rozpoznat.

---

<sup>40</sup> Užívány jsou také výrazy souvětí souřadné a podřadné.

<sup>41</sup> MÜLLEROVÁ, *Mluvený text a jeho syntaktická výstavba*, s. 42–43.

Dalším ze specifik je tedy skutečnost, že se výsledný komunikát formuje na základě možností systému automatického rozpoznávání řeči.

K dalším objektivním faktorům tvorby komunikátu za pomoci stínového řečníka patří chyby způsobené systémem rozpoznávání řeči, které stínový mluvčí není schopen během procesu přemlouvání nijak ovlivnit. Jedná se zejména o chyby ve shodě přísudku s podmětem, časté jsou také problémy s rozpoznáváním správných koncovek u určitých slovesných tvarů a rozpoznávání předložek, zejména jedná-li se o samohlásky. Obtíže způsobuje systému zejména rozpoznávání samohlásek, které stojí vedle sebe v samohláskovém trojúhelníku. Tvorba samohlásek se dá vzhledem ke vztahu k poloze jazyka ve směru nahoru a dolu pomyslně uspořádat do převráceného trojúhelníka, kdy vrchol je tvořen základní českou samohláskou A, která je umístěna uprostřed dole. Od tohoto středního nízkého A vychází dvě řady samohlásek směrem nahoru: přední a zadní. Předními samohláskami jsou středové E a vysoké I, zadními samohláskami jsou středové o a vysoké U.<sup>42</sup>

Během diktování do systému se stínový mluvčí může také setkat s problémy při rozpoznávání s tzv. „homonymy“.<sup>43</sup> Jedná se zejména o jeden z jejich typů, tzv. „nepravá homonyma“, která se shodují buď ve výslovnosti, nebo grafické podobě. Jako příklad lze uvést spojení předložky a zájmena „na to“, které systém může rozpoznávat jako „NATO“<sup>44</sup>. Zápis homonym nemůže stínový mluvčí žádným způsobem ovlivnit, závisí na jazykovém modelu systému automatického rozpoznávání řeči.

---

<sup>42</sup> HÁJKOVÁ, Český jazyk a kultura mluveného projevu, s. 32–33.

<sup>43</sup> Slova, která se shodují ve výslovnosti i grafické podobě.

<sup>44</sup> Zkratka z anglického North Atlantic Treaty Organization - Severoatlantická aliance.

Během svého projevu se mluvčí v televizním vysílání soustředí především na sdělení obsahu, často se nezabývá tím, zda je jeho projev logicky uspořádán a stínovému mluvčímu nemusí být zcela jasné, kde byla hranice mezi jednotlivými obsahovými částmi. V souvislosti se sémantickým uspořádáním textu lze hovořit o základním tématu, vedlejších tématech, o střídání témat a o tematických odbočeních. Hranice mezi nimi nejsou mnohdy zcela jasné. Nemožnost stanovit mezi těmito složkami hranice je způsobena interakcí účastníků konverzace, kteří jsou vedeni aktuálním tématem a své myšlenky sdělují dle momentálního vývoje rozmluvy.<sup>45</sup> Typickým jevem jsou tedy větné konstrukce bez explicitního vyjádření vztahu, volnost, neurčitost a významová nevyhraněnost projevu.<sup>46</sup> Obecně lze říci, že mluvčí uplatňuje během svého projevu přirozený sklon stavět na začátek promluvy informačně závažnější komponenty. Nejprve tedy sděluje to nejpodstatnější a až poté sděluje informace doplňující.<sup>47</sup> Pro stínového mluvčího je tedy klíčové, aby se během diktování do systému soustředil zejména na začátky promluv, které s nejvyšší pravděpodobností obsahují nejpodstatnější informace.

Z těchto důvodů se komunikát vytvářený stínovým mluvčím vyznačuje určitými sémantickými odlišnostmi ve vztahu k původní promluvě. Stínový mluvčí není během diktování do systému schopen přemluvit plně sémantickou hodnotu původní promluvy, zejména vzhledem k náročnosti celého procesu stínového přemlouvání a nárokům na systém

---

<sup>45</sup> MÜLLEROVÁ, *Mluvený text a jeho syntaktická výstavba*, s. 44.

<sup>46</sup> TAMTÉŽ, s. 77.

<sup>47</sup> MÜLLEROVÁ, *Mluvený text a jeho syntaktická výstavba*, s. 75.

automatického rozpoznávání řeči. Výsledné komunikáty se také vyznačují různou mírou variability, která záleží na zvyklostech způsobu přemlouvání jednotlivých stínových mluvčích. Pokud bychom nechali přemluvit různé stínové mluvčí stejnou nahrávkou, dostali bychom stejný počet odlišných výsledných komunikátů, jako byl počet stínových mluvčích. Nicméně stínový řečník musí být schopen zachovat minimálně 95 % významové hodnoty přemlouvaného projevu řečníka v televizním vysílání.

V průběhu stínového přemlouvání působí na stínového mluvčího také subjektivní stylové faktory. Vzhledem k vysokým nárokům, které jsou během stínového přemlouvání na stínového mluvčího kladeny, jsou důležité zejména jeho mentální dispozice potřebné pro vykonávání funkce stínového mluvčího, míra zkušeností nabytých během procesu tréninku, znalost tématu, které je přemlouváno, a zejména psychické rozpoložení samotného stínového řečníka, který musí být schopen zvládnout během procesu tvorby skrytých titulků mnohé stresové situace. Jako příklad lze uvést chybně rozpoznaná slova, která se odešlou do vysílání, možná nefunkčnost systému automatického rozpoznávání řeči nebo neschopnost stačit slovnímu tempu řečníka, který hovoří na obrazovce.

#### **5.4 Trenažér stínových mluvčích**

Na Katedře kybernetiky Západočeské univerzity byl vyvinut speciální tréninkový systém, tzv. „Trenažér stínových mluvčích“, který zajišťuje postupný tréninkový proces a rozvíjení schopností spojených se stínovým přemlouváním.

Trenažér nabízí možnost postupného tréninku rozděleného do čtyř etap. Pátá fáze je pak reprezentována samotnou tvorbou skrytých titulků v živém vysílání.

Po instalaci aplikace si kandidát na stínového mluvčího vytvoří svůj profil zadáním požadovaných údajů. V úvodu je také nutné vybrat odpovídající pohlaví, protože systém má k dispozici dva akustické modely. Jeden je natrénován na specifika ženských hlasů, druhý pak na zvláštnosti mužských hlasů. Důležitým krokem je také nastavení mikrofону a hlasitosti sluchátek. Nastavení mikrofónu probíhá ve dvou krocích. Nejdříve systém automaticky zjistí hladinu hluku v prostředí, ve kterém chce stínový mluvčí pracovat. Pokud je prostředí příliš hlučné, výrazně to snižuje úroveň rozpoznávání. Ve druhém kroku je pak optimalizován mikrofón a stanovena ideální hlasitost řeči. Pro usnadnění tohoto nastavení si uživatel může přehrát instruktážní video. Kandidát na stínového mluvčího má v počítači a na webu k dispozici videa roztríděná do složek odpovídajících jednotlivým fázím tréninku. Z těchto složek si pak stínový řečník vybere video, které chce přemlouvát, a zároveň s ním i odpovídající fázi tréninku, ve které se právě nachází.

První fáze umožňuje nácvik simultánního poslechu a diktování do systému. Druhá fáze pomáhá kandidátovi na stínového řečníka optimalizovat jeho projev takovým způsobem, aby vyhovoval systému automatického rozpoznávání řeči. Třetí fáze tréninku umožňuje volné přemlouvání spolu s užitím vybraných klávesových povelů. Čtvrtá fáze pak simuluje reálnou tvorbu skrytých titulků v živém vysílání včetně zadávání interpunkčních znamének. Za pátou fází lze pak považovat vytváření skrytých titulků v živém vysílání, protože i v této fázi stínovní mluvčí stále vylepšují své schopnosti.

Vzhled tréninkové aplikace simuluje vzhled programu, do kterého se diktuje promluva i během vytváření titulků v živém vysílání. Je to z toho důvodu, aby si kandidát zvykl na specifika programu již ve fázi tréninku.

Celková doba tréninku je zcela individuální. Obecně se pohybuje mezi dvěma až třemi měsíci, přičemž kandidát na stínového mluvčího musí přemluvit minimálně sto hodin nahrávek, během kterých si osvojí celý proces vytváření skrytých titulků pomocí systému automatického rozpoznávání řeči. Z těchto dat je také natrénován akustický model na hlasová specifika daného stínového mluvčího.<sup>48</sup>

## 5.5 První fáze tréninku

První fáze tréninku je klíčová pro výběr vhodných kandidátů na pozici stínového mluvčího. Cílem této etapy je vybrat ty uchazeče, kteří jsou schopni poslouchat zvukovou stopu pořadu a současně mluvit, aniž by je tyto dvě paralelně vykonávané činnosti nějak výrazně uváděly do stresového stavu. Tato fáze dosud nezapojuje systém automatického rozpoznávání řeči, to znamená, že to, co je do systému nadiktováno, není rozpoznáváno do textu.

Kandidát si v aplikaci Trenažér vybere video určené pro první fázi tréninku a ve sluchátkách poslouchá jeho zvukovou stopu. Zároveň se snaží do mikrofону diktovat to, co slyší. Přestože první stadium tréninku nevyužívá systém rozpoznávání řeči, jsou všechny výstupy kandidáta nahrávány, aby si je mohl pro vyhodnocení znovu přehrát. Vyhodnocuje je také odborný pracovník Katedry kybernetiky, který na jejich základě vybere vhodné adepty na pozici stínového řečníka a s těmi pak pokračuje v dalších fázích tréninku.<sup>49</sup>

---

<sup>48</sup> PRAŽÁK a kol., *Four-phase re-speaker training system*, s. 217–221.

<sup>49</sup> PRAŽÁK a kol., *Four-phase re-speaker training system*, s. 218.

V Trenažéru stínových mluvčích je první etapa tréninku označena jako „*fáze 1 – volný trénink komentování bez přepisu do titulků.*“

## **5.6 Druhá fáze tréninku**

Perspektivní kandidáti vybraní na základě dat z první etapy tréninku postupují do druhé tréninkové fáze, ve které je již využíván systém automatického rozpoznávání řeči, přemlouvání obsah pořadu se tedy již přepisuje do textu. Cílem této fáze je naučit kandidáta přizpůsobit jeho hlasový projev potřebám systému rozpoznávání řeči tak, aby bylo dosaženo co nejvyšší přesnosti rozpoznávání. Kandidát trénuje v této fázi doslovným přemlouváním zvukové stopy pořadu. Druhá fáze výcviku zatím neumožňuje stínovému mluvčímu mazat špatně rozpoznaná slova ani užívat klávesové povely.

Pro toto stadium tréninku jsou pro kandidáty připravena speciálně upravená videa, několikaminutové segmenty pořadů, ve kterých diskutující mluví srozumitelně, plynně a nestupují si do projevu. Nahrávka také nesmí obsahovat slova stojící mimo rozpoznávací slovník (OOV). Pro tuto etapu tréninku jsou využívány části pořadu Otázky Václava Moravce, které se vyznačují vytříbeným a předem připraveným mluveným projevem. Úkolem kandidáta je pak nahrávku mechanicky doslovně přemluvit takovým způsobem, aby dosáhl co nejvyšší přesnosti rozpoznávání.

Aby mohl systém přesnost rozpoznávání automaticky vyhodnotit, musí být k nahrávce připojen tzv. „referenční text“, tedy doslovný přepis zvukové stopy pořadu, kterou systém porovnává s rozpoznáním textem vytvořeným stínovým řečníkem. Poté, co stínový mluvčí přemluví zvolenou nahrávku, zobrazí mu systém odlišnou barvou slova, u kterých měl problém s rozpoznáním a také doplní do nahrávky symboly, které

značí vynechaná slova.<sup>50</sup> Pokud je slovo ve výsledné nahrávce podtrženo, znamená to, že došlo k záměně slova, v referenčním přepisu se nachází jiné slovo. Pokud je některé z rozpoznaných slov přeškrtnuto, znamená to, že se v referenčním textu nevyskytuje, stínový mluvčí ho nadiktoval do systému navíc. Přidá-li systém do výsledného textu červený trojúhelníček, znamená to, že v místě, kde je trojúhelník umístěn, chybí slovo uvedené v referenčním textu a stínový mluvčí ho během přemlouvání vynechal. Pokud bylo během diktátu vynecháno slov více, je před trojúhelníčkem zobrazena číslovka označující počet vynechaných slov. V rozpoznaném textu mohou být také některá slova nebo slovní spojení označena modrou barvou. Jedná se o tzv. „slova s nízkou mírou důvěry“. Během jejich rozpoznávání si systém nebyl jist, jaké slovo bylo do systému nadiktováno. Systém statisticky vyhodnotil nejpravděpodobnější možnost, která byla poté rozpoznána.<sup>51</sup>

Kandidát si také může stiskem příslušné ikony znovu přehrát část pořadu, která mu činila problémy, a porovnat ji s výsledným titulkem. Stejně tak může vypnout zvukovou stopu pořadu a poslechnout si svůj diktát. Toto umožňuje budoucímu řečníkovi i pracovníkům Katedry kybernetiky identifikovat opakující se chyby a odstranit je.

Aby stínový mluvčí i odborný pracovník Katedry kybernetiky zjistili, zda v tréninku dochází ke zlepšování schopností kandidáta, zobrazuje Trenažér také tzv. „míru důvěry“, která ohodnocuje kvalitu hlasového projevu mluvčího, tedy to, nakolik během trénování ve druhé fázi výcviku dokázal přizpůsobit systému rozpoznávání řeči svůj způsob diktátu.

---

<sup>50</sup> PRAŽÁK a kol., *Four-phase re-speaker training system*, s. 218–219.

<sup>51</sup> PRAŽÁK, *Trenažér stínových mluvčích*, s. 9.



Způsob přemlouvání by měl být neutrální, spíše monotónní, neemotivní a stabilní v hlasitosti.

Ve druhé etapě také dochází k prvotní adaptaci systému na hlasová specifika stínového mluvčího. Průběh adaptace, její aktuální fáze a množství dat potřebných pro další část adaptace se zobrazuje uživateli na obrazovce. Celý proces se skládá ze tří fází, po dokončení poslední části je adaptace ukončena a automaticky vypnuta. Cílem automatické adaptace je zlepšení míry důvěry a přesnosti rozpoznávání u konkrétního uživatele systému.<sup>52</sup>

Zpočátku může plná rychlost videa činit při doslovném přemlouvání stínovému mluvčímu potíže. Z tohoto důvodu jsou všechny nahrávky určené pro druhou fázi tréninku k dispozici ve dvou verzích. V začátcích tréninku druhé fáze si může stínový mluvčí vybrat k přemlouvání video zpomalené na 70 % rychlosti originálního vysílání. Postupně pak začíná kandidát na stínového mluvčího doslovně přemlouvat nahrávky i v plné rychlosti původního vysílání.

V Trenažéru stínových mluvčích je druhá etapa tréninku označena jako *„fáze 2 – referenční trénink komentování s přepisem do titulků a využitím referenčního textu – přesné přemlouvání.“*

## 5.7 Třetí fáze tréninku

Třetí etapa tréninku částečně simuluje tvorbu skrytých titulků v živém vysílání, protože již zahrnuje některé kroky, které umožňují stínovému mluvčímu upravovat výsledné titulky, ve třetím stadiu tréninku ještě bez doplňování interpunkce. Klíčovým bodem této fáze tréninku je ale

---

<sup>52</sup> TAMTÉŽ, s. 9.

zlepšování schopnosti stínového řečníka přemlouvat zvukovou stopu pořadu zjednodušeným a srozumitelným způsobem. Na rozdíl od druhé fáze již kandidát nepřemlouvá sledovanou debatu doslovně, ale má možnost řeč parafrázovat svými slovy. Uživatel má v této etapě tréninku možnost vidět míru důvěry reflektující kvalitu jeho hlasového projevu. Na rozdíl od druhé fáze tréninku již není možné zobrazit přesnost rozpoznávání, protože videa určená pro třetí stadium tréninku již nemají přiložena referenční text.<sup>53</sup> Absence referenčního textu komplikuje ohodnocení přemluveného textu.<sup>54</sup>

V této fázi výcviku již stínový řečník využívá některých klávesových povelů. Jedná se zejména o rozlišování jednotlivých řečníků, které se ve výsledných titulcích divákovi zobrazí odlišnou barvou. Pomocí dvou kláves odlišuje stínový mluvčí dva moderátory. Po prvním stisknutí klávesy odliší prvního moderátora (M), opětovným stiskem stejné klávesy pak moderátora druhého (M2). Stejně tak činí i u rozlišování hostů (H a H2).

Protože šedá slova dosud nebyla zpracována do výsledných titulků a odeslána ke zobrazení, lze je smazat, obsahují-li nesprávně rozpoznaný text. Povelem pro smazání šedých slov je stisk přednastavené klávesy Backspace. Tento příkaz je možno provést také pomocí hlasového příkazu „*Smazat slovo*“, který odstraní šedá, dosud neodeslaná, slova. Poté, co stínový řečník chybně rozpozná slova jedním ze způsobů<sup>55</sup>

---

<sup>53</sup> PRAŽÁK a kol., *Four-phase re-speaker training system*, s. 219.

<sup>54</sup> Hodnocením výsledných titulků se bude zabývat kapitola 6.

<sup>55</sup> Během výcviku i reálného titulkování je preferován způsob ovládní systému pomocí klávesových zkratk, který výrazně urychluje proces zobrazení výsledného skrytého titulku divákovi.

smaže, musí promluvu do systému nadiktovat znovu. Šedá slova se do vysílání odesílají automaticky, pokud tedy nebude stínový mluvčí po delší dobu do systému diktovat, odesílá je stiskem klávesy Enter, aby nedocházelo ke zbytečnému zpoždování zobrazení titulku.

System automatického rozpoznávání řeči disponuje třemi hlasovými příkazy, které může stínový řečník během přemlouvání užít, avšak nemůže je nahradit zkratkami klávesovými, jako je tomu u příkazu „*Smazat slovo*“. Jedná se o tři příkazy, které stínový řečník použije v případě, kdy je nutno, například uprostřed věty, rozpoznat slovo začínající velkým písmenem. V tomto případě pak použije povel „*Velké písmeno*“, a slovo po tomto povelu následující se rozpozná s velkým písmenem. Obdobně funguje i příkaz „*Malé písmeno*“. Číslo systém automatického rozpoznávání řeči obvykle zapisuje číslicí. V případě, že je nutné číslo zapsat slovem, je k dispozici hlasový povel „*Číslo slovy*“.<sup>56</sup>

Třetí fáze tréninku také umožňuje stínovému mluvčímu vkládat do slovníku během rozpoznávání nová slova. Stínový řečník má dvě možnosti, jak nové slovo do slovníku vložit. Děje se tak buď stiskem tlačítka pro přidávání slov, které se nachází na liště programu, nebo stiskem klávesy F9. Při procesu vkládání dojde k přerušení přenosu zvukového signálu přes mikrofon do systému, aby nedocházelo k rozpoznávání ruchu způsobeného stiskem kláves na klávesnici při zadávání slova do systému. Proces vkládání nového výrazu je nutno co nejvíce urychlit, protože přestože mikrofon je během editace nového slova odpojen, nahrávka (potažmo živé vysílání) běží i nadále. Poté, co uživatel systému slovo napíše na klávesnici, může si ještě do následujícího řádku přidat výslovnost (i několik výslovnostních možností), což zvýší úspěšnost

---

<sup>56</sup> PRAŽÁK, *Trenažér stínových mluvčích*, s. 14–15.

rozpoznání přidaného slova. Editace výslovnosti je důležitá zejména u cizích slov. Pokud by stínový mluvčí do systému nedopsal u cizího slova výslovnost, musel by přidané slovo v celém průběhu přemlouvání vyslovovat v takovém tvaru, v jakém bylo přidáno do slovníku.<sup>57</sup>

Během třetí fáze tréninku se zpočátku využívají nahrávky pořadu Hyde Park, který se vyznačuje nižší rychlostí mluveného projevu diskutujících a menším počtem hostů. Později trénují kandidáti také na nahrávkách pořadu Otázky Václava Moravce, kde je mluvený projev výrazně rychlejší a objevuje se zde více řečníků, jejichž mluva se často překrývá.

Třetí fáze tréninku je v Trenažéru stínových mluvčích označena jako *„fáze 3 – volný trénink komentování s přepisem do titulků.“*

## 5.8 Čtvrtá fáze tréninku

Poslední fáze tréninku téměř plně simuluje tvorbu skrytých titulků v živém vysílání. Zahrnuje všechny kroky, které se stínový řečník naučil vykonávat během přemlouvání ve třetí etapě tréninku, parafrázování zvukové stopy pořadu, rozlišování mluvčích, opravu a odesílání titulků a přidávání nových slov do rozpoznávacího slovníku systému. Ve čtvrté etapě tréninku jsou výše zmíněné schopnosti rozšiřovány o vkládání interpunkce do rozpoznávaného textu pomocí zvolených kláves na klávesnici počítače. Do textu stínový mluvčí přidává tečku, čárku a otazník. Klávesy, pomocí kterých stínový mluvčí interpunkci do textu vkládá, si může v systému nastavit individuálně. Vložení interpunkce na správné místo výrazně zlepšuje čitelnost výsledného titulku. Důležité je dbát na to, aby byla interpunkce vložena v mezislovní pauze, nikoliv v průběhu

---

<sup>57</sup> TAMTÉŽ, s. 10–12.

vyslovování slov. Pokud je interpunkce vložena během procesu vyslovování slova, je výraz rozdělen do dvou nesmyslných částí.<sup>58</sup> Pokud stínový mluvčí potřebuje do rozpoznávaného textu vložit i jinou interpunkci, než kterou lze přidat pomocí stisku určené klávesy, vkládá ji slovním povelům. Jedná se například o přidání dvojtečky, pomlčky nebo lomítka.

V Trenažéru stínových mluvčích je čtvrtá etapa tréninku označena jako *„fáze 4 - volný trénink komentování s přepisem do titulků se zadáváním interpunkce.“*

## **5.9 Stínové přemlouvání živého vysílání**

Pokud uchazeč na pozici stínového mluvčího projde všemi fázemi tréninku a procentuální ohodnocení všech tří stupňů hodnocení jeho skrytých titulků vytvářených v systému automatického rozpoznávání řeči překračuje 95 %, je kandidát schopen opatřovat skrytými titulky i živé vysílání České televize.

Nový stínový mluvčí zpravidla dostává výkonnější přenosný počítač, který urychluje výkon systému automatického rozpoznávání řeči. Na rozdíl od tréninku, kde má stínový mluvčí k dispozici pouze jeden jazykový komplet, jsou v počítači pro živé titulkování nainstalovány komplety pro všechny pravidelně titulkované pořady. Zpočátku noví stínoví mluvčí přemlouvají zejména sportovní přenosy, které jsou specifické tím, že se nepřemlouvá vše, co bylo ve vysílání řečeno. Později začínají přemlouvat všechny pořady, včetně obtížnějších politických debat.

---

<sup>58</sup> PRAŽÁK a kol., *Four-phase re-speaker training system*, s. 219–220.

Aby mohl stínový mluvčí opatřovat živé vysílání skrytými titulky, musí být jeho přenosný počítač propojen s anténním kabelem, který přenáší do počítače zvuk i obraz přemlouvání pořadu. Pomocí internetu je také počítač propojen s titulkovacím serverem, který výsledné titulky odesílá do vysílání. Stínové přemlouvání živého vysílání využívá aplikace Trenažeru stínových mluvčích, který je po prvním připojení k titulkovému serveru doplněn o tlačítko „Titulkovat“. Po navázání spojení s titulkovacím serverem, připojení se k televiznímu vysílání a stisku tlačítka „Titulkovat“ si stínový mluvčí vybírá z nabídky televizní stanici, na které bude vysílán pořad určený k přemlouvání. Protože je před samotným živým titulkováním nejdříve nutno vyzkoušet, zda systém automatického rozpoznávání funguje bezchybně, má stínový řečník k dispozici tlačítko „Zkouška“, po jehož stisku začne do systému diktovat libovolný text a přesvědčí se tak, zda systém rozpoznávání řeči funguje. Zároveň tento zkušební diktát může pomoci stínovému mluvčímu, aby se před samotným přemlouváním tzv. rozmluvil.

Klíčovým rozdílem mezi vytvářením titulků v Trenažeru stínových mluvčích a v živém vysílání je třívteřinový odstup mezi zvukem a obrazem, který stínový řečník vidí na obrazovce počítače. Zvuk přichází stínovému mluvčímu do sluchátek dříve z toho důvodu, aby se zkrátila zhruba pětivteřinová časová prodleva mezi vyslovením promluvy do počítače a zobrazením titulků jejich uživatelům. Třívteřinový rozdíl mezi zvukem a obrazem by neměl stínovým mluvčím činit žádné větší potíže. Hlavním úkolem během stínového přemlouvání je orientovat se na zvukovou stopu pořadu, obraz slouží pouze pro rychlou orientaci v dění ve vysílání, protože primárním úkolem stínového mluvčího je kontrola rozpoznávaného textu, nikoli sledování děje na obrazovce.

Během přemlouvání živého vysílání se může stát, že stínový mluvčí například nerozumí dialogu řečníků, je špatné spojení, nebo obraz není

doprovázen zvukem. V tomto případě má k dispozici tlačítko „*Vložit text*“, po jehož stisknutí se zobrazí nabídka předdefinovaných textů, které může stínový mluvčí vložit do rozpoznávaného textu. Zároveň, pokud je to nutné, může do řádku napsat text vlastní, který se následně přidá k rozpoznávanému textu, nikoliv však do slovníku.

Po skočení přemlouvání je na závěr skrytých titulků pomocí tlačítka „*Podpisový titulek*“ vložen text „*Živé titulky ve spolupráci se ZČU. Inovace podpořená TA ČR.*“, který je na konci doplněn o jméno stínového mluvčího, který pořad přemlouval, a soubor se skrytými titulky je odeslán pracovníkům Katedry kybernetiky pomocí tlačítka „*Odeslat titulky*“, aby nedocházelo ke zbytečnému zaplňování paměti titulkovacího serveru.

## 6 HODNOCENÍ VÝSLEDNÝCH TITULKŮ

V úvodu této kapitoly bude uveden způsob, jakým je hodnocena přesnost rozpoznávání řeči. V další kapitole pak bude představen způsob, kterým lze hodnotit titulky vyhotovené za pomoci stínového přemlouvání. Jedná se o tzv. „třístupňové hodnocení“, je využíváno při vyhodnocování stínového přemlouvání na Katedře kybernetiky Západočeské univerzity.

### 6.1 Hodnocení rozpoznávání řeči

Zhodnocení přesnosti rozpoznávání řeči (Acc) spočívá v porovnání rozpoznávaného textu s jeho referenčním přepisem. Pro výpočet je zapotřebí mít k dispozici počet slov referenčního přepisu (N), počet slov, která chybí v rozpoznávaném textu (D), počet slov, která jsou v rozpoznávaném textu navíc (I) a počet slov, která se v textech neshodují. Výsledek v procentech pak ukazuje, nakolik rozpoznávaný text odpovídá jeho referenčnímu přepisu.

Vzorec pro výpočet: 
$$\text{Acc} = \frac{N - D - I - S}{N} \times 100 \%$$

Přesnost rozpoznávání je vyhodnocována automaticky ve druhé fázi tréninku stínových mluvčích, kdy dochází k doslovnému přemlouvání zvukové stopy. Další etapy tréninku jsou pak hodnoceny odlišným způsobem, protože nahrávky nedisponují textovým přepisem promluvy.

### 6.2 Hodnocení stínového přemlouvání na KKY ZČU

Protože během přemlouvání stínový mluvčí promluvu často zkracuje a parafrázuje, nediktuje tedy zvukovou stopu do systému doslovně, je nutné rozpoznávaný text vyhodnocovat pomocí jiných metod. Pro



ohodnocování výsledných titulků byl na Katedře kybernetiky vytvořen třístupňový model, který v procentech vyjadřuje schopnost stínového mluvčího vyjádřit myšlenky řečníků v přemlouvaném pořadu. Využívá se zejména během procesu tréninku, kdy kandidáti na stínové mluvčí v pravidelných intervalech přemlouvají určené nahrávky, které jsou následně vyhodnoceny odborným pracovníkem Katedry kybernetiky. Výsledky pak ukazují budoucímu stínovému řečníkovi, zda se jeho schopnosti zlepšují. Za pomoci třístupňového systému hodnocení jsou také vyhodnocovány titulky, které jsou vyhotovovány stínovými řečníky v živém vysílání.<sup>59</sup>

### 6.2.1 První úroveň hodnocení

V první úrovni je vyhodnocována přesnost rozpoznávání za pomoci stejného vzorce jako v případě, kdy je k dispozici referenční přepis textu. Protože ale textový přepis promluvy k dispozici není, je rozpoznáný text porovnáván se stejným textem, který byl opraven podle toho, co stínový mluvčí opravdu řekl za užití pouze minimálního počtu oprav. Neopravují se drobné chyby, které nemění význam, a přeřeknutí, která byla opravena. Do výsledku se nezapočítávají změny řečníků a v případě hodnocení čtvrté fáze tréninku a titulků vyhotovených v živém vysílání ani interpunkční znaménka.<sup>60</sup>

---

<sup>59</sup> PRAŽÁK a kol., *Novel Approach to Live Captioning Through Re-speaking: Tailoring Speech Recognition to Re-speaker's Needs*, s. 1373–1374.

<sup>60</sup> TAMTÉŽ, s. 4.

## 6.2.2 Druhá úroveň hodnocení

Ve druhé úrovni je hodnocena syntaktická přesnost vět. Dochází k porovnání rozpoznaného textu z první úrovně hodnocení a stejného textu, který byl opraven takovým způsobem, aby obsahoval pouze gramaticky a syntakticky správné věty za užití minimálního počtu oprav. Upravována je každá věta zvlášť bez ohledu na kontext. Nejčastěji dochází k opravám chybného slovosledu.<sup>61</sup>

## 6.2.3 Třetí úroveň hodnocení

Ve třetí fázi je hodnocena obsahová přesnost. Přemlouvaný pořad je rozdělen do částí podle jednotlivých řečníků. K vyhodnocování je využit text opravený ve druhé úrovni hodnocení. Každá z jednotlivých částí je pak obodována na stupnici od jedné do tří, a to následujícím způsobem. Tři body, tedy nejvyšší ohodnocení, dostane stínový řečník v případě, že přemluvil vše podstatné. Tři body získá i v případě, pokud něco vypustil, například nepodstatnou řečovou výplň.

Dva body jsou stínovému mluvčímu uděleny v případě, že přemluvil vše podstatné, ale nepřesně nebo nezřetelně. A to i v případě, že se nejasně vyjádřil i původní mluvčí, protože úkolem stínového mluvčího je v případě nepřesností mluvu upravit takovým způsobem, aby byla jasná a pochopitelná.

Jeden bod je při hodnocení udělen, pokud byla v promluvě vynechána nějaká podstatná informace.

---

<sup>61</sup> PRAŽÁK a kol., *Novel Approach to Live Captioning Through Re-speaking: Tailoring Speech Recognition to Re-speaker's Needs*, s. 1373.

Body nejsou přiděleny v případě, že došlo zcela k vynechání informace, ke změně významu nebo bylo přidáno chybné sdělení.

K celkovému vyhodnocení třetí úrovně pak dochází za pomoci následujícího vzorce:

3. úroveň =  $\frac{\text{součet bodového hodnocení jednotlivých částí}}{3 \times \text{počet jednotlivých částí}} \times 2. \text{ úroveň}$   
Hodnocení nižší úrovně se vždy promítá ve výsledku hodnocení vyšší úrovně. Znamená to tedy, že hodnocení třetí úrovně hodnocení zahrnuje chyby z prvních dvou fází hodnocení.<sup>62</sup>

Z informací uvedených v této kapitole je vidět, že způsob hodnocení stínového přemlouvání užívaný na Katedře kybernetiky je částečně subjektivní, zejména způsob ohodnocování sémantického obsahu titulků.

---

<sup>62</sup> PRAŽÁK a kol., *Novel Approach to Live Captioning Through Re-speaking: Tailoring Speech Recognition to Re-speaker's Needs*, s. 1373.

## 7 KOMPARATIVNÍ LINGVISTICKÁ ANALÝZA KOMUNIKÁTŮ VYTVÁŘENÝCH STÍNOVÝMI MLUVČÍMI

V následující kapitole bude představena komparativní lingvistická analýza výsledných komunikátů, které byly vytvořeny stínovými mluvčími ve druhé a čtvrté fázi tréninku i v živém vysílání.

### 7.1 Metodika výzkumu

Během výzkumu byla analyzována data tří stínových mluvčích, kteří se v současné době aktivně podílejí na titulování živých přenosů České televize. Jedná se o nahrávky dvou žen (DT a DZ) a jednoho muže (PP) z vybraných fází tréninku a živého vysílání.

Získaná data byla analyzována v programu MegaWord<sup>63</sup> vyvinutého na Katedře kybernetiky Západočeské univerzity ve spolupráci s její spin-off firmou, společností SpeechTech, s. r. o. Program zobrazuje rozpoznaný text a současně přehrává hlas stínového mluvčího, který diktoval do systému automatického rozpoznávání řeči.

Analýza probíhala komparací rozpoznávaného textu s projevem stínového mluvčího. Tento způsob analýzy umožnil zjištění chyb systému a chyb, které do procesu rozpoznávání vnáší během diktování stínoví řečníci.

Analyzována byla videa z druhé fáze tréninku, kdy se potenciální stínoví mluvčí učí přizpůsobovat svůj projev systému automatického rozpoznávání řeči, a videa ze čtvrté fáze tréninku, ve které stínoví řečníci doplňují do rozpoznávaného textu interpunkční znaménka. Počet analyzovaných nahrávek z druhé etapy tréninku se u jednotlivých

---

<sup>63</sup> Verze 7.0.2.948.

stínových mluvčích liší. Je to z toho důvodu, že každému ze tří stínových mluvčích trval trénink ve druhé fázi odlišnou dobu, každý tedy přemluvil různý počet nahrávek, které byly následně Katedrou kybernetiky poskytnuty pro potřeby analýzy této diplomové práce. Ze čtvrté etapy tréninku byly u každého ze tří stínových mluvčích analyzovány shodně tři nahrávky, a to různých pořadů.

Chyby byly následně klasifikovány jako chyby systému automatického rozpoznávání řeči, chyby jednotlivých stínových řečníků (chyby individuální) a chyby nadindividuální, které se opakovaly u více stínových mluvčích. Za chyby systému jsou považovány všechny nedostatky, které se objevily v rozpoznaném textu i přes skutečnost, že stínový mluvčí vyslovil slovo nebo posloupnost slov zcela zřetelně.

Během výzkumu byl důraz kladen zejména na to, jakým způsobem je systém schopen rozpoznávat projev stínových mluvčích, a zejména schopnost stínových řečníků korektně vyslovovat slova tak, aby je systém rozpoznal. U komunikátů vytvořených ve čtvrté fázi tréninku a v živém vysílání pak byla sledována schopnost vytvářet srozumitelné, syntakticky korektní věty, a správné vkládání interpunkce, které může být klíčové pro plnohodnotné pochopení výsledného skrytého titulku.

Do analýzy nebyl zahrnut sémantický obsah nahrávek. Bylo tak učiněno z toho důvodu, že významovou stránku stínový mluvčí zvládá přemlouvat postupně během třetí a čtvrté fáze tréninku. S přibývajícím počtem přemluvených nahrávek se zvyšuje nejen schopnost koncentrace stínového mluvčího, ale také způsobilost zachytit a nadiktovat do systému sémantický obsah promluvy na obrazovce takovým způsobem, aby významová stránka výsledného titulku odpovídala alespoň z 95 % sémantickému obsahu televizního vysílání. Příklad vývoje schopnosti přemlouvat obsah pořadu je uveden v příloze I.

## **7.2 Analýza komunikátů vytvořených ve druhé fázi tréninku**

V následující podkapitole bude uvedena analýza výsledných komunikátů vytvořených ve druhé fázi tréninku, kde budoucí stínovní mluvčí diktují do systému doslovně to, co slyší. Úkolem stínového řečníka je přizpůsobit svůj projev potřebám systému takovým způsobem, aby byla dosažena co nejvyšší přesnost rozpoznávání. Z tohoto důvodu byly pro první část analýzy vybrány komunikáty z této etapy tréninku, protože je zde poprvé zapojen do činnosti systém automatického rozpoznávání řeči.

Vysvětlení grafického záznamu analýzy je následující: jako první je zaznamenáno slovo nebo slovní spojení, které se nacházelo v originálním znění pořadu, dále následuje rozpoznáný řetězec slov a poté objasnění, z jakého důvodu k uvedenému rozpoznání došlo.

U zpomalených nahrávek lze předpokládat, že by měl být stínový mluvčí schopen do systému diktovat zřetelně a dle pravidel správné výslovnosti, protože není nucen zrychlovat svůj projev, aby stačil tempu nahrávky v plné rychlosti. U nahrávek v plné rychlosti lze očekávat, že se vyskytnou častěji ortoepické chyby a slova nebudou od sebe striktně oddělována takovým způsobem, aby je systém dokázal bez problémů rozpoznat. Předpokládají se také častější chyby v koncovkách určitých slovesných tvarů, které mohou být, vzhledem k vyšší rychlosti projevu, vyslovovány nezřetelně.

### **7.2.1 Komunikáty vytvořené stínovou mluvčí DT**

Stínová mluvčí DT se vyznačuje zřetelným hlasovým projevem, bez vad a dialektických znaků. Z druhé fáze tréninku bylo analyzováno celkem pět nahrávek, čtyři zpomalené na 70 % rychlosti původního vysílání a jedna nahrávka v plné rychlosti původního vysílání.

### 7.2.1.1 Nahrávky zpomalené na 70 % rychlosti původního vysílání

Jedná se o úseky z pořadu Otázky Václava Moravce o různé délce v rozmezí tří až sedmi minut.

Za chyby individuální lze v těchto nahrávkách považovat:

- nezřetelné vyslovování (zejména koncovek), užití chybných koncovek  
*Není jisté* se rozpoznalo<sup>64</sup> jako *nyní jste* – výslovnost koncovky -y zněla jako koncovka -e.

reformy → reformu - řekla *reformu*

ono → ano - výslovnost o se podobala výslovnosti a

- hlasité nádechy a výdechy

k → v

hlasitý nádech → po

- pauzy během vyslovování slov, hezitační zvuky

zaváhání spojené s nádechem → pohrát

vyslovila *zak* → zákon

nominovaných → na minimum

vyslovili → jim vysloveně - pauza za předponou slovesa

---

<sup>64</sup> Slova "se rozpoznalo" budou v dalších částech práce nahrazena šipkou →.

zaváhání, zvuk podobně znějící jako že → že

co pracují → co se pracují - mezi slova vsunula hezitační zvuk

hlasitý nádech → v

hezitační zvuk → a

kauza vznikla → kauza zisk vznikla - mezi slovy byl hezitační zvuk

vám pane vicepremiére → velice projekt premiérem - udělala pauzu mezi předponou *vice-* a *premiérem*

- nedokončená slova

právník → právní - nedořekla slovo do konce

- absence mezislovní pauzy

tak já → také

- prodlužování výslovnosti hlásek, přílišný důraz na koncovky

by do poslední → by ji do poslední - protáhla výslovnost slova *by*

letošních → letošní jich - velký důraz na *ch*

motivací → motivaci ní - prodloužila výslovnost koncovky

počítali → počítali ji - kladla velký důraz na koncovku *-li*

že → že je - kvůli velkému důrazu na slovo že se navíc rozpoznalo slovo *je*

o 50. letech → o 50. let té - vyslovila *o padesátý leté*, absence hlásky *ch* a protažení samohlásky *e*

nic neudělal → nic neudělá - vyslovila [*neudělál*]



- vynechání předložek

v PS → PS

v žádném → žádném

- užívání nespisovných nebo jiných slov, slova navíc

a Jana Bursíka → a je nám musí k - vyslovila *Jana Musíka*

Bursík → Dusík - vyslovila *Dusík*

horké hlavy → horkého a hlavy - vyslovila *horké vla*

musí být navrženy → musí být návrhy že navržený - řekla *navrch že navržený*

nespokojenost → nespokojené jste - vyslovila [*nespokojést*]

rebelujete → řekla *rév rebelujete*

nepodváže → nepodám nepodváže - řekla *nepod nepodváže*

mluvím → mluvil - vyslovila krátce *mluvim* (systém nerozpozná nespisovný tvar)

země → v zemině - řekla *v země*

právníka → přát právníka - vyslovila *přát právníka*

Jiřímu Čunkovi → Jiřímu či Čunkovi - mezi obě slova vsunula slovo *či*

Za chyby systému lze považovat následující rozpoznání:

ty by → kdyby

Brusel skoupý → bruselskou prý

nepřijímat → nepřijímá

vyřčenou → listinou

reakce → reakcemi

zchladly → shlédli

sto pětáctýřicet → 100 pětáctýřiceti

hospodařením → hospodaření

že na → žena (toto ale stínový mluvčí nijak neovlivní)

neškrkali → seškrkali

či všemožné → všemožné

na růst → nárůst

desetinku → desetinkou

zaměstnanost i hospodářský → zaměstnanost je hospodářský

a ta by → zabije

ted' → TJ

let → léta

tlak i policie → tlaky policie

směřovaná → směřována

čas → část

jste → systém

Z uvedené analýzy vyplývá, že většina chyb, které se objevily v rozpoznávaných textech, byla zapříčiněna stínovou mluvčí DT a jejím způsobem diktování do systému rozpoznávání řeči. Převažují chyby způsobené hezitačními zvuky, nevhodnými pauzami uprostřed slov a prodlužováním hlásek, zejména na konci slov.

### 7.2.1.2 Nahrávka přemluvená v plné rychlosti původního vysílání

Jedná se o půlhodinový úsek z pořadu Otázky Václava Moravce. Nahrávka byla přemluvena v celé délce.

V této nahrávce byly zjištěny následující individuální chyby:

- nezřetelné vyslovování (zejména koncovek), chybné koncovky

či nepovolováním → činy nepovolováním - samohláska e zněla při vyslovení jako samohláska *i*

žalobě → žaloby - koncovka byla vyslovena chybně

naším → našim - samohláska *i* byla vyslovena krátce

je odradili → neodradili - vysloveno správně, ovšem ne dostatečně zřetelně, aby to systém správně rozpoznal

ted' → ty - vysloven byl shluk hlásek [tyd']

- hlasité nádechy a výdechy

nádech → ten

- pauzy během vyslovování slov, hezitační zvuky

ted' upřesnil → ted' u Rusů upřesnil - řečeno bylo *ted' u ru upřesnil*, systém statisticky vyhodnotil, že má rozpoznat po vyslovení hlásek *ru* slovo Rusů

hovořeno → Jacques IDS hovořeno - hezitační zvuky a poté vyslovila *hovořeno*

skupiny → skupin i - pauza uprostřed slova

mužů → muzikou - vyslovila s pauzou uprostřed slova

povolebně → kolem mě - uprostřed slova udělala pauzu

polovojenské → Polo vojenské - pauza uprostřed slova

kritizovala → s IDS kritizovala - hezitační zvuky a poté následovalo *kritizovala*

- nedokončená slova

ve všech úrovních → ve všech úrovní - hláska *ch* nebyla vyslovena

zadržných → zadržný - hláska *ch* nebyla vyslovena

- prodlužování výslovnosti hlásek, přílišný důraz na koncovky

byla skupina → byl skupinám - samohláska *a* na konci slova *skupina* byla prodloužena

akcí → akcií - vyslovila *akcí*, ale koncovka byla prodloužena

- užívání nespisovných nebo jiných slov, slova navíc

přestupkové → před Kutkové - DT to takto skutečně do systému nadiktovala

projevy → projevuje - vyslovila *projevu*

jejich souvislostech → souvislost je - vyslovila *jejich* [*souvislostjech*]

oni → on oni - tímto způsobem to bylo do systému nadiktováno

desítek tisíc korun → 1000 kde IDS korun - do systému bylo nadiktováno *tisíce kdes korun*

testují → text stojí - do systému bylo nadiktováno *textují*

připravovaný → připraveny - DT užila nespisovného výrazu [*připravovanej*], který systém nerozpozná

který → kterém - vyslovila *kterém*

Identifikovány byly následující chyby systému:

policie a policisté → policie policisté - spojka nebyla rozpoznána

pochvalu → pochvalou

radikálů → kredit

po té → poté (nelze ovlivnit)

vyrábí ty → vyrábíte

policie a bezpečnostní složky → spojka nebyla rozpoznána

pravicovým nebo levicovým → pravicový nebo levicový

státních zastupitelství → státní zastupitelství

problémy → problém

zůstalo → zůstal

abychom je → bychom - slovesný tvar rozpoznán nebyl

my máme → nemáme

říká → řekla

my to → mi to (nelze ovlivnit)

já bych → že bych

sám → nám

kritiku → kritikou

klidem → k lidem (nelze ovlivnit)

mělo → měla

co by → coby (nelze ovlivnit)

nepřisuzovali → nepřisuzoval i

Z uvedených příkladů lze vyčíst, že stínová mluvčí DT neměla při přemlouvání nahrávky v plné rychlosti problém oddělovat od sebe jednotlivá slova takovým způsobem, aby je systém bez obtíží rozpoznal. Oproti zpomaleným nahrávkám DT dbala na to, aby rozpoznávání neovlivňovaly nádechy a výdechy, nedocházelo také k vynechávání předložek. Přetrvaly ale problémy s hezitačními zvuky, nevhodnými pauzami uprostřed slov a diktování slov navíc. Naopak se neobjevily chyby v koncovkách určitých slovesných tvarů, jak bylo očekáváno v úvodu podkapitoly.

### **7.2.2 Komunikáty vytvořené stínovou mluvčí DZ**

Mluvený projev stínové mluvčí DZ nevykazuje defekty ve výslovnosti, je zřetelný, bez vad, a neobsahuje ani dialektické znaky. Z druhé fáze tréninku byla pro analýzu k dispozici jedna nahrávka zpomalené na 70 % rychlosti původního vysílání a dvanáct nahrávek v plné rychlosti původního vysílání.

### 7.2.2.1 Nahrávka zpomalená na 70 % rychlosti původního vysílání

Jedná se o nahrávku z pořadu Otázky Václava Moravce v délce 7 minut a 11 vteřin.

Za individuální lze v této nahrávce považovat následující chyby:

- nezřetelná výslovnost předložek

k jejich → jejich

v boji → boji

Za chyby systému lze považovat následující chybná rozpoznání:

tedy → ten

ty by → tedy

českou → Česko

Brusel skoupý → bruselskou ví

Jan Bursík → Jan Dusík

zelení rozpálení doběla → Zelení pro schválení doběla

U stínové mluvčí DZ se u zpomalené nahrávky, ve srovnání s DT, vyskytuje pouze jeden druh individuální chyby - nezřetelné vyslovování předložek. Do systému rozpoznávání řeči diktuje jasně a zřetelně, proto je výrazně nižší i počet chyb systému.

### 7.2.2.2 Nahrávky přemluvené v plné rychlosti původního vysílání

Analyzovány byly komunikáty vytvořené na základě nahrávek v délce od tří do třiceti minut.

Za individuální lze v těchto nahrávkách považovat následující chyby:

#### - absence předložek

v boji → boj - vyslovila *boji*, ale bez předložky

v české → české - předložka nebyla vyslovena

#### - prodlužování výslovnosti hlásek, přílišný důraz (zejména na koncovky)

a vstoupili → a to - při vyslovování kladla velký důraz na hlásky *a to*

jiným → jinými - velký důraz na hlásku *m*, který inicioval následné rozpoznání samohlásky *i*

přišel jste → slyšel jste - velký důraz na *šel jste*, systém vyhodnotil jako pravděpodobnější sloveso *slyšet* a následně ho rozpoznal

#### - nedokončená slova

vyřčenou → vyřčeno - absence samohlásky *u* při vyslovení slova

bych → by - hlásky *ch* na konci nebyla vyslovena

státních zastupitelstvích → státní zastupitelství - koncovku *ch* vynechala u obou slov

#### - pauzy během vyslovování slov, hezitační zvuky

protikrizových → oproti krizové jejich - hlasitý nádech před vyslovením slova a pauza mezi *proti* a *krizových*



v přípravě zákona → případ právě zákona - ve slově *přípravě* byla pauza

nerезignoval → rezignoval - došlo k rozpoznání antonyma, protože DZ udělala uprostřed slova pauzu

namíříme → na míříme - za předponou udělala pauzu

abychom → aby jsme - vyslovila nespisový výraz [*abysme*], mezi *aby* a [*sme*] udělala pauzu

#### - hlasité nádechy a výdechy

potom policie → potom policie do - hlasitý nádech

v boji → o boj - hlasitý nádech

#### - nezřetelné vyslovování (zejména koncovek), chybné koncovky

nim → němu - výslovnost samohlásky *i* zněla jako výslovnost samohlásky *e*

jednáním → jednání - hláska *m* byla na konci nezřetelná

zchladli → shodli - nezřetelná výslovnost

k jejich → jejich - předložka vyslovena nezřetelně

v boji → boj - předložka *i* koncovka byla vyslovena nezřetelně

dvě → je - hláska *d* byla na začátku nezřetelná

nepřijímat → nepřijme - koncovka nebyla vyslovena zřetelně

v povinných → povinných - předložka vyslovena nezřetelně

ještě je → ještě - určitý slovesný tvar vysloven nezřetelně

zopakují → zopakuje - na konci slovesného tvaru vyslovila samohlásku *e*

najevo je dává → nedává - výslovnost určitého slovesného tvaru zněla jako výraz *ne*

strategii → strategie - vyslovila slovo *strategie*

my souhlasíme → nesouhlasíme - rozpoznáno antonymum, nezřetelná výslovnost

zaznělo → zazněl - koncovka *o* vyslovena nezřetelně

vezli → veze - nezřetelná výslovnost

jak spolupracovaly → spolupracovaly - *jak* bylo nejasné, systém ho nevyhodnotil jako slovo k rozpoznání

je Ivan → Ivan - určitý slovesný tvar vysloven nezřetelně, systém ho nevyhodnotil jako slovo k rozpoznání

je dokonce → dokonce -

abychom je → bychom - určitý slovesný tvar vysloven nezřetelně

naše → naši - koncovka *e* zněla spíše jako *i*

právní → právo - vyslovila korektně začátek slova, ale koncovku již nezřetelně

#### - nejasné oddělování slov

ty by → tebe - nebyla zde jasná pauza mezi slovy

musí zde → musíte - nejasná mezislovní pauza

vy vidíte → nevidíte - absence mezislovní pauzy, rozpoznáno antonymum

#### - užívání nespisovných nebo jiných slov, slova navíc

by daň → vydání - vyslovila [vydaň]

Bursík → Dusík - nadiktovala *Dusík*

nepřijímat → nepřímo - vyslovila *nepřímat* (nespisový tvar)

boj → vývoj - vyslovila slovo *vývoj*

exekutivy → exekutivy by - rozpoznalo se navíc *by*, protože místo slova *exekutivy* vyslovila výraz *exekutivity*

a vy jste → abyste - řekla *a* [vyste]

tu běžnou → tu byly běžnou - nadiktovala navíc slovo *byly*

jsem → nejsem - DZ nadiktovala do systému antonymum

Stínová mluvčí DZ vnáší do systému během přemlouvání nahrávek v původní rychlosti nejvíce chyb svou nezřetelnou výslovností a nevhodnými pauzami uprostřed slov. Její komunikáty také vykazují chyby v důsledku nedokončených slov a chybných rozpoznání v případech, kdy klade nevhodně důraz na slabiky uprostřed slov. Naopak nevnáší do procesu rozpoznávání příliš hezitačních zvuků a nadbytečných slov. Během diktování byla ale často rozpoznávána slova opačného významu.

V těchto nahrávkách lze identifikovat následující chyby systému:

je totiž → totiž

té lobby → žaloby - výraz *žaloba* je pro systém pravděpodobnější

reakce → frakce

zchladly → shlédli

vaší → naší

čas → část

auta a jejich → auta jejich

nijak → nějak

pane Jáne → pane Já ne (nelze ovlivnit)

my připravujeme → ne připravujeme - rozpoznáno antonymum, vysloveno ale jasně a zřetelně

trh → tedy

jasné ne řekli → jasné neřekli (nelze ovlivnit)

je jejich → jejich - dvě je po sobě činí systému problém s rozpoznáváním

my v našem → nenašel

po té → poté (nelze ovlivnit)

### **7.2.3 Komunikáty vytvořené stínovým mluvčím PP**

Mluvený projev stínového mluvčího PP nevykazuje známky dialektických znaků ani závažných vad v řeči. Jeho projev se vyznačuje sklony k nezřetelné výslovnosti předložek, spojek a koncovek u určitých slovesných tvarů.

Analýza probíhala na devatenácti komunikátech vytvořených z nahrávek pořadu Otázky Václava Moravce zpomalených na 70 % původní rychlosti a na šesti komunikátech z nahrávek OVM v původní rychlosti.

#### **7.2.3.1 Nahrávky zpomalené na 70 % rychlosti původního vysílání**

Jedná se o nahrávky v délce od tří do sedmi minut.

Identifikovat lze následující individuální chyby:

- prodlužování výslovnosti hlásek, přílišný důraz (zejména na koncovky)

s → se

zcela → z zcela - velký důraz na hlásku z na začátku slova

s tím → se s tím - důraz na hlásku s

- nedokončená slova

zakazuje → zákazu - nedořekl koncovku

uskutečnit → skutečnost - na začátku nebyla vyslovena samohláska *u*

jsme chtěli → si nechtěl - nadiktoval *jsme chtěl*, absence koncovky *i*

měl na → měna - koncovka určitého slovesného tvaru *-l* nebyla vyslovena

- pauzy během vyslovování slov, hezitační zvuky

přichází → přicházejí - pokles hlasu uprostřed slova

výjimku → výjimkou - pauza před dvojhláskou *-ou*

přichází → přicházím - pokles hlasu na konci slova, který zněl jako hláska *m*

přinesl → přineslo - pokles hlasu uprostřed slova vedl k rozpoznání samohlásky *o*

na sebe nedalo dlouho čekat → na sebe nedalo toto čekat - špatně zvolené mezislovní pauzy vedly k nahrazení slova *dlouho* výrazem *toto*

řetězci → řetězcem - pokles hlasu uprostřed slova vedl k rozpoznání chybné koncovky

či modernizace → těm datům dnes roce - pauza uprostřed slova  
*modernizace*

vystavět → ty stavět - pauza za předponou

českou → v Česku - krátké zaváhání před vyslovením výrazu

vyslovili → vysvětlili - pauza uprostřed slova

rozpálení → radost pálení - pauza za předponou

ke krokům → Kypru krokům - zaváhání po vyslovení předložky

- hlasité nádechy a výdechy

nádech → k

nádech → že

nádech, hluk → i v jinam

- nezřetelné vyslovování (zejména koncovek), chybné koncovky

doběla → nebyla - nezřetelná výslovnost

vyřčenou → vyřčeno - koncovka *-u* nebyla zřetelně vyslovena

výběru → výběr - koncovka *u* na konci nezřetelně vyslovena

Českou → Česko

a o jakých → velkých - nezřetelně vysloveno, systém to vyhodnotil jako jedno slovo

mu to → motor - výslovnost se podobala spíše rozpoznávanému slovu

vzata → zato - nezřetelné vyslovení hlásky *v* na začátku slova vedlo k rozpoznání statisticky nejbližšího výrazu *zato*

nebudou → nebudu - samohláska *o* nezřetelně vyslovena

úřednického → úřednické - nezřetelně vyslovená koncovka

kabinetu → kabinet - nezřetelně vyslovená koncovka

vzniklo v době → vzniklo době - předložka nezřetelně vyslovena

za zelené → zaslané - nezřetelná výslovnost

požádal jsem → prožil jsem - nezřetelná výslovnost

bodovém → budovám - samohláska *e* vyslovena jako samohláska *a*

zchladly → skladby - hláska */* v koncovce zněla jako hláska *b*

vy → by - výslovnost hlásky *v* zněla jako výslovnost hlásky *b*

a přišel jste o židli → nepřišel jste prožili - nezřetelně vysloveno

zrychlit chod → rychle odchod - nezřetelná výslovnost hlásky *z* na začátku

a jiným → nyní - nezřetelně vysloveno

ono → ano - samohláska *o* se na podobala svou výslovností samohlásce  
*a*

- užívání nespisovných nebo jiných slov, slova navíc

jako → krok - místo slova *jako* bylo do systému nadiktováno slovo *krok*

bodovém → Borovan - do systému byl nadiktován výraz *borovém*

by lepší → byla či - nadiktoval nespisovné slovo [*lepčl*]

v → do - vyslovil shluk hlásek [*vdo*]

jste → stav - vyslovil shluk hlásek [*ste*]

státu → stav - nadiktoval *státní to*

jednak → jednat - nadiktoval výraz *jednat*

je v → jel do - nadiktoval výrazy *je do*, což vedlo k rozpoznání statisticky nejbližšího slovesného tvaru *jel*

sto pětatřiceti → 100 pětatřiceti - vhodné je nadiktovat *sto třiceti pěti*, rozpoznáno je 135

zchladly → kladly → nadiktoval kladly

- nejasné oddělování slov

ty by → tedy - nejasná pauza mezi oběma slovy

a o → o - nejasná pauza mezi oběma samohláskami

vlastně jako → hlas nějakou - nezřetelná mezislovní pauza vedla k rozpoznání zcela jiného řetězce slov

Z výše uvedené analýzy lze vyčíst, přestože se jednalo o přemlouvání nahrávek v pomalejší rychlosti, že stínový mluvčí PP měl oproti DT a DZ větší problémy přizpůsobit svůj projev systému automatického rozpoznávání řeči. Největší problémy činily systému nezřetelná výslovnost výrazů a koncovek zejména u určitých slovních tvarů a pauzy a drobná zaváhání během vyslovování slov.

V těchto nahrávkách lze identifikovat následující chyby systému:

Brusel skoupý → bruselskou i

mu to → motto

boj → boje



boj → boji

lobby → volbě

prošel → přšelo

za celou částku daně → Kodani

Brusel skoupý → bruselskou ví

stažením → stařenky

Bursík → Dusík

levice → lavice

doběla → nebyla

v tomto → Waltru tom

ty by → kdyby

té lobby → to aby

my si myslíme → a SMS víme

a Jana → a je na

Brusel skoupý → bruselskou prý

ještě je → ještě

pod → bod

### 7.2.3.2 Nahrávky přemluvené v plné rychlosti původního vysílání

Analyzovány byly komunikáty vytvořené z nahrávek v délce od tří do pěti minut.

Identifikovat lze následující individuální chyby:

#### - nejasné oddělování slov

počítá právě → počítáte - nejasná mezislovní pauza

a přišel jste o židli → přestože - absence slovních pauz

to dělá → tedy - nejasná mezislovní pauza

jde o větší → video vyčíst - absence mezislovní pauzy mezi určitým slovesným výrazem a předložkou

a přišel jste o židli → včelstev že také - absence mezislovních pauz

si stanovila → sestavila

#### - užívání nespisovných nebo jiných slov, slova navíc

by sociální → vysoce - nadiktoval vysoce

jde o větší → vyčíst - výraz větší vysloven jako *vyčíst*

#### - nezřetelné vyslovování (zejména koncovek), chybné koncovky

a poslanci → a zase - výraz poslanci byl vysloven nezřetelně

se jim nelíbí → že se nelíbí - výraz *jim* byl nezřetelně vysloven

raději → radě - koncovka *-ji* vyslovena nezřetelně

ještě je → ještě - určitý slovesný tvar nerozpoznán, nebyl vysloven zřetelně

a jiným budou → nebudou - rozpoznáno antonymum z důvodu nejasné výslovnosti

který je v → který v - určitý slovesný tvar vysloven nezřetelně

sliby → sebe - nezřetelně vysloveno

2011 → v Tisé 11 - výrazy *dva tisíce* byly nezřetelně vysloveny, což vedlo k rozpoznání slov *v Tisé*

- pauzy během vyslovování slov, hezitační zvuky

zveřejněním podrobných kroků → se zřejmě ní prostě podrobnější kroku - PP ztrácel dech, mezislovní pauzy na nevhodných místech vedly k rozpoznání zcela jiných slovních řetězců

a teprve v důsledku toho → a teprve dostat proto - nevhodné mezislovní pauzy

porovnávejte to s tou → porovná věťe dostalo - hezitační pauza uprostřed výrazu *porovnávejte*

zhruba o jedno → pro oba ledna - zaváhání při vyslovení výrazu *zhruba*

Při přemlouvání nahrávek v plné rychlosti měl stínový mluvčí PP problémy stačit tempu původního řečníka, což vedlo k drobným zaváháním a nevhodně voleným mezislovními pauzám, které způsobovaly problémy při rozpoznávání. Za další z individuálních nedostatků stínového řečníka PP lze považovat neschopnost zřetelně od sebe oddělovat jednotlivé výrazy takovým způsobem, aby byl systém automatického rozpoznávání řeči schopen rozpoznat jednotlivá slova bez chyb .

Za chyby systému lze v těchto nahrávkách považovat:

v povinných výdajích → povinných výdajích

v PS → PS

do dvou měsíců → do měsíců

místo → město

tak já → také

a za sliby → zaslat

ke sto pětatřiceti → zastupitel 30 miliardám

že na → žena

### **7.3 Shrnutí chyb zjištěných z analýzy komunikátů vyhotovených ve druhé fázi tréninku**

V následující podkapitole bude shrnuta a kvantifikována analýza dat komunikátů všech tří stínových mluvčích ze druhé fáze tréninku. Blíže budou také popsány nadindividuální chyby, které se objevily v komunikátech alespoň dvou stínových mluvčích.

#### **7.3.1 Nadindividuální chyby vnášené do systému stínovými mluvčími**

##### **1) nezřetelná výslovnost výrazů**

Nesprávná výslovnost vedla k nejčastějším chybám v rozpoznávání u všech tří stínových mluvčích. Jednalo se zejména o nejasnou výslovnost

koncovek u určitých slovesných tvarů, která nejčastěji vede k rozpoznání určitého slovesného tvaru určeného pro kategorii jiné osoby.

Dále lze do této kategorie uvést nezřetelnou výslovnost předložek a spojek, zejména těch, které jsou pouze jednohláskové. Jedná se zejména o nehláskové předložky *v*, *k*, *z* a *s*. Problémy při rozpoznávání činily systému také slabičné *o*, *u* a spojky *a*, *i*. Stínový mluvčí během diktování do systému předložku či spojku často vyslovili, ovšem nepřizpůsobili způsob jejich nadiktování potřebám systému rozpoznávání řeči. Na předložku nebo spojku nebyl při artikulaci kladen takový důraz, aby je systém rozpoznal.

Příklady:

přinesl → přineslo

přichází → přicházím

v povinných → povinných

z letos → letos

## **2) nezřetelné oddělování jednotlivých slov**

Zejména při přemlouvání nahrávek v plné rychlosti se objevoval u všech tří stínových mluvčích problém jasně od sebe oddělovat jednotlivá slova. V mnoha případech to bylo způsobeno neschopností stačit tempu projevu původního řečníka, nebo ztrátou dechu. Nejasné mezislovní pauzy vedou k chybám v rozpoznávání. Neschopnost řádně od sebe oddělovat slova pak může ovlivňovat i vkládání interpunkce v poslední části tréninku a v živém vysílání, a to z toho důvodu, že interpunkci je nutno do rozpoznávaného textu vkládat v mezislovní pauze.

Příklady:

a přišel jste o židli → přestože

to dělá → tedy

ty by → tedy

### **3) pauzy uprostřed artikulace slova**

Všichni tři stínoví mluvčí vykazovali tendence dělat během artikulace pauzy uprostřed slov. Udělat pauzu uprostřed artikulace slova je nepřijatelné, a to z toho důvodu, že to vede k závažným chybám v rozpoznávání, často výrazů, které v kontextu nedávají smysl.

Příklady:

porovnávejte to s tou → porovná věť dostalo

v přípravě zákona → případ právě zákona

vám pane vicepremiére → velice projekt premiérem

### **4) nedokončování slov**

V komunikátech všech tří stínových mluvčích se objevily chyby v rozpoznání způsobené nedokončením slov. Jednalo se zejména o koncovky u diktovaných výrazů. Příklady jsou zde uvedeny následujícím způsobem: první výraz je slovo, které mělo být rozpoznáno, výraz za šipkou je slovo, které stínový mluvčí do systému nadiktoval.

Příklady:

vyřčenou → vyřčeno

právník → právní

zakazuje → zakazu

uskutečnit → skutečnit

### 5) diktování zcela odlišných slov

V rozpoznávaných komunikátech všech tří stínových mluvčích byly identifikovány chyby v rozpoznání z důvodu nadiktování zcela odlišných slov - hovorových výrazů, nesmyslných řetězců hlásek nebo slov, která slovník systému neobsahuje. Největší problémy činily shluky nesmyslných řetězců hlásek, které systém rozpoznává jako statisticky nejbližší slova a která v kontextu rozpoznávaného textu nedávají smysl.

Příklady:

by daň → vydání - nadiktováno [vydaň]

jsem → nejsem - nadiktováno antonymum

bodovém → Borovan - nadiktován výraz *borovém*

jako → krok - nadiktováno slovo *krok*

a Jana Bursíka → a je nám musí k - nadiktováno *Jana Musíka*

právníka → přát právníka - nadiktováno *přát právníka*

### 6) prodlužování výslovnosti hlásek

Stínoví mluvčí DT a PP vnášejí do systému chyby způsobené přílišným prodlužováním výslovnosti hlásek. Systém slova s takovou výslovností s velkou pravděpodobností rozpozná, nicméně často je rozpozná s chybnou koncovkou, nebo například se samohláskou navíc.

Příklady:

s → se

zcela → z zcela - velký důraz na hlásku z na začátku slova

byla skupina → byl skupinám - samohláska a na konci slova *skupina* byla prodloužena

akcí → akcií - vyslovila *akcí*, ale koncovka byla prodloužena

### 7) hezitační zvuky, hlasité nádechy, hluk

Všichni tři stínoví mluvčí vnášeli do svých komunikátů chyby z důvodu hezitačních zvuků, nesprávného dýchání a hluku během diktování. V mnoha případech byla rozpoznána v kontextu nesmyslná slova nebo spojky a předložky, které do rozpoznávaného textu v danou chvíli nepatřila.

nádech → k

potom policie → potom policie do

ted' upřesnil → ted' u Rusů upřesnil - řečeno bylo *ted' u ru upřesnil*, systém statisticky vyhodnotil, že má rozpoznat po vyslovení hlásek *ru* slovo Rusů

hovořeno → Jacques IDS hovořeno - hezitační nádech, hluk → i v jinam

### 7.4 Analýza komunikátů vytvořených ve čtvrté fázi tréninku

V následující podkapitole budou analyzovány komunikáty vytvořené ve čtvrté etapě tréninku, kdy stínoví mluvčí vkládají do rozpoznávaného textu interpunkční znaménka. U každého ze tří stínových mluvčích byl analyzován jeden komunikát.

Grafická podoba analýzy je následovná. Z rozpoznávaného textu byly vyňaty věty (vyznačeny kurzívou), které obsahovaly chybná rozpoznání



nebo nesprávně vložená interpunkční znaménka (podtrženo), a každá z těchto chyb byla následně okomentována. Předpoklad pro analýzu byl takový, že chyby v rozpoznání budou již eliminovány a častěji se budou projevovat chyby ve vkládání interpunkce.

#### 7.4.1 Komunikát vytvořený stínovou mluvčí DT

Analyzován byl komunikát vytvořený na nahrávce pořadu Ekonomika ČT24.

Identifikovat lze následující individuální nedostatky stínové mluvčí a chyby systému:

*Povinnou benzinovou kaucí složil a jen hrstka distributorů pohonných hmot, co bude dále?*

DT do systému zřetelně nadiktovala výraz *složila*, který byl chybně rozpoznán s oddělenou osobní koncovkou.

*Strašidelné kostýmy startují halou v římské večírky, na jeden zamíříme živě.*

Rozpoznáno mělo být slovo *halloweenskéé*, které DT do systému zřetelně nadiktovala. Systém má toto slovo v rozpoznávacím slovníku. Pokud má slovo systém ve slovníku a přesto není rozpoznáno, je vhodné se problematickému výrazu vyhnout a opsat ho jinými slovy. Stejně je nutné postupovat i v případě, pokud výraz není ve slovníku systému a není ho možné z určitých důvodů přidat. Například jedná-li se o výraz ojedinělý, nebo stínový mluvčí nemá dostatek času slovo do slovníku přidávat.

*Dobrá večer, připravili jsme pořád pořád pro dnešní den, tady je.*

Rozpoznán měl být výraz *pořad*. Chyba vnesená do rozpoznávání stínovou mluvčí, která dvakrát po sobě chybně zopakovala stejné slovo. Chybný výraz i jeho druhý zápis by se v tomto případě odeslal do vysílání, protože chybné rozpoznání nebylo opraveno.

*Dlouhodobý problém vyžadovala radikální řešení.*

Chyba zapříčiněná systémem rozpoznávání řeči. DT vyslovila určitý slovesný tvar se zřetelnou koncovkou.

*Ještě v září bylo v registru zapsáno téměř 2000 distributorů, dodnes zaplatil o kauci jen 10 % z nich.*

DT nevyslovila určitý slovesný tvar zřetelně. Osobní koncovku -o vyslovila s pauzou.

*Mezi těmi, kdo kvůli nezaplacení kauci ale nejsou jen podvodnice.*

DT do systému nadiktovala namísto výrazu *podvodníci* slovo *podvodnice*. Rozpoznaná věta nedává smysl, sémantický obsah měl být lépe formulován.

*Martin Turek byl dodavatel OPEC je šest let, ale na kauci peníze nesehnal.*

Chyba způsobená systémem, DT nadiktovala správně zkratku LPG.

*Podrobněji Vít Jedlička analytik společnosti koloseum, vítám vás dobrý večer.*

Přístavek *analytik společnosti koloseum* musí být oddělen čárkami z obou stran. Čárkou měl být také oddělen pozdrav *dobrý večer*.

*Zdá se to jednoduché, zavedeme kauce a od té doby bude svět růžový, je to opravdu takto jednoduché, zafunguje to?*

Před spojku *a* měla být doplněna čárka.

*Byla to velká chyba, a ten způsob kterým se to zavedlo.*

Spojka *a* byla rozpoznána navíc z důvodu hlasitého nádechu stínové mluvčí. Před vztažným zájmenem *který* chybí čárka. Věta měla být lépe formulována.

*Právě cílem tohoto opatření mělo být zamezit těm podvodům?*

Chybná interpunkce na konci věty, na základě předchozího kontextu se jedná o větu oznamovací.

*Ta firma vznikne, nebude mít tu licenci a bude obchodovat přes bílého koně a dojde k těm daňovým únikům a ten problém bude stejný jako před.*

Rozpoznán měl být výraz *předtím*. DT ho vyslovila, ale kladla příliš velký důraz na předponu, která byla rozpoznána, zbytek slova nikoliv. Před spojku *a* měla být doplněna čárka. Souvětí je také příliš dlouhé, vhodné by bylo rozdělit ho do kratších celků.

*Nebyl by dalším možným krokem to, že by podnikatele ukázali daňové příznání za posledních pět let, tady platíme daně, tak proč platit kauci?*

Rozpoznáno mělo být slovo *nebylo*. Chyba vnesená stínovou mluvčí, která nadiktovala chybný slovesný tvar *nebyl*.

*Nemůže být celý začátek toho problému v tom, že máme extrémně mnoho čerpacích stanic v ČR, pravděpodobně nejvíce v celé Evropě.*

Chybná interpunkce na konci věty, dle původní nahrávky se jedná o otázku.

*Utáhli si opasky, přestali krizi, propouštět nemuseli.*

Chyba způsobená systémem. DT nadiktovala *přestáli*. Systém ale vyhodnotil jako statisticky pravděpodobnější výraz *přestali*.

*Staly se firmu úroků.*

Chyba způsobená systémem. DT zřetelně nadiktovala *se firmou roku*.

*Ukážeme za chvíle.*

Chyba způsobená stínovou mluvčí. DT nadiktovala do systému chybnou koncovku.

*Pro nové obchodní parky jen místo v menších městech.*

Místo slova *banky* byl do systému nadiktován výraz *parky*, jde o chybu způsobenou stínovou mluvčí. Při vyslovování určitého slovesného tvaru došlo k drobnému zaváhání, které mělo za následek rozpoznání hlásky *n*.

*Za známky negativní agentura oceňuje, že vláda provedla změny.*

Rozpoznané souvětí nedává smysl. První věta vylučuje význam věty druhé. Místo výrazu *negativní* mělo být do systému nadiktováno slovo *pozitivní*.

*Další příjem záležití podle odborníku na politické situaci.*

Chyba systému, DT nadiktovala zřetelně výraz *odborníků*.

*Tam je velká otázka jak to bude s tím provizoriem.*

Před výraz *jako* měla být doplněna čárka.

*Preferuji budu z kohoutku.*

Chyba systému, DT nadiktovala do systému výraz *vodu*.

*V nákupním košíku stále častěji končí sirupy a ochuzené šťávy.*

Chyba způsobená stínovou mluvčí. Do systému nadiktovala namísto slova *ochucené* výraz *ochuzené*.

*Dáváme jí přednost protože je zdarma.*

Chybí čárka před výrazem *protože*.

*Největší americký telekomunikace tak i operátorů uvažuje o koupi britského Vodafone.*

Chyba v rozpoznání způsobená nadiktováním pro systém neznámého výrazu. Namísto výrazu *telekomunikační* byl nadiktován výraz [telekomunický]. Rozpoznaný výraz *operátorů* je chybou systému, DT nadiktovala zřetelně slovo *operátor*.

*Případným sloučením by vznikl co do tržeb největší telekomunikační operátor na světě.*

Rozpoznaný řetězec slov *co do tržeb* je přívlastek volný, který se odděluje z obou stran čárkami.

Přijdou vám nesmyslné a zbytečné.

Chybná interpunkce na konci věty. Z kontextu komunikátu lze vyčíst, že se jednalo o otázku.

*A to na stránce bankovní poplatky. Kom.*

Chyba způsobená systémem rozpoznávání řeči a zároveň stínovou mluvčí. DT nadiktovala do systému webovou adresu *bankovnipoplatky.com*, kterou ale systém rozpoznal nesmyslným způsobem. Nadiktované webové adresy systém s velkou pravděpodobností nerozpozná, je vhodné se jejich diktování do systému vyhnout.

*Pěkný večer, na tuto párty v jednom z největších pražských barů se otevřelo před 15 minutami.*

Tato věta byla nadiktována do systému stínovou mluvčí stejným způsobem, jakým byla následně rozpoznána. Tato věta, pokud by byla odeslána bez opravy do vysílání, by zcela jistě nebyla uživatelem skrytých titulků pochopena.

*Box vyžaduje klidnou mysl a pomalý dech, kontrolu nad vlastním vědomí, pomáhá to udržet kondici i soustředění boxerů.*

Chyba způsobená stínovou mluvčí. Do systému nenadiktovala koncovku -m.

*Většinu výrobků jde do automobilek.*

Chyba způsobená stínovou mluvčí. Namísto slova *většina* nadiktovala do systému výraz *většinu*.

*Tak se vítejte, přeji vám hezký večer a po víkendu opět na viděnou.*

Úvodní fráze nedává smysl. Její chybné rozpoznání bylo způsobeno stínovou mluvčí, která do systému nadiktovala navíc výraz *se*.

Z výše uvedené analýzy vyplývá, že stínová mluvčí DT dobře přizpůsobila svůj projev požadavkům systému rozpoznávání řeči, protože se zde téměř neobjevily chyby v rozpoznávání. Její projev byl z hlediska artikulace kvalitnější než u nahrávek z druhé fáze tréninku. Přesto se zde i nadále objevovaly chyby způsobené diktováním nesprávných slov, pauzy v artikulaci slov a nedostatky způsobené nedokončenými slovy. Chyby při doplňování interpunkce do rozpoznávaného textu se objevovaly v případě spojky *a*, oddělování přístavku a přívlastku volného a výrazu *jako*.

### 7.4.2 Komunikát vytvořený stínovou mluvčí DZ

Analyzován byl komunikát vytvořený na základě nahrávky pořadu Interview Daniely Drtinové.

Identifikovány byly následující individuální nedostatky stínové mluvčí a chyby systému:

*Jak byste řešil situaci, že 1000 lidí přijdou o práci?*

Chyba způsobená systémem. DZ nadiktovala tisíce, systém rozpoznal *1000*. Toto výrazně zkreslí pochopení uživatele skrytých titulků o celkové částce lidí, kteří mají přijít o práci.

*Z tohoto může vzniknout velký sociální problém a stát je tady od toho, abych pomohl lidem, kteří se ne vlastní vinou dostanou do takové situace.*

Chyba způsobená systémem. DZ zřetelně vyslovila výraz *aby*.

*Určitě by museli mít nějakou rezervu a museli by čítat s tím, že dojde k situaci, která se bude vyhrcovat.*

Chyba způsobená systémem, mělo být rozpoznáno slovo *počítat*. DZ zřetelně vyslovila celý výraz včetně předpony.

*Ve vašem volebním programu slibujete, že dáte zeminy do pořádku.*

Chyba způsobená systémem. DZ zřetelně nadiktovala výraz *zemi*, který měl být rozpoznán.

*Není to jen věta na papíře, ale přinese to efekt?*

Jedná se o dvě dohromady spojené otázky. Ve formě jedné otázky nedává souvětí příliš velký smysl. Vhodnější by byly dvě samostatné otázky.

*Já se domnívám, když si tyto výjimky mohli vyjednat Velká Británie nebo Irsko, tak je to jen otázka toho, přesvědčit ostatní partnery v rámci EU.*

Čárka za výrazem *toho* byla vložena navíc.

*Nikdy v programu jsem nenašla, jak byste měnili penzijní reformu, nebo zda byste jí nechali tak, jak je?*

V tomto rozpoznaném textu se vyskytuje kombinace chyb zapříčiněných systémem a zároveň stínovou mluvčí. Systém chybně rozpoznal nadiktovaný výraz *nikde*, který DZ zřetelně vyslovila jako *nikde*. Naopak stínová mluvčí do souvětí vnesla chybu špatným interpunkčním znaménkem na konci. Nejednalo se o otázku, ale o větu oznamovací.

*Je to skutečný termín kteřý hovoří o tom, jak lidé přispívají do systému.*

Chybí čárka před vztažným zájmenem *kteřý*.

*Souhlasíte s názorem Jiřího Čunka, který přišel s názorem, že by vystihoval nepřizpůsobivá tele obyvatele Varnsdorfu za město?*

Chyba způsobena stínovou mluvčí. Do systému byl nadiktován řetězec slov *vystihoval [nepřizpůsobivatele]*. Výraz *[nepřizpůsobivatele]* v českém jazyce neexistuje, systém ho nemohl rozpoznat, proto rozpoznal statisticky nejbližší slova *nepřizpůsobivá* a *tele*.

*Proto navrhuje, aby byla zavedena povinná předškolní výuka, buď ve školkách nebo doma.*

Chybí čárka před výrazem *nebo*. Jedná se o poměr vylučovací, proto zde čárka musí být uvedena.

Z uvedené analýzy výsledného komunikátu vyplývá, že stínová mluvčí DZ velmi dobře přizpůsobila svůj projev systému rozpoznávání řeči. Z její



strany není do procesu rozpoznávání vnášeno takové množství chyb, jako tomu bylo u stínové mluvčí DT. DZ chybovala v nesprávné interpunkci na konci vět a před výrazy *nebo* a *který*. Na rozdíl od stínové mluvčí DT nechybovala v interpunkci před spojkou *a*. Přesto se i u DZ objevily chyby v rozpoznání z důvodu diktování nesprávných slov.

### 7.4.3 Komunikát vytvořený stínovým mluvčím PP

Analyzován byl komunikát vytvořený na základě nahrávky pořadu *Ekonomika ČT24*.

Identifikovány byly následující individuální chyby stínového mluvčího a chyby systému:

*Hezký večer u Ekonomika ČT 24.*

Chyba způsobená stínovým mluvčím. Namísto tvaru druhého pádu nadiktoval do systému tvar prvního pádu.

*O kolik a kde se cena zastaví?*

Dvě otázky byly spojeny do jednoho celku. Vhodnější by bylo celek rozdělit na otázky dvě.

*Točené chutná lépe než zdlouhavé a jeho prodej živí mnohem více lidí.*

Chyba stínového mluvčího, který nejasně vyslovil řetězec *z lahve*, což vedlo k rozpoznání zcela jiného výrazu.

*Nejvíce se pivo bylo v restauracích, ale dnes prodejem vévodí hypermarkety.*

Chyba způsobená systémem rozpoznávání řeči. Stínový mluvčí PP zřetelně vyslovil výraz *pilo*.

*Je to trend, který je po celém světě a těžko to zvrátíme tím, že zdražíme lahvové pivo.*

Před spojkou *a* měla být doplněna čárka.

*Jak jsou zákazníci u Tyla u piva citliví na cenu?*

Chyba způsobená stínovým mluvčím. Do systému nadiktoval řetězec *u Tyla*, který měl po rozpoznání smazat, což neudělal. Následně se opravil a nadiktoval správný řetězec *u piva*.

*Pane pivovarů to bude.*

Chyba způsobená stínovým mluvčím. Do systému nadiktoval namísto výrazu *několik* nespisovný výraz *pár*, který systém nerozpoznal a nahradil ho statisticky nejbližším slovem *pane*.

*Nebude to ale takové, jako v minulých letech, že se vždy čekalo na lídra trhu který zdražil a vzápětí na to zdražily všechny ostatní.*

V tomto souvětí lze identifikovat dvě chyby v interpunkci. Před vztažným zájmenem *který* a před spojkou *a* chybí čárky.

*Ted' si to myslím nedovolí.*

Výraz *myslím* měl být z obou stran oddělen čárkami. Jedná se o vsuvku.

*Co zákazníci neopouštějí a co jistota to vašem oboru?*

Chyba způsobená systémem. PP zřetelně vyslovil určitý slovesný tvar *je*, který systém ale nerozpoznal. Navíc byl rozpoznán výraz *to*.

*Budou přicházet na to společné řešení, které jim ušetří peníze, ale také umožní efektivně využít to předělané pásmo a tím pádem by zákazníci potom mohli mít v některých oblastech třeba lepší signál.*

Jedná se o příliš dlouhé souvětí. Vhodné by bylo rozdělit ho do kratších celků z toho důvodu, že prostor pro zobrazení titulku na obrazovce disponuje pouze omezeným počtem znaků. Uživatel skrytých titulků by mohl mít problém s pochopením výsledného titulku, protože takto dlouhé souvětí by bylo rozděleno do několika řádek titulků.

*Vodafone, tedy třetí operátor na tu dohodu reagoval tím, že by se kvůli tomu měla odložit aukce kmitočtů pro rychlé mobilní síť.*

V tomto rozpoznaném textu se objevují chyby systému a zároveň stínového mluvčího. Řetězec *tedy třetí operátor* musí být z obou stran oddělen čárkami, jedná se o přístavek. Chybou systému je nesprávné rozpoznání výrazu *rychlou*, kterou PP do systému zřetelně nadiktoval.

*Z hlediska Vodafone ujeto určitě velice nepříjemný krok, protože Vodafone je na našem trhu nejmenším hráče.*

Chyby způsobeny stínovým mluvčím. Chybně rozpoznáný řetězec *Vodafone ujeto* byl způsoben nadiktováním druhého pádu výrazu *Vodafone* namísto pádu prvního, který systém dokáže rozpoznat. Koncovka *-u* následně ovlivnila rozpoznání výrazů je to, které byly systémem vyhodnoceny ve spojení s koncovkou *-u* jako slovo *ujeto*.

*Sám by potřeboval silného partnera se kterým vystavěl síť.*

Chyby byly způsobeny stínovým mluvčím. Před výraz *se* nebyla doplněna čárka. Namísto výrazu *by stavěl* nadiktoval PP do systému slovo *vystavěl*.

*Na druhou stranu se nedomnívám, že to sdílení sítě jako takový by měl být důvodem pro pozastavení aukce, nebo její odložení.*

Chyby způsobeny stínovým mluvčím a zároveň systémem rozpoznávání řeči. Do systému bylo nadiktováno slovní spojení *jako takový* namísto

spisovného *jako takové*. Jedná se o přívlástek volný, musí být tedy odděleno z obou stran čárkami. Co se týče chybné koncovky u slovesného tvaru *měl*, která měla být rozpoznána s koncovkou -o, jedná se o chybu systému, protože PP zřetelně správnou koncovku nadiktoval.

*Pokud aukce skutečně bude a bude tady čtvrtý operátor, tak Vodafone by se ve výstavbě sítí, ale ať už těch původních, nebo těch sítí čtvrté generace, tak by se Vodafone mohl spojit s dalším operátorem.*

Před spojku *a* nebyla doplněna čárka. Vhodnější by bylo také souvětí rozdělit do kratších celků.

*Vypadá to z průzkumu agentury Reuters.*

Chyba způsobena stínovým mluvčím. Do systému namísto výrazu *vyplývá* nadiktoval slovo *vypadá*.

*Proč my podnikatelé, kteří jsme nic neudělali tak máme být sledováni?*

Věta uvozená vztažným zájmenem *kteří* je větou vloženou, musí být tedy oddělena z obou stran čárkami.

*Tam, kde se měli tlačit zástupy zákazníků rostou stromky.*

Chybí čárka před určitým slovesným výrazem *rostou*.

*Velká voda, která se v červnu prohnala ČR znovu strašíme jetele zatopený dům.*

Věta vložená uvozená vztažným zájmenem  *který* není oddělena z obou stran čárkami. Chyba v rozpoznání byla způsobena stínovým mluvčím. Do systému nadiktoval řetězec *straší [mejitele]*, což vedlo k rozpoznání nesmyslného řetězce slov. Namísto spojení *zatopených domů* bylo do systému nadiktováno *zatopený dům*.

*1. 3 roky jsme doopravdy pouze vysoušely.*

Pokud se rozpoznají čísla tímto způsobem, je nutné rozpoznání text smazat a užít slovního příkazu „Číslo slovy“.

*Porod čas na zateplení, které je dobré udělat před tou zimou, pokud to zdivo není dostatečně vyschlé, tak není zase tolik nákladné.*

Chyba v rozpoznání způsobena nezřetelnou výslovností stínového mluvčího. Namísto zřetelného nadiktování výrazů *to dočasné zateplení* do systému nadiktoval shluk hlásek [*todočas*].

*Polystyrén, a minerální vata.*

Čárka před spojkou *a* byla doplněna chybně.

*Ten problém je ten, který byl zmíněn, tedy problém statický který může nastat.*

Před vztažné zájmeno *který* nebyla doplněna čárka.

*Ty náklady nejsou tak vysoké a pro ten barák je to velice důležité.*

Před spojkou *a* nebyla doplněna čárka. Výraz *barák* je nespisovný, PP ho měl nahradit výrazem *dům* i v tom případě, že slovo *barák* použil původní řečník ve vysílání.

*Skutečně k tomu, že ten dům bude ohrožen?*

Věta nedává smysl. PP vynechal slovesný tvar v první části věty. Věta měla pravděpodobně znít takto: *Skutečně k tomu může dojít, že ten dům bude ohrožen?*

*Zdeněk příhoda začínal jako většina jeho kolegů s málem.*

Chyba způsobená systémem. Jazykový model nedokázal správně vyhodnotit, že se jedná o vlastní jméno, které musí být rozpoznáno s velkým počátečním písmenem.

Z uvedené analýzy výsledného komunikátu vyplývá, že stínový mluvčí PP dokázal během tréninku přizpůsobit svůj projev potřebám systému rozpoznávání řeči. Na rozdíl od analyzovaných nahrávek z druhé fáze tréninku diktoval do systému zřetelněji a nedělal pauzy uprostřed artikulace slov. Do procesu rozpoznávání nevnášel ani příliš chyb z pohledu interpunkce. Nedostatky byly zjištěny zejména v doplňování čárek před spojkou *a* a vztažným zájmenem *který*. Přesto se i u PP objevily chyby v rozpoznání z důvodu diktování nesprávných slov.

### **7.5 Shrnutí nadindividuálních chyb zjištěných v komunikátech vyhotovených ve čtvrté fázi tréninku**

Došlo k potvrzení předpokladu, že se sníží počet chyb v rozpoznávání způsobených nesprávnou artikulací a pauzami během výslovnosti výrazů. Potvrzena byla také hypotéza, že většina chyb, které se objevují v komunikátech vytvořených ve čtvrté fázi tréninku, byla způsobena nesprávně vloženou interpunkcí.

U stínových mluvčích DT a PP byly identifikovány obtíže při vkládání čárky před spojku *a*. Ve většině případů, pokud měla být čárka před spojkou vložena, interpunkční znaménko přidáno do textu nebylo.

Stínové mluvčí DT a DZ vkládaly také chybnou interpunkci na konec věty.

Stínoví mluvčí DZ a PP nevkládali v mnoha případech čárku před vztažné zájmeno *který*.

Všichni tři stínový mluvčí způsobovali systému obtíže s rozpoznáváním z důvodu diktování chybných nebo nespisovných slov, které následně systém rozpoznal jako výrazy, které neodpovídaly kontextu vět. Vyskytly se také nedostatky v oddělování přístavku a přívlastku volného čárkami z obou stran.

## **7.6 Komunikáty vytvořené v živém vysílání**

U každého ze tří stínových mluvčích byla také analyzován komunikát, který byl vytvořen během stínového přemlouvání živého vysílání pořadu Hyde Park. Během analýzy bylo zejména sledováno vkládání interpunkce a také smysluplnost rozpoznávaného textu. Tato podkapitola byla do diplomové práce zařazena z toho důvodu, aby byl jasně doložen vývoj schopností jednotlivých stínových mluvčích od druhé fáze tréninku až po přemlouvání živého vysílání.

Předpokladem je, že výsledný komunikát bude bez chyb v rozpoznávání způsobených nesprávným stylem diktování a bude obsahovat pouze minimální nedostatky v zadávání interpunkce.

### **7.6.1 Komunikát vytvořený stínovou mluvčí DT**

Identifikovány byly následující individuální chyby stínové mluvčí a chyby systému:

*Dobrý večer, je to první povolební Hyde park a to je prostor pro vás a pro vaše otázky.*

Chyby způsobené systémem a zároveň stínovou mluvčí. DT do systému nadiktovala zřetelně výraz *tu*, rozpoznáno bylo slovo *to*. Před spojkou *a* nebyla vložena čárka.

*Je pravda, že proti sobě stála Sparta a Slavie a teď se konečně začne hrát národní tým jak říká Miloš Zeman.*

Do rozpoznávaného souvětí nebyly na odpovídající místa vloženy čárky. Čárky měly být vloženy před spojku *a* a také před výraz *jako*. Zároveň z předchozího kontextu vyplývá, že se jednalo o otázku, proto měl být na konec věty vložena otazník.

*Výsledky známé a vítěz je vždy jen jeden.*

Chyba způsobena systémem. DT zřetelně nadiktovala výraz *známe*. Před spojkou *a* měla být vložena čárka.

*Pod ním se rychle přidává hlavní aktér volebního víkendu, na Miloše zemana si jeho příznivci museli chvíli počkat.*

Chyby způsobeny systémem. DT nadiktovala zřetelně řetězec *k nim*, který byl rozpoznán jako *pod ním*. Vlastní jméno *Zemana* mělo být rozpoznáno s velkým písmenem.

*Potom nám velmi dobře vycházeli také může 45 až 55 let.*

Chyba rozpoznání způsobena systémem. DT nadiktovala zřetelně výraz *muži*.

*Jestli ty skupiny, které se hlásili k voličům před tím nakonec zvolili Karla Schwarzenberga?*

Výraz *před tím* měl být rozpoznán dohromady. Jedná se o již zmiňované obtíže systému rozpoznat správně nepravá homonyma. Před výraz *nakonec* měla být vložena čárka.

*Když říkáte, že v prvním kole jste cíle na určitou skupinu a pak se to neukázalo.*



Jedná se o chybu způsobenou systémem. DT zřetelně vyslovila výraz cílili. Před spojkou a měla být vložena čárka. Věta měla být také lépe formulovaná, takto rozpoznaná nedává příliš smysl.

*My jsme si dělali nějaké vlastní odhady, náš vlastní interní výzkum, ze kterého vycházela určitá data, která přibližně na konci prosince, kdy jsme uvažovali o tom druhém kole a dělali jsme si scénáře, jak na Miloše Zemana, tak na pana Fišera, tak na pana Dienstbiera, v případě Miloše Zemana to vycházelo tak, že ten odliv od různých kandidátů, že půjdou ti voliči směrem k němu, což se potvrdilo druhý týden v lednu.*

Jedná se o příliš dlouhé souvětí, které mělo být rozděleno do kratších celků.

*Ukázalo se, že tam půlmilionový rozdíl a bylo jasné, že klíčový budou voliči ze starších segmentů.*

DT do systému nenadiktovala určitý slovesný tvar *je*.

*To je zajímavý názor, ten pán by jistě souhlasil, že by se těžce měřila účinnost kampaně, kdy ta kampaň má nějaký efekt nebo nemá.*

Před výraz *nebo* měla být vložena čárka, protože se jedná o poměr vylučovací.

*Vy jste si v prosinci udělali testovací modely, kdyby Schwarzenberg postoupil do druhého kola s určitým kandidátem anatom jste to stavěli.*

Chyby způsobeny systémem a zároveň stínovou mluvčí. Rozpoznaný výraz [*anatom*] měl jazykový model vyhodnotit jako tři oddělená slova, protože DT *je* nadiktovala zřetelně s mezislovními pauzami. Z předchozího kontextu je zřejmé, že se jednalo o otázku, na konec věty měl být vložen otazník.

*Pak následovala ta programová část, co bude jako prezident těla a ta finální část měla být o tom, jaký bude prezident, jaké bude zastávat postoj, co bude dělat, to už bylo myšleno na to druhé kolo.*

Chyba způsobená systémem rozpoznávání. DT zřetelně nadiktovala výraz *dělat*. Před spojkou a měla být vložena čárka.

*Volební manažér, jeho zástupce, tak tam lidé dohlíželi na všechny konce termínu, zástupce pro sociální sítě, někdo dohlížel na rozpočet, člověk na plánování rozpočtu Karla Schwarzenberga, tak kampaň opravdu nebyla cílená pouze na Prahu, další lidé pracovali na tisku, například pracoval na tom, aby návrh odešel na billboardy, já bych nerada na někoho zapomněl.*

Chyby v rozpoznání způsobeny systémem. DT zřetelně nadiktovala výrazy *pak* a *zapomněla*. Jedná se o příliš dlouhé souvětí, které mělo být rozděleno do kratších celků.

*Co na tom bylo úplně nejhezčí, všichni ví, že má ochranku, protože ministr zahraničí, ale z toho, co jsem potkala v životě hodně politiku, aby si pohlídal, jestli se najedla ochranka a také si vždy pohlídal, jestli se najedli všichni lidé týmu, jak se mají, jestli si stihli odpočinout, tak se o nás staral.*

Chyba způsobena systémem. DT zřetelně nadiktovala určitý slovesný tvar *je*, který systém nerozpoznal. Souvětí mělo být také rozděleno do kratších celků, v této podobě mohlo uživatelům skrytých titulků činit potíže rozpoznatý text pochopit.

*Bylo to problém zejména v první části kampaně, bylo to v říjnu a v listopadu a tam měl ještě velký počet zahraničních cest a nejen že ministr zahraničí, tak chodil do PS, na vládu, takže se s tím časem muselo nakládat velice úsporně, často ta jednání začínala v šest nebo sedm večer a byly do dvou hodin do noci.*

Chyba způsobena systémem. DT zřetelně nadiktovala určitý slovesný tvar *je*, který systém nerozpoznal. Druhá z chyb byla způsobena stínovou mluvčí, která namísto výrazu *byla* nadiktovala do systému slovo *byly*.

*Kdyby toto stanovisko neměl na začátku, doporučovala byste mu aby skončil?*

Před výraz *aby* měla být vložena čárka.

*Máte měsíční taktiky, které měníte a v závěru kampaně se některé věci museli dělat ze dne na den.*

Před spojkou *a* měla být vložena čárka.

*Takže si vás vyslechlo a pak se rozhodl sám?*

Chyba způsobena systémem. DT zřetelně nadiktovala výraz *vyslechl*.

*Průzkumy těch preferencí vycházely pravidelně od června a července, my jsme si nechali udělat výzkum, který spíše testoval, s čím si lidé Karla Schwarzenberga testují, kde cítí jeho silné stránky, kde cítí jeho slabiny, jak ho vnímají vůči jeho konkurenci, to nám pomohlo v té strategii.*

Souvětí příliš nedává smysl. Spíše než výraz *testují* měl být do systému nadiktován výraz *spojují*.

*Q mnoho let podporuje různé charitativní organizace, když jsme mu říkali, že musí zdůraznit, co za ním je, tak on velmi striktně řekl, že se nebude ničím chlubit.*

Chyba způsobena systémem. DT zřetelně nadiktovala osobní zájmeno *on*.

*Byl tam problém s tím, že není Čech a že na to lidé reagují špatně?*

Před spojku *a* měla být vložena čárka.

*Před druhým kolem jsme si zadali průzkum těsně před první volbou, abychom věděli, kdy si stojíme.*

Chyba způsobena nezřetelnou výslovností stínové mluvčí. Její artikulace výrazu *kde* se podobala spíše rozpoznávanému výrazu *kdy*.

*Velká část té kampaně, která logicky není vidět byla kontaktní, měli jsme velký počet mladých lidí, jak dobrovolníků, tak podporovatelů, kteří obcházeli různá města.*

Před určitý slovesný tvar *byla* měla být vložena čárka.

*Celá ta situace začala pod účelu ČT, kdy když jsme vyhodnocovali tu debatu, tak jsme si říkali, ta situace není ideální, příště musí okamžitě zazní, že se nesmí nikomu nic dělat s majetkem, to jsme mu řekli.*

Chyba způsobena systémem. DT zřetelně nadiktovala řetězec *po duelu*.

*Mohl mít a vše ten potenciál, aby přiměl voliče, aby šli volit, ale neměl určitě ten rozhodující potenciál.*

Chyba způsobena systémem. DT nadiktovala zřetelně výraz *avšak*.

*Ano v Británii je to také jasné, nicméně u nás to nikdy zvykem nebylo.*

Za výraz *ano* měla být vložena čárka.

*Například i ta mobilizace lidí, že se vizuálně označoval, komu dávají přednost, stejně tak i ty deníky.*

Chyba systému. DT zřetelně nadiktovala *označovali*. Souvětí mělo být lépe formulováno, nedává příliš smysl.

*Vy jste z Fakulty sociálních věd, to znamená že média řešíte.*

Před výraz *že* měla být vložena čárka.

*Já na to nemám data, říkalo se, jestli jsme to na sociálních sítích tzv. nepřipálily.*

Chyba způsobena systémem. DT zřetelně nadiktovala slovo *nepřepálili*.

V komunikátu vytvořeném při přemlouvání živého vysílání se vyskytovaly chyby v rozpoznávání způsobené systémem automatického rozpoznávání řeči a zejména v chybném vkládání interpunkce. Především se jednalo o vkládání interpunkce před spojky *a*, *aby*, *že* a *jak*. DT by avšak měla více kontrolovat průběžně rozpoznávaný text. Přestože do systému diktovala slova jasně a zřetelně, v některých případech systém rozpoznal chybné koncovky nebo nesmyslné výrazy, které by měly být smazány a přemluveny znovu, což DT v mnoha případech neudělala a nechala v rozpoznávaném textu chyby, které se zobrazily uživateli skrytých titulků, a mohly vést k obtížnému porozumění výsledného titulku. Vhodné by také bylo nevytvářet příliš dlouhá souvětí, nýbrž kratší věty, které usnadňují uživateli skrytých titulků porozumění.

Komunikáty stínové mluvčí DT ze čtvrté fáze tréninku a vytvořené při přemlouvání živého vysílání pořadu Hyde Park vykazují zlepšení co se týče míry chybovosti rozpoznávání a počtu chyb, které do procesu rozpoznávání vnášela stínová mluvčí, oproti nahrávkám, které byly vyhotoveny v druhé fázi tréninku. DT odstranila ze svého projevu hezitační zvuky, pauzy během artikulace slov a diktování chybných výrazů, které zapříčiňovaly chybná rozpoznání ve výsledném komunikátu.

### **7.6.2 Komunikát vytvořený stínovou mluvčí DZ**

Identifikovány byly následující individuální chyby stínové mluvčí a chyby systému:

Správy v osm.

Jedná se o chybu systému, kterému činí obtíže rozpoznávat homonyma.

*Návrh byl pozměněn do konečné podoby na počítače Zdeňka Koudelky, bývalého poradce ČSSD a také prezidenta Klause.*

Chyba způsobená stínovou mluvčí, která do systému nadiktovala výraz počítače namísto počítači.

*Býval ministrem práce a sociálních věcí, míval premiérem a eurokomisařem.*

Chyba způsobena systémem. DZ nadiktovala zřetelně výraz býval.

*To už se blížila první přímá volba prezidenta a opět rostla moc může, se kterým si Vladimír Špidla nikdy moc nesesl, Miloše Zemana.*

Chyba způsobena systémem. DZ zřetelně nadiktovala výraz muže.

*Bavil jste se, snaha jste se během toho příspěvku, co vás tak pobavilo?*

Chyba způsobena systémem. DZ nadiktovala zřetelně slovo smál.

*Vždy je zajímavé podívat se na nějaký průřez života, který je udělán nikým, kdo má svoji vlastní představu.*

Chyba způsobena systémem. DZ nadiktovala zřetelně výraz někým.

*Nebo vládu přímo viníte zad, že jsme do recese spadly?*

Chyba způsobena systémem. DZ nadiktovala zřetelně výrazy za to.

*Rakousko není větší než jsme my.*

Před předložku než musí být v tomto případě vložena čárka.

*První věc je připravit rozpočet na příští období, protože pravděpodobně volby budou na konci května, bude na vládě, aby zásadně revidovala rozpočet na příští období, aby přeorganizovala priority, aby po mém soudu je nasměrovala na vzdělání a inovace, to je podstatné.*

Příliš dlouhé souvětí, které by bylo vhodné rozdělit do kratších celků.

*A bylo by trapné, kdybychom postupovali stejně, jako v tomto období.*

Čárka před výraz *jako* byla vložena chybně.

*Nicméně říká se, že za ten největší problém mohou dva aspekt. Ty jedna, poslankyně Šedivá, a pak se říká druhá věc, a to byly příliš velkorysé propočty, které byly v tom zákoně, které vyšly z dílny pana Ambrozka.*

Chyba způsobena stínovou mluvčí, která nevložíla interpunkci v mezislovní pauze, což způsobilo rozdělení rozpoznávaného slova na dvě části.

*A tady se musíme zeptat, proč byla základní sazba dána na 12 50 a ne na 8 40.*

Rozpoznávaná čísla nedávají příliš smysl. Mezi 12 a 50 a 8 a 40 měly být vloženy čárky, tečky popřípadě slovním povelem dvojtečka.

*A ten systém vyžaduje, abychom jej revidoval.*

Chyba způsobena systémem. DZ zřetelně nadiktovala výraz *revidovali*.

*Já jsem definoval problém energetické chudobě.*

Chyba způsobena systémem. DZ zřetelně nadiktovala výraz *chudoby*.

*Pro mě je nepřijatelná energie, která není rozumět dostupná, protože je příliš drahá.*

Chyba způsobena systémem. DZ zřetelně nadiktovala výraz *rozumě*.

*Já si myslím, že využívat řepku olejnou jako příměs do paliva, je správné, ale závisí na míře.*

Čárka před určitým slovesným tvarem *je* byla vložena chybně.

*EU chce, aby sedávala jako příměs.*

Jedná se o chybu systému. Jazykový model měl identifikovat, že se jedná o dvě slova.

*A dokonce se mluví o tom, že ta pětiprocentní hranice je příliš, co se přidává do paliva.*

Rozpoznané souvětí nedává příliš smysl. Vhodné by bylo jej přeformulovat, což je jeden z úkolů, které musí stínový mluvčí během přemlouvání vykonávat.

*A využívají se nejrůznějšími způsoby řadě technologií, od chemických, biologických až k energetice.*

Chyba způsobena systémem, který nerozpoznal jasně vyslovenou předložku *v*.

*Pro mě ta věc uzavřená.*

Chyba způsobena systémem. DZ nadiktovala určitý slovesný tvar *je* zřetelným způsobem.

*Já vím, že to je asi 40 jich šest desetin %, 4,6 %.*

Chyba v rozpoznání způsobena opravou provedenou stínovou mluvčí. Do vysílání se odeslal chybně rozpoznaný řetězec, DZ ho následně opravila. Pokud k tomuto případu dojde, je vhodné za chybný výraz nadiktovat



slovo *oprava*, které bude rozpoznáno, a následně zopakovat výraz znovu správně. Uživatel skrytých titulků jasně vidí, že předchozí chyba byla opravena.

Komunikáty stínové mluví DZ ze čtvrté fáze tréninku a vytvořené při přemlouvání živého vysílání pořadu Hyde Park jsou na stejně kvalitní úrovni. DZ diktuje výrazy do systému zřetelně a to včetně koncovek. Neobjevují se zde nesmyslně rozpoznané řetězce slov způsobené chybnou artikulací nebo diktováním jiných popřípadě nespisovných slov. V ojedinělém případě se objevila chyba způsobená chybně načasovaným vložením interpunkce během artikulace slova, což vedlo k rozdělení výrazu na dvě části. Některá souvětí by bylo vhodné rozdělit do kratších celků.

### 7.6.3 Komunikát vytvořený stínovým mluvěním PP

Identifikovány byly následující individuální chyby stínového mluvěního a chyby systému:

*Šestý den od začátku povodní voda opadá a sčítají se škody hosty speciálního Hyde parku budou hydrolog Jakub Langhammer starosta Všechovic Miloš Čapek a generální ředitel pražského Dopravního podniku Milan Křístek.*

Jména jednotlivých hostů měla být oddělena čárkami.

Velké míře by odklizení škod mělo začít v neděli.

Chyba způsobena systémem. PP zřetelně vyslovil spojení ve *velké*.

Očekáváme pokles hladiny na osm m, kdyby to mělo výrazně změnit život ve městě a to, že by mohl být opět zprovozněny oba silniční mosty přes

*řeku Labe a také by mohlo být potom zprovozněna vstupní hala do hlavního nádraží ČD.*

Souvětí je příliš dlouhé, mělo by být rozděleno do kratších celků. Chyby v rozpoznání osobních koncovek u sloves byly způsobeny nezřetelnou artikulací stínového mluvčího.

*V metropoli dále zůstávají uzavřen čtyři desítky silnic.*

Chyba v rozpoznání způsobena nezřetelnou artikulací osobní koncovky.

*Probíhalo vše tak, jak mělo.*

Z kontextu je jasně patrné, že se jednalo o otázku. Stínový mluvčí zadal na konec věty chybné interpunkční znaménko.

*Místa nebyl proud tak silný, že se silnice nedala přejít.*

PP vyslovil určitý slovesný tvar správně, systém ale rozpoznal antonymum, které PP neopravil, což vedlo k jeho odeslání do vysílání.

*V pondělí začali menší toky kulminovat, nikoliv však berunka, Sázava nebo Vltava.*

Chyba způsobena systémem. PP zřetelně vyslovil výraz *Berounka*.

*Zcela zaplavený těch Prahy se stal jeden ze symbolů letošních povodní.*

Chyba způsobena systémem. PP vyslovil zřetelně slovo *jih*.

*Majetek se v podstatě tuto chvíli neřeší.*

Chyba způsobena stínovým mluvčím. PP do systému předložku v nenadiktoval.

*Čekám na to, kdy to skončí abychom mohli pracovat.*

Před výraz *abychom* měla být vložena čárka.

*Je zapotřebí být připravený na eventualitu, že situace, kterou zažíváme není stále u konce.*

Před výraz *není* měla být vložena čárka.

*Pane docente, vy tomu asi jako hydrolog budete mít co nejvíce říci.*

Chyba způsobena systémem. PP předložku *k* nadiktoval zřetelně.

*To, že se atmosféra v posledních letech chová nestandardně a to, že s velice mocnými prostředky a modely je velice těžké předpovědět stav atmosféry, takto je zřejmé.*

Jedná se o nedostatek jazykového modelu. Rozpoznána měla být dvě slova.

*Jsme to simulovali matematickými modely po povodních v roce 2002, tak to bylo zhruba o hodinu a čtvrt.*

Chyba způsobena stínovým mluvčím, který osobní zájmeno *my* nadiktoval do systému nezřetelně.

*To se může zdát, jako málo, ale v případě rychlého nástupu povodně, tak každá minuta může být drahá pro evakuaci obyvatel, nebo přípravu záchranných opatření.*

Čárka před výraz *jako* byla doplněna navíc.

*Pane starosto, vy žijete na břeh řeky poměrně dlouho.*

Chyba způsobena systémem. PP zřetelně vyslovil výraz *břehu*.

*To je otázka daleké minulosti, kdy územní plány byly pochopitelně páhány v takovém rytmu, že zastavit co nejvíce a stavět všude možně.*

Chyba způsobena systémem. PP vyslovil zřetelně infinitivní tvar *zastavět*.

*To dnes asi není téma, které bychom tedy vyřešili.*

Chyba způsobena systémem. PP vyslovil zřetelně výraz *tady*.

*Odtékat a těmi kanály do jímek, ze kterých se to potom čerpadly vyčerpá na povrch.*

Chyba způsobena systémem. PP vyslovil zřetelně spojení *odtéká to*.

*Vím, že se dávali pytle například na invalidovně, nebo v Křížkově.*

Čárka před výraz *nebo* byla vložena navíc.

*Kdybychom nedělali preventivní opatření, tak aby se vylila stejně, jako v roce 2002.*

Čárka před výraz *jako* byla vložena navíc.

*Podle povodňového plánu jsme začali ty etapy proti původní plnit a zabránili jsme tak vyplavení metra.*

Chyba v rozpoznání způsobena stínovým mluvčím, který do systému namísto výrazu *povodní* nadiktoval slovo *průvodní*.

*Chci říci, proč je tak důležité nenechat zaplavit m a udržet ho suché.*

Rozpoznáno mělo být slovo *metro*, který PP do systému nadiktoval nikoli zkratka pro metr.

*Dají se ty komory uzavřít tak, že by se zavřely autobusy a dostat by se nedostala voda?*

Chyba v rozpoznání způsobena systémem. PP zřetelně nadiktoval že by se zavřely *tubusy* a do stanic.

*Tyto prohlídky a testy probíhají v nočních hodinách a druhé ranní, do páté ranní, kdy metro je vypnuté a lidé se mohou fyzicky m pohybovat.*

Chyba způsobena systémem. Rozpoznáno mělo být slovo *metrem*, které systém vyhodnotil jako jednotku délky a rozpoznal zkratku *m*.

*Myslím si, že ta opatření která byla nastavena jsou skutečně v maximální možné míře užitečná pro prevenci.*

Věta uvozená zájmenem *kte*rá měla být oddělena z obou stran čárkami, jedná se o větu vloženou.

*Máte pocit, že někdo zaspal a mělo se upouštět dříve a více, proto aby se udělala větší rezerva?*

Čárka měla být vložena před výraz *aby* nikoliv před výraz *proto*. Chyba v rozpoznání způsobena systémem. PP vyslovil zřetelně infinitivní tvar *zastavět*.

*Po povodních máme v ČR máme mnoho hospodářům souboru.*

Do systému navíc nadiktovaný slovesný tvar *máme* měl být smazán, byl však odeslán do vysílání. Chyba v rozpoznání způsobena systémem. PP zřetelně nadiktoval spojení s *vodou*.

*Kdybychom ale pokaždé když mají být povodně snižovali stavy vody, tak to máme snížené pořád.*

Věta uvozená výrazem *když* je věta vložená, měla být oddělena z obou stran čárkami.

*Ta situace současnosti podle nich ale tak vážná, že nikdo z jeho kolegů ani on se k nám do Hyde parku nevydal.*

Chyba způsobena systémem. PP zřetelně nadiktoval určitý slovesný tvar *je*.

*V puse byla postavena protipovodňová zed'.*

Chyba v rozpoznání způsobena stínovým mluvčím. Namísto výrazu *v kuse* nadiktoval výraz *v puse*.

*Ona byla nějaká.*

Chyba způsobena systémem. PP nadiktoval zřetelně výraz *nijaká*.

*Těžko se tomu nějak vyjadřovat, ale když na stránkách Krajského úřadu najdu dvojstránkový list s krizovým štábem, všemi jmény a není tam ani jeden telefon, tak i my ten list k ničemu.*

Chyba způsobena systémem. PP nadiktoval zřetelně výrazy *tak je mi*.

*Znamená to že kraje v tomto nejsou moc funkční?*

Před výraz *že* měla být vložena čárka.

*Dokonce z poloviny máme umyto obec.*

Chyba způsobena systémem. PP nadiktoval zřetelně výraz *umytou*.

*Budu mluvit za Dopravní podnik za sebe, jako člena krizového štábu hlavního města Prahy.*

Chyba způsobena systémem. PP nadiktoval zřetelně spojku *a*.

*Z mého pohledu, když jsme koordinovali činnost Synot městskými částmi, například jsme pomáhali s evakuací a vypravili jsme 15 autobusů a dalších 10 jsme měli připravených záloze na evakuaci například Nemocnice na Františku apod.*

Chyba v rozpoznání způsobena stínovým mluvčím. Namísto výrazu *Synot* mělo být rozpoznáno spojení s *jednotlivými*, které ale PP do systému nadiktoval nezřetelně. Souvětí je také příliš dlouhé, bylo by vhodné ho rozdělit na kratší celky.

*Teď když nebudou spadnou srážky, tak i srážky se okamžitě přetvoří pórkový odtok.*

Pauza během artikulace slova *na půdu*, které bylo následně rozpoznáno jako *nebudou*.

*Nikdo nedovede říci, kam ta voda poteče a kolik bude, protože se s tím nepočítá.*

Chyba způsobena systémem. PP zřetelně nadiktoval výraz *jí*.

*Protože se počítá s tím, že se voda někam vylije, kam se být nemůže.*

Chyba v rozpoznání způsobena stínovým mluvčím, který udělal pauzu během artikulace výrazu *vylít* pauzu za předponou, což vedlo k rozpoznání zcela jiného slovesa.

*Odhadnout, kdy to vyklidíme je velice těžké.*

Před určitý slovesný tvar *je* měla být vložena čárka.

*Pokud by nepršelo a ta situace se dále nekomplikovala, za jak dlouho předpokládáte, že by se to dalo vyřešit.*

Z kontextu věty lze poznat, že se jednalo o otázku. Na konec bylo vloženo chybné interpunkční znaménko.

Teď to nebude silnice, tak nemáme šanci ani napojit ale triku.

Chyby v rozpoznání byly způsobeny stínovým mluvčím. Do systému byla nadiktována nespisovná slova [dokavad] a *elektriku*, které systém nedokáže rozpoznat.

*V kolika kubících to je a jak daleko jsme od toho?*

Před spojkou *a* měla být vložena čárka.

*Primární záměr metra bylo ochránit obyvatelstvo protichemickému, nebo atomovému nebezpečí.*

Chyba způsobena jazykovým modelem systému, který obtížně rozpoznává homonyma. Výraz měl být rozpoznán zvlášť.

*Kdyby přišla tedy něco takového, tak metro by svojí funkci splnil.*

Z kontextu této věty vyplývá, že se jedná o otázku. Na konec měl být vložen otazník.

*Proto jsem neodmítnou, ale řekl jsem, že to máme vyřízeno.*

Chyba způsobena systémem. PP zřetelně vyslovil výraz *neodmítnul*.

*Misi pomoc přeci musíme od někoho dovědět.*

Věta je nesmyslná.

*Přeji vám hezký večer pane starosto. Hezký večer pane hejtmane.*

Oslovení musí být odděleno čárkou.

Stínový mluvčí PP vykazuje při tvorbě komunikátu v živém vysílání oproti tréninku výrazné zlepšení. Přesto je většina chyb v rozpoznání způsobena jeho nezřetelnou artikulací a diktováním nesprávných slov, které systém nerozpozná. Také by měl s větší pečlivostí kontrolovat rozpoznávaný text. Z výše uvedené analýzy vyplývá, že ve skrytých titulcích se uživatel



zobrazila chybně rozpoznaná slova, v jednom případě dokonce slovo se zcela opačným významem. Velké nedostatky byly také zjištěny ve vkládání interpunkce, zejména před spojkou *a*, výrazy *abychom*, *nebo*, *jako*, *který* a *že*. V několika případech byla také zaznamenána chybně vložená interpunkce na konci věty. Komunikát vytvořený stínovým mluvčím PP v živém vysílání vykazoval největší míru chyb v rozpoznání a ve vkládání interpunkce.

### **7.7 Shrnutí nadindividuálních chyb zjištěných v komunikátech vyhotovených v živém vysílání**

Za vážnou chybu lze považovat v případě stínových mluvčích DT a PP nedůslednou kontrolu rozpoznávaného textu. V jejich komunikátech, zejména v těch, které byly vytvářeny v živém vysílání, byly identifikovány chyby v rozpoznání i interpunkci, které nebyly opraveny, přičemž mohly způsobit uživateli skrytých titulků problémy s jejich pochopením. Oba stínoví mluvčí také do systému diktují nespisovné výrazy, které systém nedokáže rozpoznat. Dále PP a DT nekládají správně interpunkci do rozpoznávaného textu. Jedná se zejména o neznalost pravidel vkládání interpunkce před spojkou *a*, výrazy *který*, *než*, *nebo*, *že* a *aby*.

U všech tří stínových mluvčích se také v několika případech objevila ve výsledném komunikátu příliš dlouhá souvětí, která by bylo vhodné rozdělit do kratších vět jednoduchých.

### **7.8 Chyby systému ovlivňující analyzované komunikáty**

V následující podkapitole budou shrnuty nedostatky systému, které snižují míru přesnosti rozpoznávání jednotlivých stínových mluvčích.

#### **1) zakončení slov**

Systému automatického rozpoznávání slov činí problémy rozpoznat správnou koncovku u výrazů, u kterých koncovce bezprostředně předchází následující hlásky nebo jsou na tyto hlásky zakončeny:

**-d** Příklad: podvod → rozpoznáno *podvodům*

**-j** Příklad: boj → rozpoznáno *boji* nebo *boje*

**-l** Příklad: přinesl → rozpoznáno *přineslo*

**-m** Příklad: problém → rozpoznáno *problémy*

**-t** Příklad: projekt → rozpoznáno *projekty*

**-z** Příklad: přichází → rozpoznáno *přicházím*

**-ch** Příklad: státních zastupitelstvích → rozpoznáno státních zastupitelství

## **2) rozpoznávání jednoslabičných slov**

Nedostatky v rozpoznávání vykazuje systém zejména u jednoslabičných slov. Výraz je buď rozpoznán s jinou hláskou, nebo není rozpoznán vůbec.

Příklady:

ono → ano; ano → ono

vy → by; ty → vy; by → ty

určitý slovesný tvar *je* → v mnoha případech není vůbec rozpoznán

## **3) neschopnost graficky odlišit nepravá homonyma**

Graficky odlišovat nepravá homonyma je úkolem jazykového modelu systému automatického rozpoznávání řeči. Chybně zapsaná homonyma

nemění výrazně smysl rozpoznávaného textu, jejich nesprávný zápis může ale působit rušivě.

Příklady:

že na → žena

po té → poté

my to → mi to

co by → coby

musely → museli

#### **4) rozpoznávání po sobě bezprostředně následujících slov, která začínají na stejnou hlásku**

V nahrávkách byly identifikovány hlásky způsobující chyby při rozpoznávání začínají-li na ně dvě po sobě bezprostředně následující slova.

**j** - u řetězců slov *je ještě a je jejich* (i v případě, že jsou slova naopak)

V případě této hlásky systém nerozpozná s nejvyšší pravděpodobností nerozpozná určitý slovesný tvar *je*

**d** - u řetězců slov *do dvou*

V případě této hlásky systém nerozpozná s nejvyšší pravděpodobností nerozpozná předložku *do*

**v** - u řetězců slov *ve velké*

#### **5) další chyby rozpoznávání**

Jedná se o takové chyby rozpoznávání, které nelze zařadit do výše uvedených kategorií. Jedná se zejména o rozpoznaná slova, která vůbec neodpovídala tomu, co stínoví mluvčí do systému nadiktovali.

Příklady:

Brusel skoupý → bruselskou ví; bruselskou prý; bruselskou i; Rosa skoupí

vyřčenou → listinou

stažením → stařenky

v tomto → Waltru tom

pod → Waltr

## 7.9 Klasifikace zjištěných chyb

V předchozích podkapitolách byly během analýzy zjištěny chyby, které byly způsobeny jak činností stínových mluvčích, tak systému automatického rozpoznávání řeči. V této podkapitole budou tyto chyby rozděleny do tří kategorií podle závažnosti.

### 7.9.1 Závažné chyby

Jedná se o chyby, které výrazným způsobem mění význam výsledného titulku, zkreslují ho, nebo vytváří význam zcela nový.

Do této kategorie lze zařadit následující chyby, které byly zanechány ve výsledných komunikátech:

Antonyma, která zcela mění význam rozpoznaného textu,

Výrazy rozpoznané chybně nebo navíc,

Nesprávně vložená interpunkce.

### **7.9.2 Standardní chyby**

Jedná se o chyby, které mohou narušit rozpoznaný text a mohou způsobit uživateli skrytých titulků obtíže rekonstruovat z výsledného titulku význam promluvy.

Do této kategorie lze zařadit následující chyby, které byly zanechány ve výsledných komunikátech:

Chybné koncovky u rozpoznaných slov, zejména u slovesných tvarů,

Příliš dlouhá souvětí, ve kterých může uživatel skrytých titulků při jejich čtení ztratit orientaci,

Opis výrazu, který se nenachází ve slovníku systému rozpoznávání řeči, jinými slovy, což může částečně vést k obtížím při pochopení významu skrytého titulku.

### **7.9.3 Chyby téměř neovlivňující pochopení výsledného titulku**

Do této kategorie lze zařadit problémy systému rozpoznávání řeči způsobené chybným rozpoznáváním nepravých homonym a nedostatků ve shodě podmětu s přísudkem, která jazykový model systému nedokáže spolehlivě identifikovat, a stínový mluvčí nemůže žádným způsobem ovlivnit jejich správné rozpoznání. Uživateli skrytých titulků ale toto nijak neztěžuje pochopení výsledného skrytého titulku.

## 8 ZÁVĚR

Tato diplomová práce představuje koncept stínového překloubání jako nové metody opatrování živého televizního vysílání skrytými titulky. Stínoví mluvčí jsou vysoce kvalifikovaní pracovníci, kteří během tvorby skrytých titulků za pomoci systému rozpoznávání řeči musí disponovat schopností vykonávat několik paralelních činností najednou, což na ně klade nesmírně vysoké nároky. V praxi je běžné, že většina kandidátů na stínové mluvčí nezvládne náročné požadavky, které jsou na ně kladeny již během fáze tréninku, a nemohou funkci stínového řečníka vykonávat.

Během své činnosti vytváří stínoví mluvčí zcela nový a specifický komunikát vykazující vlastnosti psaného textu s určitou mírou prvků specifických pro text mluvený. Úroveň výsledného komunikátu záleží na schopnostech stínového mluvčího nabytých během čtyř fází tréninku.

Cílem této diplomové práce bylo analyzovat komunikáty vytvářené stínovými řečníky ve dvou z fází tréninku a v živém vysílání. K dispozici byla data tří stínových mluvčích, která byla poskytnuta Katedrou kybernetiky Západočeské univerzity v Plzni. V současné době stínoví mluvčí překloubají na jednotlivých kanálech České televize politické debaty, zábavní pořady a sportovní přenosy. Pro účely této diplomové práce byly vybrány komunikáty vytvořené při překloubání debat zejména na politická a ekonomická témata, která tvoří stěžejní hlavní část pořadů překloubaných stínovými mluvčími.

Během výzkumu komunikátů byly identifikovány a zhodnoceny individuální chyby jednotlivých stínových mluvčích, jejichž data byla k dispozici. Dále byly zjištěny chyby tzv. nadindividuální, tedy chyby, kterými se vyznačovali alespoň dva ze stínových mluvčích. Klasifikovány byly také chyby způsobované systémem rozpoznávání řeči, které byly rozděleny do několika kategorií. Jejich identifikace a rozdělení do skupin by mohlo v budoucnu sloužit jako podnět k případnému vylepšení

efektivitu systému rozpoznávání řeči. Případné opravy nedostatků systému by výrazně ulehčily vykonávání profese stávajícím stínovým mluvčím i trénink novým adeptům.

Analyzovány byly nahrávky z druhé fáze tréninku, ve které se kandidáti na stínové mluvčí učí přizpůsobit svůj projev potřebám systému rozpoznávání řeči. Zde se nejčastěji projevovaly chyby zaviněné stínovými mluvčími, které byly způsobeny nesprávnou výslovností a nezřetelnými mezislovními pauzami. Závažné chyby způsobovaly pauzy a drobná zaváhání během artikulace slov, které vedly k rozpoznávání nesmyslných řetězců slov. Stínoví mluvčí do systému také často diktovali zcela odlišná slova, v mnohých případech nespisovná, které systém nedokázal rozpoznat. Problémy v rozpoznání byly také v mnoha případech způsobeny prodlužováním výslovnosti hlásek a hezitačními zvuky.

Dále byly analyzovány nahrávky ze čtvrté fáze tréninku. V této etapě výcviku by již kandidáti na post stínového mluvčího měli být seznámeni se specifiky, která vyžaduje systém automatického rozpoznávání pro efektivní rozpoznávání jejich mluveného projevu. V této fázi také stínoví mluvčí vkládají do rozpoznávaného textu interpunkční znaménka. Čtvrtá fáze již téměř plně simuluje tvorbu skrytých titulků během živého vysílání. Současně s komunikáty vytvořenými ve čtvrté etapě tréninku byly analyzovány také komunikáty, které vytvořili stínoví mluvčí během živého přenosu pořadu Hyde Park. Za nejzávažnější nedostatek v těchto výsledných komunikátech lze považovat fakt, že stínoví mluvčí nedbají na důslednou kontrolu rozpoznávaného textu a odesílají do vysílání chybně rozpoznané výrazy a v mnohých případech zcela nesmyslné věty. Důsledná kontrola rozpoznávaného textu je nutná, protože chybně rozpoznaná slova odeslaná do skrytých titulků mohou zcela změnit význam a znemožnit uživatelům skrytých titulků plnohodnotné pochopení významu původní promluvy. Další typ chyb byl zjištěn ve vkládání

interpunkčních znamének. Stínoví mluvčí v mnohých případech chybovali ve vkládání čárky před spojky *a* a *nebo*. Chybně vložená interpunkční znaménka se objevovala v komunikátech také na konci vět. Často se také objevovaly nedostatky v dělení promluv do vhodných celků. V mnoha případech odesílali stínoví mluvčí do vysílání souvětí o několika větách. Dlouhá souvětí mohou výrazně ztížit uživatelům skrytých titulků orientaci v ději, protože titulek se jim navíc zobrazí s několikavteřinovým zpožděním. Během analýzy byly zjištěny chyby, které nemohou stínoví mluvčí během přemlouvání žádným způsobem ovlivnit. Jedná se zejména o shodu podmětu s přísudkem a správné rozpoznávání homonym. Další typy chyb jsou způsobovány samotnými stínovými mluvčími. Jedná se zejména o nesprávnou výslovnost, nezřetelné koncovky zejména u slovesných tvarů, hezitační zvuky nebo pauzy během vyslovování slov. Všechny tyto nedostatky vedou k chybám v rozpoznávání.

V konečném důsledku bylo na základě analýzy zjištěno, že systém automatického rozpoznávání řeči má své nedostatky, ale největší vliv na celkovou kvalitu výsledného komunikátu má lidský faktor, tedy stínový mluvčí. Kompetentní stínový mluvčí by měl být schopen vytvářet smysluplné, bezchybné a sémanticky relevantní skryté titulky, které umožní jejich uživatelům plnohodnotné pochopení obsahu přemlouvaného pořadu.

Stínové přemlouvání je přelomový přístup v tvorbě skrytých titulků v živém vysílání a zcela jistě se bude rozvíjet i v budoucnosti. Díky přítomnosti kompetentního stínového mluvčího jsou eliminovány téměř všechny nedostatky, které by způsobovalo rozpoznávání řeči pouze za pomoci počítače. Stínový mluvčí může také pracovat v prostředí, které by bylo pro použití počítače nevhodné.



Tato diplomová práce přispěje ke zlepšení efektivity činnosti stávajících sedmi stínových mluvčích, kteří pravidelně opatřují vysílání České televize skrytými titulky, a svým návrhem funkčně orientované taxonomie zásad tvorby komunikátu přispěje k rychlejšímu a jednoduššímu tréninku nových adeptů na post stínového mluvčího. V neposlední řadě slouží jako podnět pro pracovníky podílející se na vývoji systému rozpoznávání na Katedře kybernetiky Západočeské univerzity v Plzni na vylepšení vlastností jimi vyvinutého systému.

## 9 SEZNAM POUŽITÉ LITERATURY A ZDROJŮ

HÁJKOVÁ, Eva. *Český jazyk a kultura mluveného projevu*. Praha: Univerzita Karlova v Praze, Pedagogická fakulta, 2003, 48 s. Texty pro distanční studium. ISBN 80-729-0117-6.

KOPLÍK, Raimund a Věra STRNADOVÁ. *Problematika tvorby skrytých titulků k předtočeným programům*. Vyd. 1. Praha: Česká komora tlumočnicků znakového jazyka, c2008. ISBN 978-808-7153-826.

MÜLLEROVÁ, Olga. *Mluvený text a jeho syntaktická výstavba*. Vyd. 1. Praha: Academia, 1994. 145 s. Studie a práce lingvistické; Sv. 27. ISBN 80-200-0489-0.

PRAŽÁK, Aleš et al. *Captioning of live TV programs through speech recognition and re-speaking*. Text, speech and dialogue: 15th international conference, TSD 2012: Brno, Czech Republic, September 3-7, 2012: proceedings. Berlin: Springer, 2012. Lecture Notes in Computer Science, 2012, 7499, s. 513-519. ISSN 0302-9743. ISBN 978-3-642-32789-6.

PRAŽÁK, Aleš et al. *Four-phase re-speaker training system*. In: Proceedings of the International Conference on Signal Processing and Multimedia Applications: SIGMAP 2011. [S.l.]: INSTICC, 2011. s. 217-220. ISBN 978-989-8425-72-0.

PRAŽÁK, Aleš et al. *Novel approach to live captioning through re-speaking: tailoring speech recognition to re-speaker's needs*. In: 12th Annual Conference of the International Speech Communication

Association 2012. Red Hook: Curran Associates, 2012. s. 1370-1373. ISBN 978-1-62276-759-5.

PRAŽÁK, Aleš. *Trenažér stínových mluvčích - návod pro užívání aplikace*. Trenažér verze 6,4.2.556 ze dne 11. 2. 2014, LVCSR 6.2.0.1284 (8 x CPU + OpenCL). Copyright © SpeechTech, s.r.o. + Západočeská univerzita v Plzni, 2009-2014.

PSUTKA, Josef et al. *Mluvíme s počítačem česky*. Vyd. 1. Praha: Academia, 2006. 746 s. Česká matice technická; roč. 111, č. spisu 502. ISBN 80-200-1309-1.

ŠVEC, Jan et al. *Web text data mining for building large scale language modelling corpus*. Text, speech and dialogue: 14th international conference, TSD 2011: Pilsen, Czech Republic, September 2011: proceedings. Berlin: Springer, 2011. Lecture Notes in Computer Science, 2011, 6836, s. 356-363. ISSN 0302-9743. ISBN 978-3-642-23537-5.

TRMAL, Jan et al. *Online TV captioning of Czech parliamentary sessions*. Text, speech and dialogue: 13th international conference, TSD 2010: Brno, Czech Republic, September 2010: proceedings. Berlin: Springer, 2010. Lecture Notes in Computer Science, 2010, 6231, s. 416-422. ISSN 0302-9743. ISBN 978-3-642-15759-2.

Zákon č. 231/2001 Sb., o provozování rozhlasového a televizního vysílání. In: *Sbírka zákonů*.

### **Další zdroje**

MÜLLER, Luděk. *Automatické titulkování živých pořadů České televize – současný stav a výhled do budoucna*. [online]. [cit. 15. 12.2013.] Konference INSPO 2012, TU Liberec, 2012. Dostupné z www: [http://www.helpnet.cz/data/articles/down\\_52440.pdf](http://www.helpnet.cz/data/articles/down_52440.pdf)

2013 – *Pořady s titulky a/nebo se znakovým jazykem ve vysílání České televize* [online]. [cit. 31. 1. 2014].

Dostupné:[http://img.ceskatelevize.cz/boss/image/contents/sledovanost/ci-sla-o-vysilani/2013/4q2013\\_4\\_neslysi.pdf](http://img.ceskatelevize.cz/boss/image/contents/sledovanost/ci-sla-o-vysilani/2013/4q2013_4_neslysi.pdf)

*Internetová jazyková příručka* [online]. [cit. 20. 3. 2014]. Dostupné z [www.prirucka.ujc.cas.cz](http://www.prirucka.ujc.cas.cz)

*Program MegaWord*. Verze 7.0.2.948 ze dne 17. 3. 2014, LVCSR 6.2.0.1297 build 137 (8 x CPU + OpenCL) Copyright © 2000–2014 SpeechTech, s.r.o.

*Aplikace Trenažér stínových mluvčích*. Trenažér verze 6,4.2.556 ze dne 11. 2. 2014, LVCSR 6.2.0.1284 (8 x CPU + OpenCL). Copyright © SpeechTech, s.r.o. + Západočeská univerzita v Plzni, 2009–2014.

## **10 RESUMÉ**

This thesis deals with the topic of Comparative linguistic analysis of parallel utterances generated by a shadow speaker. This is a new form of making closed captions during live broadcasting with the help of the speech recognition system.

This theses introduces the speech recognition system, the ways of making live closed captions in the main Czech broadcast, Czech Television and the main goal is to analyse the final closed captions made by the re-speakers.

Afterwards this thesis introduces the way of Four-phase re-speaker training system which helps to train the skilled re-speakers.

The final part of the thesis is focused on the linguistic analysis of closed captions generated during the two phases of training and during the live subtitling. The mistakes were divided into categories to help to improve the speech recognition system and the capabilities of current re-speakers and the future candidates.

## **11 PŘÍLOHY**

I. VÝVOJ ZLEPŠOVÁNÍ SCHOPNOSTI PŘEMLOUVÁNÍ  
SÉMANTICKÉHO OBSAHU PŘEMLOUVANÝCH BĚHEM FÁZE  
TRÉNINKU NA PŘÍKLADU DVOU STÍNOVÝCH MLUVČÍCH

II. FUNKČNĚ ORIENTOVANÁ TAXONOMIE ZÁSAD TVORBY  
KOMUNIKÁTŮ BĚHEM STÍNOVÉHO PŘEMLOUVÁNÍ

## I. VÝVOJ ZLEPŠOVÁNÍ SCHOPNOSTÍ PŘEMLOUVÁNÍ SÉMANTICKÉHO OBSAHU NAHRÁVEK PŘEMLOUVANÝCH BĚHEM FÁZE TRÉNINKU NA PŘÍKLADU DVOU STÍNOVÝCH MLUVČÍCH

STÍNOVÝ MLUVČÍ	1. NAHRÁVKA	2. NAHRÁVKA	3. NAHRÁVKA
A	58, 78 %	75, 59 %	80, 66 %
B	68, 19 %	77, 95 %	83, 61 %

## II. FUNKČNĚ ORIENTOVANÁ TAXONOMIE ZÁSAD TVORBY KOMUNIKÁTŮ BĚHEM STÍNOVÉHO PŘEMLOUVÁNÍ

### 1. SLOVNÍ ZÁSoba VHODNÁ PRO DIKTOVÁNÍ DO SYSTÉMU

Do systému automatického rozpoznávání řeči je nutno diktovat pouze spisovné výrazy. Nespisovné výrazy systém nerozpozná.

### 2. ZŘETELNÁ ARTIKULACE

Během analýzy komunikátů bylo zjištěno, že největší počet chyb v rozpoznání byl způsoben nezřetelnou artikulací stínových mluvčích.

Nutné je zřetelně vyslovovat koncovky, zejména u slovesných tvarů, předložky, spojky a krátká, především jednoslabičná slova.

Chyby v rozpoznávání jsou také často způsobovány neadekvátním prodlužováním výslovnosti hlásek, což vede nejčastěji k chybnému rozpoznávání koncovek u slov.

Pokud začínají dva bezprostředně po sobě jdoucí výrazy na stejnou hlásku, nebude jeden z výrazů s největší pravděpodobností rozpoznán. Pokud se tak stane, je vhodné výraz nahradit synonymem nebo ho opsat za pomoci jiných slov takovým způsobem, aby byla zachována sémantická stránka promluvy. Systém rozpoznávání řeči má také v mnoha případech problémy rozpoznat určitý slovesný tvar *je*, na základě zjištění z analýzy nedojde s velkou pravděpodobností k jeho rozpoznání.

Důležitým faktorem jsou také zřetelné pauzy mezi jednotlivými slovy. Pokud není například jasně oddělena předložka od bezprostředně následujícího výrazu, nebude s největší pravděpodobností rozpoznána.

Během stínového přemlouvání je nutno důsledně diktovat celá slova včetně koncovek. Pokud není slovo jasně vysloveno celé, systém rozpozná v mnoha případech pouze jeho část.

### **3. HEZITAČNÍ ZVUKY, HLASITÉ NÁDECHY, HLUK**

Během procesu stínového přemlouvání je nezbytné nevnášet do projevu hezitační zvuky, je nutné správně dýchat takovým způsobem, aby hlasitý proud vzduchu nebyl zachycen mikrofonom a následně rozpoznán jako slovo navíc. Do systému je také nutno diktovat v naprosto tichém prostředí s rukama umístěnými nejlépe na klávesnici, aby nebyly do mikrofonu přenášeny hluky způsobující rozpoznávání slov navíc.

Důležité je také vyvarovat se pauz během vyslovování výrazů diktovaných do systému. Drobná zaváhání způsobují chyby v rozpoznávání.

### **4. NEUSTÁLÁ KONTROLA ROZPOZNAVANÉHO TEXTU**

Během stínového přemlouvání musí stínový mluvčí neustále kontrolovat průběžně rozpoznávaná slova. Chybně rozpoznané výrazy musí stínový



mluvčí opravovat, aby nebyly odesílány do výsledných skrytých titulků. Toto pravidlo musí být dodržováno zejména při přemlouvání pořadů v živém vysílání.

## **5. SOUSTŘEDĚNÍ SE NA ZAČÁTEK PROMLUV**

Sémanticky nejrelevantnější informace bude s nejvyšší pravděpodobností vyslovena původním mluvčím na začátku promluvy. Je tedy důležité soustředit se zejména na úvod jednotlivých promluv.

## **6. ZÁSADY VKLÁDÁNÍ INTERPUNKCE**

Před vstupem do čtvrté fáze tréninku je vhodné nastudovat si základní pravidla vkládání interpunkce do vět. Jedná se především o vkládání interpunkce před následující jevy:

a,

nebo,

než,

přívlastek volný a těsný,

přístavek.

Relevantním zdrojem pro nastudování interpunkčních principů je webová stránka [www.prirucka.ujc.cas.cz](http://www.prirucka.ujc.cas.cz) vytvářená pracovníky Ústavu pro jazyk český AV ČR.

## **7. NÁPRAVA CHYBNĚ ROZPOZNANÝCH ŘETĚZCŮ ODESLANÝCH DO VYSÍLÁNÍ**

Přestože stínoví mluvčí procházejí náročným tréninkem, může se během přemlouvání živého vysílání stát, že se do skrytých titulků odešle chybně rozpoznané slovo, které může v krajním případě zcela změnit význam

původní promluvy. V tomto případě je vhodné do systému nadiktovat slovo „oprava“, které systém rozpozná a bezprostředně za ním nadiktovat opravenou verzi chybně rozpoznaného řetězce. Stínový mluvčí tímto napraví svou chybu a uživatel skrytého titulku není uveden v omyl.

## **8. ZPŮSOBY STÍNOVÉHO PŘEMLouvÁNÍ**

Během tréninku si každý z kandidátů vyvine svůj individuální způsob, jakým přemlouvá nahrávky. Obecně lze identifikovat dva způsoby, jakými je možno pojmout proces stínového přemlouvání.

Prvním způsobem je doslovné přemlouvání, které lze užít zejména v případech, kdy je projev řečníka v televizním vysílání gramaticky správný a významově ucelený. Jedná se zejména o takové typy pořadů, kde je řeč předem připravena. Příkladem může být přemlouvání projevu moderátora Václava Moravce.

V případech, kdy je projev řečníka v televizním vysílání gramaticky nesprávný, neucelený nebo je narušován projevem jiného mluvčího, je nutné původní projev parafrázovat takovým způsobem, aby byl text zobrazený ve výsledném titulku pro jeho uživatele pochopitelný. Zjednodušování lze užít i v případě, je-li projev původního řečníka pro stínového mluvčího příliš rychlý a není ho tak možné přemlouvat doslova.

Obecně platí, že čím více si stínový mluvčí poslechne informací, tím větší má prostor pro parafrázování. Dochází ale k výraznějšímu zpoždění zobrazení výsledního titulku jeho uživateli.

## **9. SLOVO STOJÍCÍ MIMO SLOVNÍK SYSTÉMU**

V případě, že nadiktované slovo není systém schopen rozpoznat, a to ani opakovaně, nenachází se ve slovníku systému. Stínový mluvčí ho tedy musí buď přidat do slovníku, což může výrazně zpomalit zobrazení

výsledného titulku jeho uživateli, nebo se může inkriminovanému výrazu vyhnout a opsat ho jinými slovy, pokud je to možné.

Pokud se stínový mluvčí rozhodne slovo přidat do slovníku, je vhodné si k němu přidat také odpovídající výslovnost.

## **10. VHODNÁ DÉLKA VĚT VE VÝSLEDNÉM KOMUNIKÁTU**

Protože je prostor pro zobrazení titulku omezen počtem znaků, je vhodné namísto dlouhých a komplikovaných souvětí dělit promluvu do kratších celků. Tímto se uživateli skrytých titulků usnadní orientace v ději ve vysílání a také se eliminuje míra chybovosti ve vkládání interpunkce do rozpoznávaného textu.

## **11. STÍNOVÉ PŘEMLouvÁNÍ ŽIVÉHO VYSÍLÁNÍ**

Během tréninku mají stínoví mluvčí k dispozici nahrávky pořadů, které jsou přemlouvány i vytrénovanými stínovými mluvčími v živém vysílání. Klíčovou odlišností mezi tréninkem a živým vysíláním je třívteřinový rozdíl mezi přijímaným zvukem a obrazem. Tento časový rozdíl je důležitý z toho důvodu, aby byla časová prodleva mezi rozpoznáním řetězcem slov a jeho zobrazením na televizní obrazovce co nejmenší. Nicméně tato několikavteřinová prodleva může nově vytrénovaným stínovým mluvčím činit zpočátku potíže.

## **12. RADY A TIPY**

Před samotným stínovým přemlouváním je vhodné procvičit mluvidla, například přečtením několika libovolných vět. Vzhledem k vysokým nárokům na hlasivky stínových mluvčích je důležité mít připravenou nádobu s tekutinou.

Před přemlouváním živého vysílání je vhodné propojit počítač s titulkovacím serverem a televizním vysíláním s několikaminutovým

předstihem. Je to důležité zejména pro kontrolu bezchybného fungování systému.

Zejména při přemlouvání prvního živého přenosu se může stát, že stínový mluvčí bude nervózní a svou nervozitou zapříčiní například rozpoznání chybných výrazů, které se odešlou do vysílání. Je nutné zachovat klid a maximální míru soustředění a dále klidně pokračovat v přemlouvání. Stínoví mluvčí vylepšují své schopnosti nejen během tréninku, ale i během samotného procesu vytváření skrytých titulků v živém vysílání.