

ZÁPADOČESKÁ UNIVERZITA V PLZNI

FAKULTA APLIKOVANÝCH VĚD

KATEDRA KYBERNETIKY

TBD

DIZERTAČNÍ PRÁCE

SRPEN 2019

AUTOR: ING. PETR STANISLAV

ŠKOLITEL: PROF. ING. JOSEF PSUTKA, CSc.

OBOR: KYBERNETIKA

# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
<b>2</b>	<b>Příčiny a léčba trvalé ztráty hlasu</b>	<b>2</b>
2.1	Příčiny ztráty hlasu . . . . .	2
2.1.1	Rakovina hrtanu . . . . .	3
2.2	Rehabilitace hlasu po totální laryngektomii . . . . .	7
2.2.1	Foniatrické metody . . . . .	8
2.2.2	Chirurgicko-protetická metoda . . . . .	12
2.2.3	Hrtanu podobné struktury . . . . .	15
2.2.4	Transplantace hrtanu . . . . .	17
2.2.5	Shrnutí . . . . .	18
<b>3</b>	<b>Řečová rozhraní</b>	<b>21</b>
<b>4</b>	<b>Experimenty</b>	<b>22</b>
4.1	Analýza dat a první modely . . . . .	23
4.1.1	Vytvoření korpusu EL promluv . . . . .	24
4.1.2	Analýza získaných dat . . . . .	26
4.2	Provnání člověk vs. stroj . . . . .	26
4.3	Protahování na příýnácích . . . . .	27
4.4	Modely pracující s protaženými daty . . . . .	27

<b>5</b>	<b>Trenažér</b>	<b>29</b>
<b>6</b>	<b>Závěr</b>	<b>30</b>

# Kapitola 1

## Úvod

# Kapitola 2

## Příčiny a léčba trvalé ztráty hlasu

Lidská řeč tvoří jeden ze stavebních kamenů lidského dorozumívání. Pro člověka postiženého dočasnou či trvalou ztrátou hlasu představuje běžná lidská komunikace mnohem náročnější úkol než pro člověka zdravého. Takový jedinec se musí dennodenně potýkat s problémy, které by za normálních okolností řešit nemusel. V mnoha případech doprovází ztrátu hlasu i zvýšená psychická zátěž, například strach z reakce okolí, a proto se problematice lidí trpících ztrátou hlasu věnuje nemalá pozornost. V této kapitole si nejprve v části 2.1 přiblížíme možné příčiny ztráty hlasivek, a tedy i trvalé ztráty hlasu, a následně v části 2.2 představíme dostupné metody rehabilitace hlasu.

### 2.1 Příčiny ztráty hlasu

Nejčastěji je trvalá ztráta hlasu zapříčiněna chirurgickým zákrokem zvaným totální laryngektomie<sup>1</sup> neboli úplné odstranění hrtanu. Odstraněním hrtanu, a tím i hlasivek (glottis), přichází člověk o schopnost rozvibrovat vzduch vycházející z plic, který je dále modulován artikulačním ústrojím. Nejběžnější příčinou vedoucí k totální laryngektomii představuje rakovina hrtanu v pokročilém stádiu. V mnohem nižší míře je na vině rakovina hltanu či poškození hrtanu automobilovou nebo jinou traumatickou

---

<sup>1</sup>laryngektomie: larynx, laryngos - hrtan, ectos, ectomia - odstranění, vynětí

nehodou. Podle [0] přibude v České republice ročně přibližně 400 nových onemocnění rakoviny hrtanu, z toho je přibližně jedna třetina léčena pomocí totální laryngektomie. To představuje více než 100 nových případů trvalé ztráty hlasu každý rok.

### 2.1.1 Rakovina hrtanu

Jak již bylo zmíněno, rakovina hrtanu je jedním z hlavních důvodů odstranění hrtanu. Tento typ rakoviny postihuje převážně muže ve věku 50-60 let. V posledních letech je však zřejmý trend snižujícího se průměrného věku pacientů [0]. Z celkovém počtu pacientů zhruba 20% představují ženy.

Přesná příčina vzniku nádorového onemocnění hrtanu doposud není známa, ale z průzkumů je zřejmá korelace mezi vznikem rakoviny a konzumací alkoholu či kouřením. Jinými slovy, mezi rizikovou skupinu patří lidé, kteří jsou pravidelně vystavováni vlivu kouření, ať již aktivně (sami kouří) či pasivně (vdechují cigaretový kouř) a zároveň si dopřávají nemalé množství alkoholu. Podle [0] 90% pacientů aktivně kuří.

### Příznaky onemocnění

Vznikající nádorové onemocnění v oblasti hrtanu se může projevovat různými způsoby. Mezi hlavní faktory ovlivňující počáteční příznaky patří umístění a velikost nádoru.

Již ve velmi raném stádiu, kdy je nádor umístěn přímo na hlasivkách, je příznakem **chrapot**. Ve většině případů se samozřejmě jedná o krátkodobé postižení hlasivek virovou infekcí. Nicméně pokud chrapot trvá déle než tři týdny, je již doporučováno navštívit odborného lékaře. U nádorů nacházejících se ve vchodu do hrtanu a v polykacích cestách se mohou jako příznak objevovat **polykací obtíže**. Mezi další možné příznaky patří **bolesti v krku**, jednostranné bolesti vystřelující do ucha či nepříjemný pocit při polykání. I v tomto případě krátkodobý výskyt nemusí nutně znamenat rakovinu hrtanu, nicméně při obtížích trvajících déle než měsíc je doporučováno důkladné vyšetření lékařem. Na základě umístění nádoru se může objevovat

**dráždivý kašel** s možným vykašláváním krve. Dalším možným příznakem je vznik **zduření na krku**. V tomto případě je vhodné neprodleně vyhledat lékaře.

Z výše uvedených příznaků je zřejmé, že prvotní indicie o vážném onemocnění mohou být podceněny a tím značně snížena šance na plné uzdravení pacienta. V případě včasného diagnostikování rakoviny hrtanu či hltanu je možnost úplného vyléčení pacienta bez trvalých následků více než 90% [0].

### **Léčba nádorového onemocnění**

U nádorovitých onemocnění se v drtivé většině případů využívá **chirurgické léčby**, **aktinoterapie** neboli ozařování a **chemoterapie**. Nejinak tomu je i v případě rakoviny hrtanu a polykacích cest obecně. Majoritní část pacientů je zpravidla nejprve konfrontována s chirurgickou léčbou. Mezi nejčastější zákroky patří **tracheostomie**, parciální laryngektomie, **totální laryngektomie** a chordektomie. V rámci této práce budou blíže popsány pouze léčebné postupy přímo související s úplným odstraněním hrtanu a hlasivek.

Totální laryngektomie (TL), jak už název napovídá, představuje chirurgický zákrok, při kterém je úplně odstraněn hrtan. V určitých případech může být odstraněna i menší či větší část hltanu. Tento zákrok, ve spojení s léčbou rakoviny hrtanu, poprvé vykonal Dr. Theodor Billroth 31. prosince roku 1873 ve Vídni [0] a do dnešní doby přežil v podstatě nezměněn. Cílem této operace je odstranění orgánu zasaženého rakovinným bujením, a tedy zamezení dalšího šíření nemoci. Součástí hrtanu je také hrtanová příklopka (latinsky epiglottis), která zamezuje vdechnutí potravy nebo tekutin do dýchacích cest. Po odstranění hrtanu by tedy potrava a tekutiny mohly být vdechnuty do plic a z tohoto důvodu jsou jícen a průdušnice trvale odděleny. Rozdíl mezi zdravým člověkem a osobou po totální laryngektomii je znázorněn na obr. 2.1.

Odstraněním spojení průdušnice a jícnu je zahrazena cesta vzduchu do plic. Z tohoto důvodu je nezbytné společně s TL vykonat také **tracheostomii**. Cílem tohoto

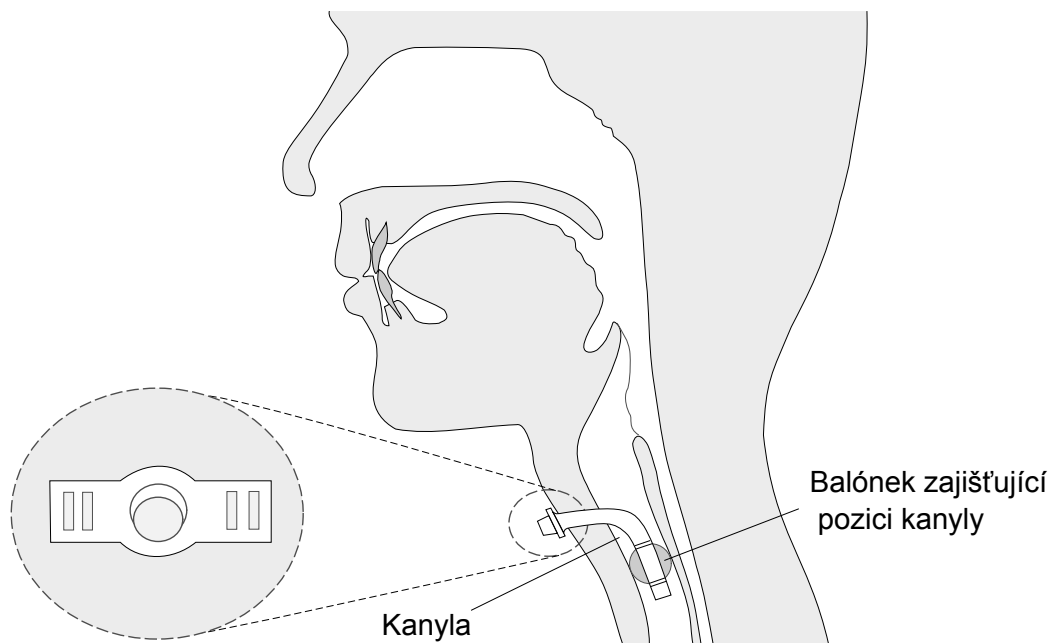


Obrázek 2.1: (A) Schéma dýchacích cest zdravého člověka (B) Schéma dýchacích cest po totální laryngektomii

zákroku není léčba nádorovitého onemocnění, nýbrž vytvoření vstupu pro vzduch směřující do plic a z nich. Samotný zákrok se využívá i v situacích, kdy dojde k uzávěře hrtanu a postižená osoba se dusí. K tomuto může dojít například při alergické reakci na včelí bodnutí, otoku hrtanu, úrazu apod. Při tracheostomii se provádí řez skrz kůži a průdušnici. Do vzniklého otvoru se zavádí kanyla, která slouží k dýchání. Místo výkonu zákroku a princip kanyly je znázorněn na obr. 2.2. Pro lepší názornost je znázorněna tracheostomie se zavedenou kanylou u zdravého člověka. Výsledek operace může být dočasný (například v případě alergické reakce) nebo trvalý.

Další významnou metodou využívanou k léčbě nádorových onemocnění představuje **aktinoterapie** neboli léčba ozařováním. Podstatou postupu je periodické vystavování buňky ionizujícímu záření. Energie z tohoto záření je předávána buňce, která je tím poškozována. Tento postup se opírá o předpoklad, že nádorové buňky jsou náchylnější k poškození. Záření ovšem ovlivňuje i zdravé buňky, a proto je tato léčba pro organismus velkou zátěží. Aktinoterapie je možné využít jednak u případů, kdy je cílem terapie úplné vyléčení, tak i v případech, kdy naprosté odstranění onemoc-





Obrázek 2.2: Tracheostomie

nění není možné. V druhém případě je metoda využívána k prodloužení a zkvalitnění života [0].

Aktinoterapii je možné využít jako hlavní léčebnou metodu (primární aktinoterapie) nebo i ve spojení s ostatními metodami. V případě primární se k léčbě využívá pouze ozařování a cílem je úplné odstranění všech defektních buněk. Z podstaty metody, zejména dopadů léčby na lidský organismus, je zřejmé, že tímto způsobem je ve většině případů možno léčit pouze malé nádory.

Ve spojení s chirurgickou léčbou rozlišujeme předoperační, pooperační nebo tzv. sandwich (tj. před a po chirurgickém zákroku) aktinoterapii. Předoperační ozařování je užíváno v případech, kdy není možné původní nádor vyoperovat. Cílem je zmenšení tumoru do takové míry, aby jej bylo možné chirurgicky odstranit. Někdy je předoperační ozařování spojeno s chemoterapií. U pooperační aktinoterapie je záměrem odstranění potencionálních mikroskopických zbytků tumoru, které by mohly znovu začít růst.

Velmi často se ve spojení s léčbou rakoviny mluví i o proceduře zvané **chemoterapie**. Podstatou je podávání léků zastavujících buněčné dělení, tzv. cytostatik. Zjednodušeně řečeno se jedná o velmi toxický koktejl látek sloužící k zahubení buňky tím, že poškodí určitou její část a zastaví tak proces dělení. Na tuto léčbu jsou citlivé převážně rychle se dělící buňky. Právě defektní buňky v tumoru mají obvykle určitým způsobem poškozeny opravné mechanismy a cytostatiky zasažená rakovinná buňka tak s větší pravděpodobností zahyne. Samozřejmě nelze u chemoterapie hovořit o přesně zacílené léčbě. Cytostatika postihují všechny buňky v lidském těle, a proto je možná namíste srovnání s kobercovým bombardováním. S aplikací cytostatik je tak spojena celá řada vedlejších rizik. Mezi nejzávažnější patří poškození ledvin nebo poškození krvetvorby.

Z výše uvedeného je zřejmé, že postižená osoba má velkou šanci na kompletní vyléčení. V mnoha případech má však pacient trvalé následky (trvalá ztráta hlasu) z důvodu podcenění prvotních příznaků vážného onemocnění.

## 2.2 Rehabilitace hlasu po totální laryngektomii

Nesporná výhoda totální laryngektomie neoddiskutovatelně spočívá v likvidaci primárního nádorového onemocnění. Následky operace však s sebou nesou obrovský zásah do kvality života pacienta. Okem nejviditelnější změnu představuje přítomnost tracheostomie a s ní spojený způsob dýchání. Tato skutečnost má spoustu, na první pohled ne úplně očividných, následků. Postižený člověk ztrácí přirozené zvlhčování, ohřev a filtraci vdechovaného vzduchu, jež má za následek vyšší náchylnost k respiračním onemocněním. Příčina spočívá v průchodu vzduchu do průdušnice přes tracheostomii a nikoli přes nosní dutiny.

Pro samotného pacienta je však nejspíše nejobtížnější se vypořádat s trvalou ztrátou vlastního hlasu. Z tohoto důvodu se již samotný autor procedury doktor Billroth

zaobíral otázkou rehabilitace hlasu. Jeho první pokusy s kovovou tracheostomickou kanylou sice umožňovaly pacientovi hovořit, ale svou konstrukcí pacienta spíše ohrožovaly na životě [0]. Proto se více uchytila metoda tzv. jícnového hlasu [0]. Ve stejnou dobu, tedy přibližně začátkem minulého století, se začaly objevovat první interní a externí hlasové aparáty. V současnosti je rehabilitace hlasu možná pomocí:

- **foniatrických metod**, mezi které patří jícnový hlas a elektrolarynx,
- **chirurgicko-protetickým způsobem**, který spočívá ve vytvoření kanálku skrze stěnu mezi průdušnicí a jícnem,
- **vytvoření hrtanu podobných struktur chirurgickým způsobem**,
- **transplantace hrtanu**.

Z uvedeného výčtu se může zdát, že máme k dispozici relativně širokou škálu možností, jak pacientovi vrátit schopnost vyjadřování pomocí mluvené řeči. Ovšem je nutné si uvědomit, že je potřeba volit konkrétní metodu podle stavu a možností pacienta. Jinými slovy, ne každá metoda se hodí pro každého pacienta a žádná z metod není univerzální pro všechny pacienty.

### 2.2.1 Foniatrické metody

Ačkoli odstranění hrtanu vyústí ve ztrátu hlasu, neznamena to, že by byla úplně eliminována schopnost produkovat řeč. V procesu vytváření hlasu zastává odstraněný orgán pouze (i když velmi zásadní) roli generátoru zvuku. Zbylé orgány (hrdelní, nosní a ústní dutina a další) zůstávají nedotčeny a mohou i nadále plnit svou funkci. Logicky se tak nabízí myšlenka nahradit chybějící zdroj zvuku jiným. Mezi nejpoužívanější metody patří jícnový hlas a elektrolarynx.

## Jícnový hlas

Počátek této metody se datuje do roku 1922, kdy si prof. MUDr. Miloslav Seeman [0] uvědomil, že funkci štěrbiny mezi hlasivkami (rima glottidis) přebírá tzv. pseudo-glottis, která se vytváří na úrovni horního jícnového svěrače. Zároveň vypracoval a popsal metodiku vytváření jícnového hlasu, při které se vzduch neplní do plic, ale do jícnu. Tato metoda se nazývá **aspirační**. Princip spočívá v aktivním otevření jícnového svěrače, nasáváním a vtlačováním vzduchu do jícnu pomocí polykání. Naplněním jícnu vzduchem si pacient připravuje potřebný vzduch k následné eruktaci<sup>2</sup> vzduchu a produkci řeči. Vlastní jícnový hlas poté vzniká na přechodu jícnu a hypofaryngu (spodní část hltanu). V této oblasti horního jícnového zúžení dochází k rozkmitání sliznice a podslizniční vrstvy a produkci zvuku, který je následně modulován stejně jako v případě přirozené produkce řeči. Princip tvorby „základního“ tónu jícnového hlasu je znázorněn na obr. 2.3.



Obrázek 2.3: Princip tvorby jícnového hlasu. Průchodem vzduchu přes zúžení vzniká základní tón jícnového hlasu.

---

<sup>2</sup>eruktace - latinsky název pro proces říhání (popřípadě krkání), při kterém dochází k úniku plynů pocházejících ze žaludku dutinou ústní.

Kromě aspirační metody je ještě možné se setkat s metodou **injekční**. Hlavní rozdíl spočívá v principu plnění vzduchu do jícnu. Při aspirační metodě se využívá polykání, zatímco v tomto případě je využito kořene jazyka, kterým je vzduch vtlačován do jícnu. Následný princip produkce hlasu je již shodný s původní metodou. S tímto principem se můžeme setkat u pacientů, kterým byla při laryngektomii odstraněna jazylka a aspirační náplň není možná.

Proces učení jícnového hlasu by měl začít co možná nejdříve po operaci. Pokud je to možné, tak se s výukou začíná ještě za pobytu pacienta na ORL klinice nebo krátce po propuštění. V první fázi se pacient učí pouze slabiky sestávající z explosivy a souhlásky. Postupně se však přidávají slabičné shluky, které sice nedávají smysl, ale pomáhají v osvojení potřebné techniky. V případě úspěšného zvládnutí se přistupuje k nácviku frází a souvislé řeči. Potřebnou dobu k nácviku jícnového hlasu nelze přesně určit, protože je závislá na mnoha faktorech. V literatuře se uvádí, že je potřeba 30 až 50 hodin velmi intenzivního tréninku k osvojení jícnového hlasu.

Míra úspěšnosti nácviku srozumitelného hlasu se uvádí v rozsahu od 14% do 75%. Takto obrovský rozsah značí o mnoha faktorech, které mohou ovlivnit úspěšné osvojení jícnového hlasu. Mezi možné příčiny neúspěchu patří fyziologické nebo anatomické problémy, psychologické problémy, nebo jednoduše neadekvátní podpora při řečové terapii [0]. Velkou roli také hraje snaha a odhodlání samotného pacienta.

Nepopíratelnou výhodou této techniky rehabilitace je nezávislost pacienta na lékaři po úspěšném osvojení jícnového hlasu a permanentní oddělení dýchacích a polykacích cest bez rizika vniknutí potravy do dýchacích cest. Mezi nesporné výhody také patří volné ruce při vytváření řeči. Za nevýhody se obecně považuje srozumitelnost produkovaného hlasu. Je to způsobeno jednak „břišním“ zabarvením, které je už z podstaty metody přítomné, a dále také nízkou intenzitou a krátkou výdrží při tvorbě tónu. Za negativum se dá také považovat množství pacientem vynaloženého úsilí potřebného k osvojení techniky. Velmi často se také mluvčí ostýchají jícnový

hlas používat, protože mají pocit, že je společensky nevhodné dorozumívat se formou blízkoúhlové řeči. Z tohoto důvodu se odhaduje, že v běžném životě využívá jícnový hlas pouze 20 až 30% pacientů, kteří se začali tuto techniku učit [0].

### **Elektrolarynx**

Rehabilitace hlasu pomocí elektrolarynxu se řadí mezi tzv. elektromechanické metody. Princip spočívá v přikládání zařízení, které obsahuje generátor zvuku nazývaný elektrolarynx. Přiložením do oblasti spodiny úst a aktivací zařízení se generovaný zvuk a vibrace přenášejí do dutiny ústní a dalších přilehlých artikulačních orgánů. Následnou artikulací je pacient schopen hovořit. Znázorněno na obr. 2.4.



Obrázek 2.4: Princip rehabilitace hlasu pomocí elektrolarynxu.

Takto generovaná řeč se vyznačuje několika charakteristickými rysy. V první řadě řeč budí velmi mechanický dojem. Důvodem je samozřejmě samotný elektrolarynx, jelikož se jedná o elektromechanický generátor zvuku s konstantním buzením, je také základní frekvence produkovaného hlasu více či méně konstantní. Řečník tak má velmi

omezené možnosti, jak řeč emotivně zabarvovat. V průběhu času se objevily snahy průběžně měnit frekvenci zařízení a tím ovlivňovat základní frekvenci produkované řeči [0, 0, 0]. Hlavním problémem všech těchto zařízení je docílit změnu fundamentální frekvence na základě toho, co chce řečník říci. V současné době existují pouze experimentální zařízení, která umožňují ve velmi omezené míře změnu frekvence [0]. Další charakteristický rys představuje nižší srozumitelnost řeči, která se ještě snižuje s rostoucím okolním hlukem. Velmi často se stává, že posluchač, který se s takto produkovanou řečí setkává poprvé, není schopen plně porozumět. Se srozumitelností souvisí i další charakteristický rys, kterým je přítomnost zvukového podkresu produkovaného samotným přístrojem.

Za hlavní výhodu elektrolarynxu se považuje rychlost osvojení schopnosti produkovat řeč. Zároveň je tato metoda vhodná pro téměř všechny pacienty postižené ztrátou hlasu způsobenou léčbou karcinomu hrtanu. Z tohoto důvodu se hojně užívá u pacientů, kteří si neosvojili jícnový hlas nebo u nich není možné využití ostatních chirurgických metod. Za nevýhody se obecně pokládá kvalita produkované řeči, tedy monotonní a mechanicky znějící hlas. Dále potom zaměstnání jedné ruky držením nebo spouštěním zařízení.

Samostatnou kapitolou může být psychologický dopad na pacienta. Stejně jako u jícnového hlasu se řeč produkovaná pomocí elektrolarynxu jeví odlišně od řeči přirozené. Navíc se ještě přidává potřeba využití nějakého zařízení. Člověk proto v mnoha případech cítí ostych a bojí se na veřejnosti mluvit.

### 2.2.2 Chirurgicko-protetická metoda

Další možnost rehabilitace hlasu představuje tracheoezofageální (zkr. TE) protéza. První zmínka o vytvoření fistule<sup>3</sup> mezi průdušnicí a jícnem pochází z roku 1932.

---

<sup>3</sup>fistule (česky píštěl) je abnormální otvor mezi dvěma dutými orgány, nebo mezi dutým orgánem a kůží.

V tomto roce doktor Guttman poprvé vytvořil tracheoezofageální shunt<sup>4</sup> („umělá píštěl“). Hlavní myšlenka spočívá ve vytvoření cesty prostřednictvím píštěle, pomocí které u tracheostomovaného člověka může proudit vzduch z plic do úst. Za normálních okolností vzduch proudí skrze tracheostomii a do úst se tak nedostane. Zacpe-li si pacient stomu, může proud vzduchu proudit skrze píštěl do úst. Vzduch procházející přes fistuli naráží do stěn jícnu a je rozvibrován. Tyto vibrace jsou následně modulovány pomocí artikulačních ústrojí a tak vzniká řeč. Tento ojedinělý zákrok otevřel cestu k chirurgické hlasové rehabilitaci. Vzniklo několik operačních metod, které se navzájem lišili víceméně jen umístěním fistule [0].

Hlavní snahou chirurgů bylo vytvoření bezpečné, správně nasměrované píštěle umožňující tvorbu hlasu. Bohužel v mnoha případech byly tyto zákroky spojené s vážnými komplikacemi (infekce, zápalý či těžká krvácení). Důležitým problémem, se kterým se jednotliví tvůrci museli vypořádat, byla stálost vytvořeného otvoru tak, aby jím neprotékaly tekutiny špatným směrem a nedocházelo k zatékání do dýchacích cest a orgánů. Jelikož se jednalo o velmi náročné techniky, a bylo s nimi spojeno velké množství rizik, došlo v 80.letech 20.století k opadnutí snah tyto metody aplikovat.

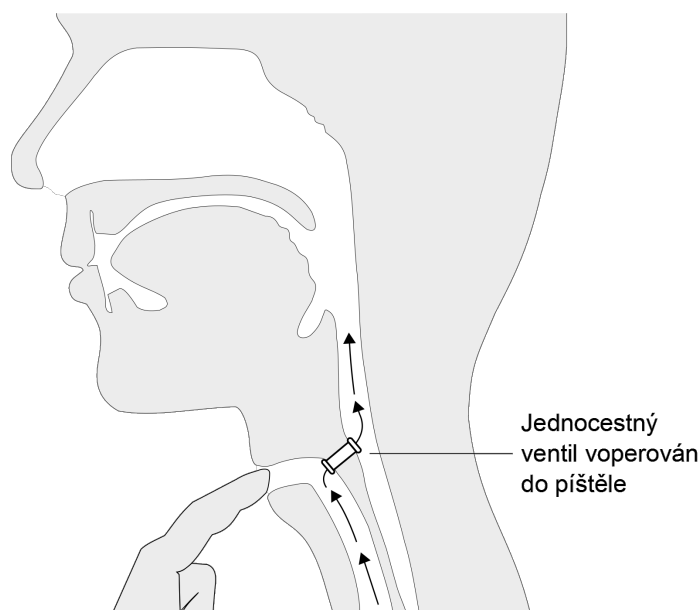
Svou renesanci zažily s vložením jednocestného ventilu, který umožňoval pouze jednosměrný průchod tekutin skrze píštěl, jak je ilustrováno na obr. 2.5. První komerčně dostupná protéza se objevila v 80.letech 20.století v USA. Na obr. 2.6 jsou zobrazeny příklady různých typů protéz. Na používané protézy jsou kladeny přísné nároky a musí vyhovovat určitým požadavkům. Předně se musí vyrábět z biokompatibilního materiálu, který odolává biodegradaci. Tím je zaručena dlouhodobá trvanlivost a správná funkce. Potřebný tlak k otevření faryngoezofageálního segmentu by měl být co nejnižší, aby bylo možné vytvářet plynulou řeč. První vyráběné protézy měly tento tlak příliš vysoký a omezovaly tak množinu potencionálních pacientů. Nejmodernější

---

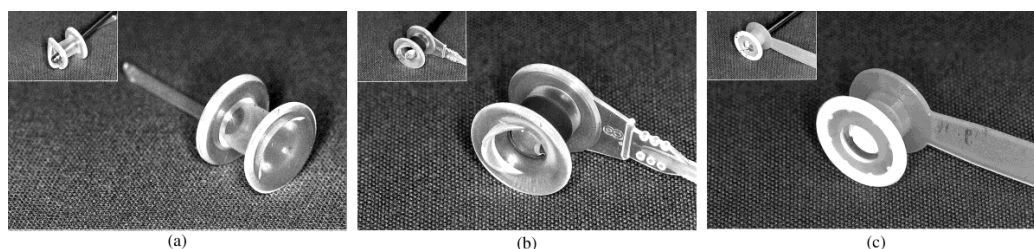
<sup>4</sup>shunt - kanál, kterým je tekutina odkloněna z přirozené dráhy. Tento kanál může být vytvořen chirurgicky nebo pomocí syntetické trubice.



protézy se již vyznačují velmi nízkým otevíracím fonačním tlakem. V neposlední řadě by měla být protéza samofixační a snadno vyměnitelná.



Obrázek 2.5: Průchod vzduchu tracheoezofageální protézou.



Obrázek 2.6: Ilustrace používaných TE protéz (a) Gronigenova nízkotlaká protéza, (b) Provox2 a (c) Blom-Singer protéza.

V praxi se používá několik druhů protéz. Hlavním rozdílem mezi nimi však je zda se pacient přímo účastní výměny ventilu, jehož fundamentální funkcí je vytvoření průchodu pro vzduch proudící z průdušnice do jícnu. U protéz, které jsou vyměňovány operačně, se doba používání pohybuje od 3 do 6 měsíců. Tento interval velmi významně ovlivňuje tvorba biofilmu na povrchu náhrady. K tvorbě dochází následkem přímého kontaktu protézy s tělními tekutinami a potravou. Rychlost tvorby biofilmu

ovlivňuje tvar a materiál, ze kterého je náhrada vytvořena [0]. U typů, které si nositel může měnit sám, se předpokládá, že budou čištěny nebo měněny přibližně jednou za dva týdny.

Samotný zákrok zavedení protézy je možné provést zároveň s výkonem totální laryngektomie (tzv. primární zavedení hlasové protézy) nebo až po zotavení pacienta z náročné léčby nádorového onemocnění (tzv. sekundární zavedení). Primární zavedení umožňuje začít s hlasovou rehabilitací krátce po odstranění hrtanu. Zároveň pacient nemusí v krátké době podstupovat druhou operaci, při které by se vkládal jednocestný ventil do vytvořené fistule.

V praxi se ukázalo, že úspěšnost rehabilitace je více než 80% [0]. Důležitým faktorem, stejně jako u jícnového hlasu, je funkčnost faryngozofageálního segmentu. Dále také otvírací tlak horního jícnového svěrače. Hlas tvořený protézou se vyznačuje vysokou kvalitou, dobrou srozumitelností, individuálním zabarvením a relativně dlouhou fonační dobou dosahující průměrně 20 sekund [0]. Oproti jícnovému hlasu není potřeba tak intenzivní edukace pacienta k plnému osvojení hlasu. V současnosti se jedná o nejpoužívanější metodu rehabilitace hlasu.

### 2.2.3 Hrtanu podobné struktury

S rozvojem mikrovaskulárních<sup>5</sup> transplantátů se začaly objevovat postupy, které umožňovaly rehabilitovat hlas pouze pomocí chirurgického zákroku. Tyto techniky umožňují permanentní spojení hypofaryngu s tracheou pomocí vlastní tkáně pacienta.

První takovouto metodu představil v roce 1984 doktor Ehrenberger [0], který popsal tzv. „řečový sífón“ (angl. **speech siphon**). Tento sífón je vytvořen z části tenkého střeva zvané lačník (jejunum). Spojení mezi hrtanem a hltanem je dvakrát esovitě zahnuto tak, aby bylo minimalizováno riziko sekundární aspirace. Schéma „řečového sífónu“ podle Ehrenberga je znázorněno na obr. 2.7 A. Již na první pohled

---

<sup>5</sup>mikrovaskulární - část oběhového systému složeného z nejmenších cév, jako jsou kapiláry, žilky aj.

je zřejmé, že se jedná o velmi náročný chirurgický zákrok. První články publikované autorským kolektivem prezentovaly velmi dobré funkční výsledky metody. Podle [0] bylo doposud operováno přibližně 60 pacientů.

V roce 1990 byla popsána laryngoplastika podle Hagena. V tomto případě se vytváří tzv. **neolarynx**, k jehož vytvoření se používá štěp z předloktí. Vnitřek neolaryngu je kryt kůží. Neoglottis je vyztužen chrupavkou a překrývá vchod do neolaryngu tak, aby nedocházelo k sekundární aspiraci. Laryngoplastika podle Hagena je znázorněna na obr. 2.7 B. Doposud bylo operováno přibližně 300 pacientů [0].



Obrázek 2.7: A) Schéma „řečového sifónu“ tak jak jej představil Ehrenberg. B) Laryngoplastika podle Hagena

Bohužel v současné době tyto metody nenacházejí širší uplatnění. Především je to způsobeno chirurgickou náročností samotných metod, kvůli které se velmi těžko prosazují na dalších pracovištích. Dalším aspektem, který limituje tyto metody, je vliv na samotného pacienta. Metody předpokládají další chirurgický zákrok vykonaný po totální laryngektomii. Tento zákrok představuje další zátěž pro pacienta nemluvě o možných komplikacích. I přes nedostatky těchto metod je pochopitelná snaha lékařů o intenzivní výzkum v této oblasti. Při úspěšné léčbě je pacient schopen produkovat

hlas velmi dobré kvality a ve většině případů nepotřebuje žádnou péči ze strany lékařů ORL.

## 2.2.4 Transplantace hrtanu

Nejkomplexnější možnost rehabilitace hlasu představuje transplantace hrtanu. V tomto případě pacient obdrží implantovaný hrtan od dárce. Pokud je transplantace úspěšná, přebírá transplantovaný orgán plně funkci původního orgánu a velmi významně zvyšuje šance pacienta na plné zotavení bez trvalých následků.

První informace spojené s výzkumem možností provedení transplantace hrtanu se objevují již v 60. letech 20. století<sup>6</sup>. Přesto byla první totální hrtanová transplantace provedena až profesorem Marshallm Stromem v roce 1998 [0] a do dnešních dnů byly provedeny pouze 2 kompletní transplantace.

Prvním pacientem, který podstoupil transplantaci, byl čtyřicetiletý muž z USA. K laryngektomii v jeho případě vedla motocyklová nehoda, při které si pacient rozdrtil hrtan. K incidentu došlo 20 let před transplantací. Před zákrokem používal k produkci řeči elektrolarynx. Dárce orgánu byl taktéž čtyřicetiletý muž, který zemřel na mozkové aneurysma. Úspěch transplantace se na příjemci projevil již třetí den po operaci, kdy poprvé po 20 letech promluvil (vyslovil anglické slovo „hello“). Přibližně po 36 měsících od transplantace byl produkován hlas srovnatelný s hlasem zdravého člověka. Podle vlastních slov pacienta se po operaci jeho kvalita života „nesmírně“ zlepšila. [0] Doposud poslední úspěšně vykonaná transplantace byla zaznamenána v říjnu 2010.

Mezi hlavní důvody takto malého počtu zákroků patří množství pacientů vhodných pro tuto proceduru. Jelikož se jedná o transplantaci dárcovského orgánu je nutné použití imunosupresiv, tedy medikamentů zabráňujících odmítnutí orgánu. Imunosupresiva jsou však v současné době nepoužitelná u lidí trpících rakovinou hrtanu z důvodu

---

<sup>6</sup>Vůbec první úspěšná transplantace orgánu (ledvin) se uskutečnila v roce 1954.

velmi vysokého rizika rozšíření rakoviny [0]. Další problém představuje náročnost samotného zákroku. Předně je potřeba provést reinervaci a obnovení krevního oběhu v implantovaném orgánu. U první provedené transplantace se nepodařilo dosáhnout kompletní reinervace. Výsledkem tak byl velmi kvalitní generovaný hlas, ale zároveň nebylo možné pomocí hrtanu zabezpečit bezproblémové dýchání a bylo proto nutné ponechat tracheostomii.

Poslední výzkum v oblasti imunosuprese však naznačuje, že by v dohledné době mohlo dojít k pokroku a umožnit transplantaci hrtanu i u lidí trpících rozsáhlou rakovinou v oblasti krku [0]. Prozatím je však tato metoda vhodná pro pacienty netrpící rakovinou, případně ty, u kterých převažovaly benigní nádory a již 5 let nedošlo k recidivě.

### 2.2.5 Shrnutí

Rehabilitaci pacientů, kteří prodělali chirurgické odstranění hrtanu, je ve vyspělých zemích věnována značná pozornost, jelikož následky této operace, oproti jiným druhům léčby, velmi významně ovlivňují kvalitu života pacientů. V první řadě se léčený musí vyrovnat se ztrátou hlasu. Tato situace je již sama o sobě velmi náročnou psychickou zkouškou. Ztráta hlasu je však pouze jedním z vícero problémů, se kterými je potřeba se vypořádat. Mezi další patří možná ztráta čichu či vyšší náchylnost k respiračním onemocněním. Neméně významnou roli sehrává i fyzická odlišnost a z toho pramenící psychická zátěž pacienta po absolvované léčbě.

V současnosti nejpoužívanějšími metodami rehabilitace hlasu jsou **tracheoezofageální píštěl** (popsáno v části 2.2.2), **jícnový hlas** (2.2.1) a použití **elektrolarynxu** (2.2.1). Existují samozřejmě i další a přehled v současnosti používaných je uveden v tab. 2.1.

Většina pacientů je tedy rehabilitována pomocí tracheoezofageálního píštěle, který principiálně vychází z jícnového hlasu, jehož negativa se snaží eliminovat. O úspěchu

	Kvalita	Výhody	Nevýhody
<b>Tracheoezofageální píštěl</b>	Vysoká	Vysoká míra osvojení, dlouhá fonační doba	Zanášení píštěle a s ním spojené čištění, případně dodatečná lékařská péče
<b>Jícnový hlas</b>	Dobrá	Volné ruce při mluvení, není potřeba dodatečné lékařské péče	Velmi náročná metoda k naučení, nepřirozený hlas
<b>Elektrolarynx</b>	Nízká	Snadné k naučení	Monotonní až robotický hlas, nutné nosit externí elektrické zařízení
<b>Hrtanu podobné struktury</b>	Vysoká	Nezávislost pacienta na pravidelné lékařské péči	Velmi náročná chirurgická procedura, která pacienta vystavuje dalším možným rizikům
<b>Transplantace hrtanu</b>	Velmi vysoká	Transplantovaný hrtan přejímá funkci odstraněného orgánu	Velmi náročná chirurgická procedura, která je vhodná jen pro malé procento pacientů

Tabulka 2.1: Přehled dostupných metod rehabilitace hlasu

rehabilitace, stejně jako u jícnového hlasu, tak především rozhodují vlastnosti faryngoezofageálního segmentu. Pokud pacient není schopen si osvojit jícnový hlas, případně nemá voperován píštěl, je použit elektrolarynx. Bohužel tyto metody neřeší další problémy spojené s odstraněním hrtanu, a proto se lékaři stále snaží zdokonalovat rehabilitační metody. Za nejkomplexnější se dá považovat úplná transplantace hrtanu, která řeší víceméně všechny problémy spojené s odstraněním hrtanu. Bohužel tento zákrok je velmi náročný a vhodný pouze pro malou část pacientů. I když je tedy

v současné době lékařská věda schopna rehabilitovat hlas, tak zde zůstává otevřený prostor pro inovace a tím zlepšení kvality života lidí postižených ztrátou hrtanu.

# Kapitola 3

## Řečová rozhraní

Jedním z hlavních důsledků TL (popsané v [xx]) je ztráta hlasivek a tím i hlasu.

TBD

Problematikou komunikace pomocí mluvené řeči i v situacích, kdy akustický řečový signál není k dispozici, se zabývají systémy zpracovávající „tichou“ řeč (angl. Silent speech interface, zkr. SSI). Ve většině případů se snaží získat informaci, která je normálně zakódována v akustickém signálu získat jinou cestou.

Produkce mluvené řeči je komplexní proces, který začíná v možku a končí produkcí slyšitelného zvuku. Pokud odstraníme komponentu starající se o vznik zvuku, ještě to neznamena, že i ostatní komponenty také ztrácejí svou funkci. Tento fakt je základní premisou pro funkci všech v současnosti vyvíjených SSI systémů.

- lehce popsat technické přístupy
  - NAM
  - magnety
  - brain interface



# Kapitola 4

## Experimenty

V předchozích sekcích bylo zmíněno, že velmi významný problém představuje psychologický faktor ztráty hlasu. Ten sebou nese i na první pohled néúplně zřejmou komplikaci. A tou je obtížnost získávání řečníků po TL, kteří jsou ochotní spolupracovat na výzkumu.

Jedním z možných řešení je, že se výzkumníci sami naučí používat jednu z výše [c4l1])zmíněných metod produkce řeči. Pro účely našeho výzkumu a rámec této práce by se jednalo o elektrolarynx. I když se tato možnost jeví poměrně jednoduše, tak zdání klame. Přestože je zdravý řečník schopen obstojně mluvit (pomocí EL), za relativně krátkou dobu tak k tomu, aby produkováná řeč svými parametry odpovídala řečníkovi po TL, je potřeba relativně dlouhá doba.

V našem případě se podařilo získat pouze jednoho zkušeného řečníka<sup>1</sup>, který podstoupil TL před více než 15 lety a EL používá aktivně každý den. Je to jeho jediná možnost jak produkovat slyšitelnou řeč.

Díky spolupráci s tímto řečníkem jsme byli schopni získat cca 15 hodin promluv (více o získaných datech v 4.1 a 4.2), které byly použity pro všechny dosavadní experimenty. Z pohledu standardních obecných systémů rozpoznávání řeči se to může

---

<sup>1</sup>Jedná se o ženu v důchodovém věku, která ale stále působí na akademické půdě a jednou za čas i přednáší.

jevit jako velmi malé množství, ale jak ukáží následné experimenty, tak to není velký problém. Jelikož už před prvními experimenty se jevilo jako prozaičtější vytvářet individuální modely pro každého řečníka a výsledky experimentů toto jen podpořily.

Při rozhovorech s řečníkem se potvrdil psychický dopad ztráty hlasu na člověka po TL. Konkrétní osoba ještě dlouhá léta po operaci nebyla schopna telefonovat, natož mluvit na veřejnosti. Kvalita života se tímto velmi snížila a trvalo prý velmi dlouhou dobu, než se daná osoba odvážila i jen odpovědět na nečekaný telefonní hovor. To pro nás představovalo nezanedbatelnou porci motivace v práci.

V následujícím textu si přiblížíme a zanalyzujeme pořízená data (4.1) a na základě těchto prvotních experimentů nadefinujeme cíle a určíme další postup. V části (4.2) si popíšeme normalizaci dat a porovnáme schopnosti člověka a stroje. V sekci 4.3 rozebereme prvotní ověřovací experimenty s upravenými daty a v sekci 4.4 představíme výsledky modelů zohledňujících i délku fónému.

## 4.1 Analýza dat a první modely

Rozpoznávání řeči se věnuje nemalé usílí již od 50. let 20. století a v současné době nikoho nepřekvapí téměř bezchybně fungující obecný rozpoznávač v mobilních zařízeních. Pro obecné systémy dokonce existují korpusy s desítkami ne-li stovkami a více hodin promluv, které je možné využít při vytváření těchto systémů.

Tyto korpusy však obsahují ve většině případů pouze „standardní“<sup>2</sup> řeč. Pokud je snaha vytvořit nebo ověřit funkčnost systému za specifických podmínek (ať už se jedná o rušné prostředí či speciální typy promluv), tak je nezbytné získat potřebná data.

---

<sup>2</sup>Slovením spojením „standardní řeč“ je myšlena řeč neobsahující výrazné řečové vady, případně jiné formy produkce a často v nepříliš akusticky náročném prostředí.

### 4.1.1 Vytvoření korpusu EL promluv

Na začátku všeho je idea, v tomto případě o pomoci skupině lidí mající problémy s přirozenou řečí. Vůbec prvním předpokladem, na cestě k úspěšnému dosažení vůbec nějakého cíle, jsou data. Jelikož se jedná o velmi specifická data, tak je potřeba zajistit co možná největší množství kvalitních<sup>3</sup> a přesných dat.

V části 2.1 bylo zmíněno, že ročně se objeví více než 100 nových případů trvalé ztráty hlasu ročně. Zároveň bylo řečeno [0], že více rizikovými osobami jsou starší lidé, kteří intenzivně kouří a konzumují alkohol. Přesto je velmi patrný trend nárůstu mladších pacientů a tím i případnému nárůstu případů ztráty hlasu. Přičteme-li již zmíněný psychologický aspekt jeho ztráty, tak je zřejmé, jak komplikované je získat ke spolupráci i jen jednoho řečníka ochotného podstoupit náročné<sup>4</sup> nahrávání.

Při libovolné práci s pacienty po TL, dřív nebo později dojde k určité formě spolupráce s oddělením ORL, které má na starosti péči o tyto pacienty. V našem případě nejprve s ORL klinikou při Fakultní nemocnici v Plzni a poté i s ORL klinikou Fakultní nemocnice v Motole. S jejich pomocí jsme získali ke spolupráci jednoho řečníka. Konkrétně se jedná o dámu v duchodovém věku, která podstoupila TL před více než 15 lety. Po překonání ostychu<sup>5</sup> se byla schopna naplno vrátit do běžného života a dokonce v určité formě opět přednášet o stomatologii na Lékařské fakultě v Plzni Univerzity Karlovy.

S její pomocí jsem, v 1. etapě nahrávání, byli schopni pořídit přes 10 hodin promluv, viz tabulka 4.1. Nahrávání probíhala v relativně spartánských podmínkách v běžných prostorách katedry za plného provozu. K nahrávání byl použit profesionální mikrofon (TBD), ale mezi ostatními komponentami byla již běžná externí zvuková

<sup>3</sup>Kvalitou je myšlena věrnost dat dané doméně dále se mluví o přesnosti ve smyslu bezchybnosti přepisů.

<sup>4</sup>I pro zdravého člověka je někdy několika hodinové nahrávání vysilující. Pro jedince po TL to je z mnoha důvodů ještě řádově náročnější.

<sup>5</sup>Podle jejích vlastních slov nebyla schopna několik let po operaci ani zvednout nečekaný telefonní hovor natož mluvit na veřejnosti.

typ mik-  
rofonu

karta a soukromý notebook. Celé nahrávání bylo rozděleno do 14 samostatných sezení a probíhalo od prosince roku 2010 do května roku 2011. Každé sezení trvalo přibližně dvě hodiny během kterých jsme byli schopni získat necelou hodinu akustických dat. Samotné nahrávání se sestávalo z 10 - 20 minutového úseku pořizování nahrávky a přibližně 10 minut dlouhého odpočinku. Ten byl nezbytný z důvodu únavy řečníka.

Pro potřeby nahrávání byly vytvořeny 2 sady vět:

1. sada obsahující všechny fonémy - *40 vět.*
2. sada obsahující věty s reálnou četností fonémů - *5000 vět.*

Pořízené nahrávky vždy odpovídají 10 - 20 minutovému úseku nepřerušovaného nahrávání a soubory tak vždy obsahují několik vět, které jsou od sebe odděleny minimálně 5 sekundovým úsekem ticha. Nahrávky dále mohou obsahovat opakování chybně vyslovené věty, přerušování, kýchnutí a další neřečové události.

Přestože byly nahrávky pořízeny na základě předpřipravených sad vět, bylo nezbytné je anotovat. Ještě před samotným anotováním byly nahrávky, podle úseků s tichem, rozsekány na menší části. Jelikož řečník mohl, např. kvůli potřebě hlubšího nádechu nebo úpravě pozice EL, udělat pauzu i uprostřed věty, je výsledný počet souborů vyšší<sup>6</sup> než skutečný počet nahraných vět, viz tabulka 4.1. K anotaci posloužil interní nástroj určený k tomuto účelu a podíleli se na něm celkem 3 anotátoři z řad studentů, kteří si vzájemně kontrolovali své anotace. Ačkoli bylo potřeba anotovat relativně malé množství dat (cca 10 hodin audio záznamu), tak anotace zabrala přibližně 2 měsíce. Hlavním důvodem byla relativně dlouhá doba, po kterou se anotátoři adaptovali na specifická EL řeči. Hlavním problémem bylo to, že nejprve nebyli vůbec schopni porozumět obsahu promluvy a tím pádem jej správně přepsat.

---

<sup>6</sup>Zvýšení je způsobeno i faktem, že ne všechny věty se podařilo nahrát na první pokud. Celkově se tímto zvýšil počet souborů o 5 – 10%.

Výsledný korpus představuje 5040 unikátních vět rozdělených do 6385 souborů (viz tabulka 4.1), které v průměru obsahují 7 slov o průměrné délce 5 znaků. Tento korpus poslouží jako základ pro všechny budoucí experimenty.

Nahrávání	Délka $[HH:MM:SS]$	Počet vět	Počet souborů
2010.12 - 2011.05	11:42:42	5040	6385

Tabulka 4.1: Infoemace o korpusu nahrávek z 1. etapy nahrávání.

### 4.1.2 Analýza získaných dat

Model	WER	
	8 kHz [%]	16 kHz [%]
GMM 1024	21.37	19.20
GMM 2048	20.45	18.91
GMM 4096	20.21	18.80

Tabulka 4.2: Vliv frekvence na kvalitu modelu.

- popis jak byla data získána
- informace o datech (frekvenční rozsah, atp.)
- prvotní experimenty k určení parametrů modelu
- popsat výsledky, výsledky převážně na HTK (HMM + GMM)

## 4.2 Provnání člověk vs. stroj

- popsat důvody proč potřebujeme další nahrávky (slova, která se liší znělostí)
- popsat algoritmus výběru slov/vět k nahrávání

- napsat něco o tom, že mezi nahráváním byl velký časový rozestup a taky se změnila technika nahráváním
- problém taky s tím, že první sada nahrávek je mnohem větší (10h) než ta nová (cca 30 min) a modely tak nefungovali na nových datech v testovací sade
- použití neuronovek (Kaldi)
- popis a výsledky experimentu „člověk vs. stroj“

### 4.3 Protahování na příznacích

- problém se znělostí (systém moc nefunguje pokud je promluva krátká)
- popsat důvody pro protažení
- popis a výsledky experimentů na uměle protažených datech na příznacích (není použitelné reálně)
- aktualizace experimentu „člověk vs. stroj“

### 4.4 Modely pracující s protaženými daty

- reálné (neceločíselné) protažení na audio (použitelné reálně)
- algoritmus protažení
  1. natrénovat model na původních nahrávkách
  2. získat co možná nejpresnější zarovnání
  3. podle zarovnání protáhnout inkriminované fonémy
  4. natrénovat nový model
- vysvětlit proč je získání zarovnání kritické

- přiblížit možnosti jak vytvořit model, který bere v potaz i délku
  1. změna topologie modelu
  2. aplikace rescoringu
- výsledky model na „reálných“ protažených datech
- aktualizace experimentu „člověk vs. stroj“

# Kapitola 5

## Trenažér

Idea o trenažéru, který by mluvčího naučil protahovat. Otázkou je jestli to nedat jen jako sekci do předchozí kapitoly, případně jako součást rozsáhlejšího závěru.



## Kapitola 6

### Závěr