Graphical user interface, text, application, letter, email

Description automatically generated

Text, letter

Description automatically generated

Abstrakt

Práca analyzuje prozodické a spektrálne vlastnosti oboch smerov ATC komunikácie. Popisuje triedenie komunikácie podľa prepisu a následne Výsledkom analýz je porovnanie práve týchto vlastností medzi rečou pilotov a riadiacich veží.

Klíčová slova

Abstract

Keywords

Bibliografická citace

SIMONIDES, Jakub. *Analýza prosodických a spektrálních vlastností hlasové komunikace v řízení letového provozu* [online]. Brno, 2021 [cit. 2021-12-11]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/139062>. Semestrální práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce Zdeněk Smékal.

Prehlásenie autora o pôvodnosti diela

**Meno a priezvisko študenta:** *Jakub Simonides*

**VUT ID študenta:** *212580*

**Typ práce:** *Semestrálna práa*

**Akademický rok:** *2021/22*

**Téma záverečnej práce:** *Analýza prosodických a spektrálních vlastností hlasové komunikace v řízení letového provozu*

Prehlasujem, že svoju záverečnú prácu som vypracoval samostatne pod vedením vedúcej/ho záverečnej práce a s použitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej záverečnej práce ďalej prehlasujem, že v súvislosti s vytvorením tejto záverečnej práce som neporušil autorské práva tretích osôb, najmä som nezasiahol nedovoleným spôsobom do cudzích autorských práv osobnostných a som si plne vedomý následkov porušení ustanovení § 11 a nasledujúcich autorského zákona č. 121/2000 Sb., vrátane možných trestnoprávnych dôsledkov vyplývajúcich z ustanovení časti druhej, hlavy VI. diel 4 Trestného zákonníku č. 40/2009 Sb.

.

V Brne dňa: 13. decembra 2021

podpis autora

Obsah

[1. ATC 10](#_Toc90233050)

[1.1 História letectva 10](#_Toc90233051)

[1.2 História ATC komunikácie 10](#_Toc90233052)

[1.3 Rádiová komunikácia 12](#_Toc90233053)

[1.4 Frazeologický slovník využitý v ATC komunikácii 13](#_Toc90233054)

[1.4.1 Označovanie výšok 15](#_Toc90233055)

[1.4.2 Letová hladina (Flight level) 15](#_Toc90233056)

[1.4.3 Minimálny pokles a rozhodujúca výška 15](#_Toc90233057)

[1.4.4 Čas 15](#_Toc90233058)

[1.4.5 Výškomer 15](#_Toc90233059)

[1.4.6 Rýchlosť a smer vetra 15](#_Toc90233060)

[1.4.7 Smer letu 16](#_Toc90233061)

[1.4.8 Čísla pristávacích dráh 16](#_Toc90233062)

[1.4.9 Rádiové frekvencie 16](#_Toc90233063)

[2. Prozódia 17](#_Toc90233064)

[2.1 Reč 17](#_Toc90233065)

[2.2 Rozpoznávanie reči 18](#_Toc90233066)

[2.3 Špecifické vlastnosti reči rečníka 21](#_Toc90233067)

[2.4 Pravdepodobnosť rozpoznania 21](#_Toc90233068)

[2.5 Jazykovo špecifické aspekty reči 22](#_Toc90233069)

[3. Prevedenie 23](#_Toc90233070)

[3.1 Formátovanie dát 23](#_Toc90233071)

[3.2 Delenie podľa smeru komunikácie 23](#_Toc90233072)

[3.3 Generovanie grafov 24](#_Toc90233073)

[4. Záver 26](#_Toc90233074)

Zoznam obrázkov

[Obrázok 1.1 – Štandardná frazeológia pre čísla a písmená 14](#_Toc90233076)

[Obrázok 1.2 – Číslovanie dráh 16](#_Toc90233077)

[Obrázok 2.1 Priebežná zmena F0 pre text “The MAN in the STREET“, kde veľkými písmenami sú označené slová s dôrazom. [2] 17](#_Toc90233078)

[Obrázok 2.2 Euklidovská vzdialenosť LPCC vektorov 19](#_Toc90233079)

[Obrázok 2.3 Kontúry F0 pre (a) TIMIT, (b) NTIMIT a (c) CTIMIT rovnakého rečníka pre tú istú vetu 19](#_Toc90233080)

[Obrázok 2.4 Zmena kontúry F0 pri rozdielnom emočnom vyslovení 20](#_Toc90233081)

[Obrázok 2.5 Rozdiel v dynamike kontúr F0 pri rozprávaní čísloviek od 1 do 10 dvomi rozdielnymi ženami. 21](#_Toc90233082)

[Obrázok 3.1 Prepis pred formátovaním 23](#_Toc90233083)

[Obrázok 3.2 Prepis po formátovaní 23](#_Toc90233084)

Zoznam tabuliek

[Tabuľka 1.1 - Svetelné signály [4] 11](#_Toc90233085)

[Tabuľka 1.2 – Využité frekvencie v ATC 13](#_Toc90233086)

Úvod

Riadenie letovej prevádzky, ATC, sa zaoberá poskytovaním pokynov a informácií o letovej prevádzke tak, aby sa zabezpečila bezpečná a plynulá prevádzka lietadiel. Motiváciou práce je, aby sa dalo s istotou identifikovať, akým smerom komunikácia prebieha. Kvôli veľkému širokospektrálnemu šumu v pozadí a zložitým termínom používaných v reči, v porovnaní s bežnou komunikáciou, totiž nie je vôbec jednoduché určiť smer komunikácie. Podľa publikácie Misunderstandings in ATC Communication, častokrát až spätný pohľad na komunikáciu vyjasní, že si ju interpretovali nesprávne. [1] V reálnom čase by to znamenalo, že korekciu zle podanej informácie, je treba opraviť hneď a predísť tým prípadným nehodám ktoré by v rámci leteckej prepravy mohli mať katastrofálne následky.

Cieľom tejto práce a zároveň aj jej dôvodom výberu by v konečnom následku malo byť zjednodušenie identifikácie zdroja ATC komunikácie a tým aj jednoduchšie rozpoznanie zdroja komunikácie pre osobu tretej strany. Keďže sa letecká doprava stále rozširuje a je stále viac a viac populárnejšia aj v rámci vnútroštátnych letov, je treba nesmierne dbať na bezpečnosť pasažierov všetkými možnými bezpečnostnými prvkami. Jedným z nich je práve aj ATC komunikácia, keďže práve tá dohliada na bezpečnú nepretržitú prevádzku, nielen na letiskách. Je nevyhnutnou súčasťou celej modernej leteckej infraštruktúry a letecká doprava je od tejto komunikácie nesmierne závislá. Ak by komunikácia medzi riadiacou vežou a pilotom náhle zlyhala na celom svete, malo by to dopad na celú ľudskú spoločnosť, nielen kvôli osobnej, ale aj nákladnej preprave. Celý svet by sa jednoducho zastavil.

# ATC

## História letectva

Riadenie letovej prevádzky sa vzťahuje na „predchádzanie zrážok lietadiel a podporu usporiadaného toku letovej prevádzky“. Okrem toho sa vzdušný priestor, kde sa vykonáva riadenie letovej prevádzky, nazýva riadený vzdušný priestor (anglicky: Controlled airspace). Amplitúdová modulácia sa celosvetovo používa ako metóda modulácie rádiových vĺn pre rádiotelefóny používané na riadenie letovej prevádzky. Spočiatku, pri skorých začiatkoch letectva, nebola žiadna, verbálna ani neverbálna, komunikácia medzi pilotom a letiskom. Spočiatku preto, lebo letiská ani neexistovali. Každý človek si mohol skonštruovať nejaký letecký stroj. Tieto prvé lietadlá teda často havarovali, najmä kvôli ich technickej nedokonalosti. Časom sa ale aj napriek prvotným obavám, začalo brať v úvahu užitočnosti leteckej dopravy. V roku 1911 bola letecky doručená prvá pošta. Pri zapojení Spojených Štátov Amerických do Prvej Svetovej Vojny, sa letectvo ukázalo byť užitočné aj za účelmi skúmania a prepravy zbraní. [4]

## História ATC komunikácie

Pred skorými 30. rokmi minulého storočia bol jediný bezpečnostný prvok let cez deň a za priaznivých podmienok. Postupne sa ale bezpečnosť vyvíjala. Najskorší spôsob ATC komunikácie, bol človek, ktorý stál na viditeľnej pozícii letiska a mával farebnými vlajkami na letisku St. Louis, Missouri v 1929. Kvôli viacerým problémom, ako napríklad lety so zlou viditeľnosťou, sa postupne vlajky prestali používať a nahradili ich svetelné pištole, ktoré vedeli zasvietiť lúč svetla priamo na konkrétne lietadlo. Táto metóda sa využíva v takmer všetkých riadiacich vežiach dodnes, ak náhodou nie je možné využiť rádiovú komunikáciu. Svetelné pokyny sa líšili podľa toho, či bolo lietadlo práve vo vzdušnom priestore alebo na letisku. [4]

|  |  |  |
| --- | --- | --- |
| **Farba a typ signálu** | **Význam z pohľadu pilota na zemi** | **Význam z pohľadu pilota vo vzduchu** |
| Stála zelená | Povolenie k odletu | Povolenie k pristátiu |
| Blikajúca zelená | Povolenie k pojazdu | Vrátenie sa k pristávaniu |
| Stála červená | Stop | Pokračovať v krúžení a dať prednosť inému lietadlu |
| Blikajúca červená | Pojazd mimo práve požitú pristávaciu dráhu | Letisko je nebezpečné na pristátie - nepristávať |
| Blikajúca biela | Návrat na štartovací bod z letiska | - |
| Meniaca sa červená a zelená | Dbajte na mimoriadnu opatrnosť | Dbajte na mimoriadnu opatrnosť |

Tabuľka 1.1 - Svetelné signály, prevzaté z [4]

Od 30. rokov sa značne tieto signály nezmenili. Keďže dopyt po leteckej preprave bol veľký, tak ani tento spôsob už nebol dostatočný na pomaly sa zhusťujúcu leteckú prepravu. Kvôli tomu bolo treba vyvinúť nejaký obojsmerný typ komunikácie, ktorým bola rádiová komunikácia. Počiatky rádiovej ATC komunikácie sú z letiska v Clevelande, Ohio. Práve tu bola postavená veža na starom hangári a boli v nej používané 15W vysielačky s dosahom až 15 míľ pri rôznych poveternostných podmienkach. Problém však bol s ich počiatočnou veľkosťou a hmotnosťou alebo, pri malých lietadlách, s nedostatkom energie na ich napájanie. Keďže veža bola najvyšším bodom letiska, pracovníci veže mali dokonalý prehľad o celom priebehu dopravy na letisku. Postupne sa, po vzore toho letiska, transformovali aj iné rušné letiská k tomuto typu udržiavania kontroly nad plynulosťou letovej premávky. Problémom bola stále hmotnosť a veľkosť celého systému rádiovej komunikácie v porovnaní s tým o koľko ziskov by sa letecké spoločnosti pripravili, keby tento priestor nevyužijú na miesto pre pasažierov. Rovnako ešte neboli zriadené žiadne pravidlá v rámci rádiovej komunikácie, čo niekedy vedelo byť viac na príťaž ako na úžitok. Niektorí piloti kontaktovali ATC až pri pristávaní, pričom niektorí už keď boli 5 až 10 míľ ďaleko. V konečnom dôsledku sa ale rádiová ATC komunikácia potvrdila byť dobrou cestou pre budúcnosť leteckej prepravy.

Ako sa rádiové veže stávali bežnou súčasťou letísk, leteckí inžinieri už boli možní vyrábať nové lietadlá, ktoré boli veľkokapacitné, rýchlejšie ale hlavne, ktoré boli osadené rádiovými systémami pre umožnenie komunikácie. Umožnil sa tým let v podmienkach, kedy by sa 10 rokov dozadu nemohol uskutočniť. Piloti boli možní pristávať pri viditeľnosti menej ako 2 míle. Ako sa ale letectvo stávalo rozšírenejším, ľudia začínali mať strach pre prípadné zrážky lietadiel alebo poškodenie majetku spôsobené lietadlom a tak žiadali, o zákaz letovej dopravy ponad isté územia. Toto by znamenalo veľmi rýchly koniec komerčnej leteckej dopravy. V 1934, pre tieto opodstatnené obavy, Kongres schválil Predsedníctvo Obchodného Letectva, ktoré zodpovedalo za reguláciu dopravy národných liniek. Tieto všetky pokroky v letectve sa ukázali ako prínosné a postupne sa schválil tzv. The Air Commerce Act, ktorý mal dopomôcť rýchlo sa rozširujúcej leteckej doprave prostredníctvom letísk a rôznych navigačných prvkov. Postupne sa ku komunikácii pridávali aj ďalšie technológie pre leteckú bezpečnosť, napríklad radar. [4]

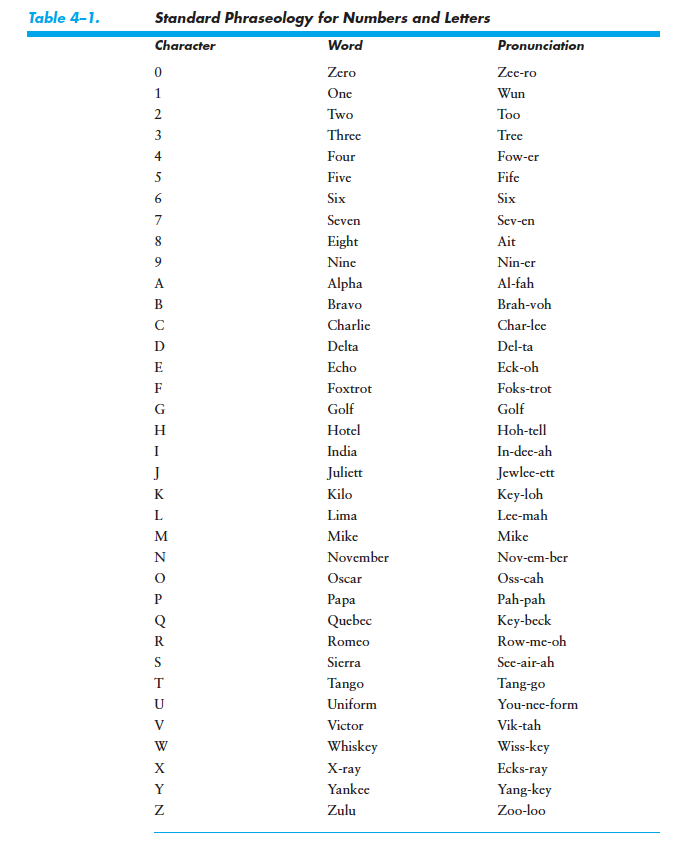
## Rádiová komunikácia

Aj keď sa rádiová technika za ten čas rôzne menila, od počiatkov z Clevelandu zostala rádiová komunikácia hlavným spôsobom ATC komunikácie. Prvým typom komunikácie, bola jednostranná, kedy sa z veže dal kontaktovať pilot no nie naopak, keďže spočiatku bol problém s veľkosťou inštalácie aj vysielača aj príjmača. Riešením v rámci bezpečnosti ale muselo byť použitie príjmacej technológie vo veži a vysielacej v lietadle. Aby sa eliminovala interferencia, vysielače v lietadlách používali inú frekvenciu ako tie, ktoré boli vo vežiach. Tento systém sa nazýva ako systém duplex komunikácií. Systém má isté nevýhody, avšak to bol prvý systém, ktorý umožňoval obojsmernú komunikáciu pilota s ATC. Časom sa z neho vyvinul systém simplex ktorý pracuje na jednej frekvencii obojsmerne. Keďže letecká doprava nepozostáva len z komerčných letov, bolo treba prideliť frekvencie k reálnym využitiam. Väčšina rádiových komunikácii, momentálne využívaných v letectve, vie komunikovať takmer s každým komunikačným kanálom, ktorý je k dispozícii, či už to je s lietadlom, alebo v rámci veží. [4]

|  |  |
| --- | --- |
| **Frekvencie** | **Použitie** |
| 108.000–117.950 | Navigačné pomôcky |
| 118.000–121.400 | ATC |
| 121.500 | Núdzové hľadanie a záchrana |
| 121.600–121.925 | Služby letiska a ELT test |
| 121.950 | Letecká inštruktáž a podpora |
| 121.975 | FSS poradenstvo pre súkromné lety |
| 122.000–122.050 | Pomocná služba pre prebiehajúce lety |
| 122.075–122.675 | FSS poradenstvo pre súkromné lety |
| 122.700–122.725 | UNICOM |
| 122.750 | Medzilietadlová vo vzduchu |
| 122.775 | Letecká inštruktáž a podpora |
| 122.800 | UNICOM |
| 122.825 | Vnútroštátne VHF |
| 122.850 | Multicom |
| 122.875 | Vnútroštátne VHF |
| 122.900 | Multicom |
| 122.925 | Multicom |
| 122.950 | Unicom |
| 122.975–123.000 | Unicom |
| 123.050–123.075 | Unicom |
| 123.100 | Letecké hľadanie a pomoc |
| 123.125–123.275 | Letecké skúšobné stanice |
| 123.300 | Letecká podpora |
| 123.325–123.475 | Letecké skúšobné stanice |
| 123.500 | Letecká podpora |
| 123.525–123.575 | Letecké skúšobné stanice |
| 123.600–123.650 | FSS poradenstvo leteckého dopravcu |
| 123.675–128.800 | ATC |
| 126.200 | vojenské ATC |
| 128.825–132.000 | Vnútroštátne VHF |
| 132.025–136.975 | ATC |

Tabuľka 1.2 – Využité frekvencie v ATC, prevzaté z [4]

## Frazeologický slovník využitý v ATC komunikácii

Aby sa predišlo zle pochopeným správam, je na to prispôsobený aj samotný slovník a forma skladby informácie pri ATC komunikácii. Štandardne sa používa formát kde ako prvá časť správy je identifikácia cieľového príjmateľa správy. To slúži na upozornenie príjmateľa na prichádzajúcu správu. Druhou informáciou v poradí je identifikácia odosielateľa správy, pre určenie toho, kto je zdroj. Tretím v poradí je obsah správy ktorý je bežne vo formáte štandardom FAA. Nakoniec je uzatvorenie komunikácie, ktoré prebieha spôsobom, že strana ktorá uzatvára komunikáciu sa identifikuje, pre jednoduchosť spätnej identifikácie. Viacero čísel aj písmen môže znieť podobne a vzhľadom aj na viaceré prízvuky pri rôznych jazykoch bolo potrebné vytvoriť určitý štandard ktorý by sa využíval v ATC komunikácii, aby sa predošlo možným chybám. Práve kvôli tomuto, bol vytvorený frazeologický slovník špeciálne pre ATC. Pre čísla sa využíva hovorenie po číslovkách. Číslo 10 by sa teda povedalo ako one zero ak sa nenachádza v skupinovej forme. Vyslovenie skupinovej formy byt bolo napríklad pre číslo 132 ako one thirty-two a samostatne one three two. [4]

Obrázok 1.1 – Štandardná frazeológia pre čísla a písmená, prevzaté z [4]

### Označovanie výšok

Označenie výšok sa v ATC bežne udáva v nadmorskej výške (MSL), s výnimkou oblakov, kedy sa využíva nadzemská výška. Pri nadzemskej výške sa na to však dopredu upozorňuje. Pri vyslovovaní sa rozdeľujú Tisícky a stovky zo skupín. Napríklad výška 12 500 nad morom by sa vyslovila ako one two thousand five hundred. [4]

### Letová hladina (Flight level)

Letovej hladine vždy predchádzajú slová Flight level a každá číslovka by sa mala vyslovovať samostatne. Pri letovej hladine 180 by to teda bolo Flight level one eight zero. [4]

### Minimálny pokles a rozhodujúca výška

Vždy sa udávajú podľa tabuliek momentálneho stavu. Pri vyslovovaní MDA 1950 sa znovu číslovky uvádzajú samostatne ako minimum descent altitude one niner five zero. [4]

### Čas

Aby sa predišlo nedorozumeniam, všade na svete v ATC sa používa štandardný čas UTC, ktorý je rovnaký ako Greenwich mean time (GMT). Odstraňuje sa ním nutnosť vyjadrovania A.M. a P.M., kvôli využívaniu 24 hodinového formátu. Frazeologicky sa vždy vyslovuje ako len štvorčíslie a následne slovom zulu, kvôli ujasneniu, že sa jedná o čas. Čas 7:20 AM by sa teda zmenil na 0720 ako time zero seven two zero zulu a 7:20 PM by bol 1920. one niner two zero zulu. [4]

### Výškomer

Každý pilot musí mať dáta o aktuálnom barometrickom tlaku aby výškomer mohol byť správne nakalibrovaný. Pri vyslovovaní sa vyslovujú číslovky samostatne a desatinná čiarka sa vynecháva. Predchádza tomu slovo altimeter. Pri hodnote výškomera 29.92 by sa teda vyslovilo altimeter two niner niner two. Piloti, ktorí majú výškomer v metrických jednotkách musia svojvoľne konvertovať hodnotu, alebo požiadať o konverziu jednotiek. [4]

### Rýchlosť a smer vetra

Smer vetra na letisku sa vždy určuje vo vzťahu k magnetickému severu a opisuje smer odkiaľ vietor fúka. Smer sa vždy zaokrúhľuje k najbližším desiatim stupňom. Vietor ktorý prúdi zo severu na juh bude teda 360° vietor a vietor z východu je 90° vietor. Medzinárodný štandard na určenie rýchlosti vetra sa udáva v uzloch. Frazeologicky sa vyslovuje započatím so slovom wind, nasleduje údaj smeru vetra, potom prichádza slovo “at“ a za ním nasleduje rýchlosť vetra. Ak sa smer vetra mení tak prichádza slovo variable, ak prístroje na meranie vetra nefungujú, používa sa slovo estimated a ak sa mení rýchlosť tak sa pred rýchlosť pridáva slovo gust. Pri premenlivom vetre z juhovýchodu pri rýchlosti 12 uzlov miestami 35 uzlov by sa vyslovilo wind one five zero variable at one two gusts to three five. [4]

### Smer letu

Smer letu sa rovnako uvádza vo vzťahu k magnetickému serveru. Ak obsahuje menej ako tri číslice, pridávajú sa na začiatok nuly, aby sa z toho stalo trojciferné číslo. Predchádza im slovo heading. Pri smere letu 225°je výslovnosť heading two five five. [4]

### Čísla pristávacích dráh

Pristávacie dráhy sú očíslované tiež podľa magnetického severu. Pristávacia dráha smerom na sever by teda bola označená 36 z vynechaním poslednej nuly. Vyslovovaniu vždy predchádza slovo runway, nasleduje číslo po číslovkách a ak ich je viac na rovnaký smer, za číslom nasleduje right center alebo left. Ak teda je dráha smerom 138°a jej číslo je 14R, tak výslovnosť je Runway one four right. [4]Diagram

Description automatically generated

Obrázok 1.2 – Číslovanie dráh, prevzaté z [4]

### Rádiové frekvencie

Pri frekvenciách sa číslovky znovu udávajú samostatne a frekvencie sú od seba vždy vzdialené o 25kHz. Prvé desatinné číslo sa vždy vyslovuje, aj keby bolo nulové. Tretie desatinné sa nevyslovuje nikdy, keďže je vždy 0 alebo 5 a dá sa predpokladať. Frekvencia sa explicitne vyslovuje len ak je v kHz. Pri frekvencii 131,725 mHz by sa vyslovovalo one three one point seven two. [4]

# Prozódia

## Reč

Reč je skladbou viacerých foném, ktoré sa následne spájajú do slabík a vytvoria slovo. Reč, ako ľudstvu najprirodzenejší komunikačný prostriedok, si teda rýchlo zabrala prvenstvo v ATC komunikácii. Zároveň je aj najefektívnejšia. Priemerne dokáže človek povedať 120 až 250 slov za minútu a tak niet pochýb, že v rámci hustej leteckej dopravy, je reč tým najlepším komunikačným kanálom. Prozódia sa vzťahuje na všetky fonetické vlastnosti, ktoré sa vyskytujú v reči. Zahŕňa intonáciu alebo výšku tónu, prízvuk, dôraz, relatívnu dĺžku tónu, rytmus, energiu, pauzy, kvalitu hlasu a intonáciu. Vzťahuje sa na tie prvky, ktoré sa môžu líšiť v závislosti od kontextu a dôrazu. Sú jazyky, kde je dôraz v určitom slove permanentne určený, no keďže v ATC komunikácii sa používa prevažne angličtina, musíme brať ohľad na to, že v angličtine má viacero rovnakých slov iný význam práve podľa toho, na ktorú slabiku dáme dôraz. Preto je aj v anglickej ATC komunikácii nesmierne dôležité, ktorým slovám dáme dôraz na aké miesto.

Intonácia je dôležitým aspektom prozódie. Intonáciu vieme vyjadriť ako zmenu výšky hlasu v čase. Z toho sa dá v konečnom dôsledku odvodiť aj základná frekvencia tónu hlasu F0 v čase. Pre mužov je to zväčša medzi 80 až 200 Hz, pre ženy 180 až 400 Hz. Medzi týmito hodnotami F0 buď stúpa alebo klesá. [2]

A picture containing chart

Description automatically generated

Obrázok 2.1 Priebežná zmena F0 pre text “The MAN in the STREET“, kde veľkými písmenami sú označené slová s dôrazom. [2]

Ďalšou prozodickou vlastnosťou je rytmická zmena. Rytmickou zmenou sa rozumejú rôzne dĺžky zvukov v slove. Slabika v slove je základnou jednotkou časovania v reči. Čo sa týka analýzy všetkých týchto vlastností reči, sú jedinečné u každého jedinca a tak sa stávajú akýmsi unikátnym identifikátorom človeka. Práca sa zameriava na konkrétnu analýzu týchto vlastností pochádzajúcich z viacerých zdrojov a ich následné porovnanie, za účelom analýzy podobných a rozličných aspektov komunikácie. [2]

## Rozpoznávanie reči

Prozódia je taktiež podstatným faktorom v rámci rozpoznávania reči. Rozpoznávanie je hrubo založené na spektrálnych/kepstrálnych vlastnostiach získaných pomocou spektrálnej analýzy. Hlavným faktorom k dosiahnutiu

očakávaných výsledkov je získanie hlasového záznamu v dobrých akustických podmienkach. Pri zlej akustike sa totiž v spektrálnej analýze reči vyskytuje oj chyba, ktorá vznikne zlými akustickými podmienkami, čo znamená že v analýze nebude len čistá reč ale aj napríklad ruch, ktorý vzniká akustikou miestnosti. Samotné spektrum hlasu hovoriacej osoby je teda značne rozdielne v akusticky ideálnej miestnosti, vo vonkajšom prostredí a taktiež v kokpite pilota alebo v samotnej riadiacej veži. Rovnako tento problém môžeme pozorovať pri ATC komunikácii, kde nám spektrum ovplyvňuje šum prenosového kanálu. Pri analýze ATC komunikácie teda musíme brať v úvahu všetky faktory ktoré ovplyvňujú spektrum a vo výsledkoch ich aj zohľadniť. Pôsobenie týchto faktorov si je možné všimnúť na ukážke nižšie pri vyslovení frázy “Don’t carry an oily rig like that“. Kanály pomenované TIMIT, NTIMIT a CTIMIT sú ukážkou variácie F0 pri nahrávaní blízko pri mikrofóne, cez kanál s vysokým obsahom šumu a v prostredí mobilnej siete. Pri charakteristikách kontúr F0 sa dá potvrdiť, že robustnosť prozodických vlastností je natoľko silná, že zmeny kontúr F0 sú minimálne, no stále postrehnuteľné. [2]

Chart

Description automatically generated

Obrázok 2.2 Euklidovská vzdialenosť LPCC vektorov

Letter

Description automatically generated with low confidence

Obrázok 2.3 Kontúry F0 pre (a) TIMIT, (b) NTIMIT a (c) CTIMIT rovnakého rečníka pre tú istú vetu

Pri reči netreba zabúdať ani na emóciu. Rozdielne kontúry F0 môžeme získať z toho istého zdroja len pri emočnom rozdiele v reči. Takisto emočný faktor vplýva aj na rozdiely tempa a páuz. [2]

**Text, letter

Description automatically generated**

Obrázok 2.4 Zmena kontúry F0 pri rozdielnom emočnom vyslovení

## Špecifické vlastnosti reči rečníka

Každý rečník má svoje unikátne charakteristické črty, buď podľa rečníkových fyziologických charakteristík hlasových orgánov alebo nejakých zvykov. Rozdiely môžu byť zhrnuté nasledovne:

1. Veľkosť a tvar vokálneho traktu
2. Excitačné charakteristiky
3. Prozódia
4. Idiolekt
5. Sémantika

Prvé tri z vyššie uvedených čŕt vieme vyzistiť zo signálu a zvyšné dve sú prítomné v samotnom obsahu textu. [2]A picture containing graphical user interface

Description automatically generated

Obrázok 2.5 Rozdiel v dynamike kontúr F0 pri rozprávaní čísloviek od 1 do 10 dvomi rozdielnymi ženami.

## Pravdepodobnosť rozpoznania

Rozpoznanie sa dá nazvať ako priradenie najpravdepodobnejšieho rozprávača, jazyka, emócie alebo nejakej zvukovej jednotky C\* vstupnej reči z množiny existujúcich rozprávačov, jazykov, emócii ale zvukových jednotiek. Ak je určené že označuje množinu existujúcich rozprávačov, jazykov, emócii ale zvukových jednotiek a O označuje pozorovania zistené zo vstupnej reči, tak rovnica rozpoznávanie v rámci pravdepodobnosti bude znieť nasledovne:

|  |  |
| --- | --- |
| , | (2.1) |

kde je aposteriórna pravdepodobnosť členu , pre daný výskyt reči vyjadrený pre . [2]

## Jazykovo špecifické aspekty reči

Je viacero rečových aspektov, ktoré sú rozdielne pri rozličných jazykoch ako napríklad akustická fonetika, prozódia, fonotaktika, slovník používaných slov a slovná skladba. Keďže ATC komunikácia pri lokálnych letoch neprebieha len v anglickom jazyku, tieto aspekty treba brať v úvahu aj pri analýze v rozličných jazykoch kde sa napríklad intonácia, rytmus alebo dôraz pri vyslovovaní rôznych slov mení v závislosti od hovoreného jazyka. [2]

# Prevedenie

Cieľom semestrálnej práce, bol zber dát na samotnú analýzu z nahrávok nepretržitej ATC komunikácie medzi pilotmi a vežami z viacerých letísk. Zvukové súbory, spolu s prepismi komunikácie, boli poskytnuté vedúcim práce. Zvukové súbory boli vo formáte .sph, čo je audio súbor, vytvorený vo formáte NIST SPHERE. Daný formát sa využíva práve v hlasovom rozpoznávaní a vlny v tomto formáte, majú bežne 16-bitovú pulzne kódovú moduláciu a vzorkovaciu frekvenciu o výške 16 KHz. Formát .sph však python neprečíta a tak bolo nutné danú nahrávku prekonvertovať do formátu, s ktorým python problém mať nebude. To bolo docielené za použitia programu SoX, ktorý audio súbor prekonvertoval na Waveform Audio File Format.

## Formátovanie dát

Analýze samotných dát predchádzalo rozčlenenie nepretržitej, cca. dvojhodinovej komunikácie na segmenty. Jeden segment komunikácie sa rozumie ako jeden súbor s príponou .wav kde je extrahovaná práve jedna pasáž komunikačného kanála. Rozdelenie bolo dosiahnuté pomocou skriptu v programovacom jazyku python, kde v prvom kroku, je modifikovaný textový dokument s prepisom komunikácie, za účelom dosiahnutia formátu, ktorý obsahuje jedine údaj smeru komunikácie a počiatočný a konečný čas segmentu.

Text

Description automatically generated with medium confidence

Obrázok 3.1 Prepis pred formátovaním



Obrázok 3.2 Prepis po formátovaní

## Delenie podľa smeru komunikácie

V druhom kroku, za použitia knižnice pydub, sa z modifikovaného textového dokumentu načítajú dáta a pomocou cyklu “for“, program vyhodnotí, že ak riadok začína na PA (pilot veži) resp. AP (veža pilotovi), tak čítanie prejde do ďalšieho riadku. Ďalší riadok sa rozdelí na dva stringy, ktoré sa uložia do premenných var1 a var2. Stringy je potrebné konvertovať na dátový typ float, keďže sa jedná o desatinné čísla a následne ich vynásobiť tisícimi. Násobenie je nutné vykonať, kvôli prevodu zo sekúnd na milisekundy. Nakoniec sa audio súbor segmentuje podľa časových údajov a následne je uložený do príslušného priečinku podľa smeru komunikácie, rovnako do súboru s príponou .wav. Súbory sú uložené číselne vzostupne, pre obe strany komunikácie.

from pydub import AudioSegment

x=0

with open('stampsfromatted.txt') as f:

for line in f:

if line.startswith('PA1'):

linez = next(f)

x=x+1

var1, var2 = linez.split()

var1 = float(var1)

var2 = float(var2)

print(var1, var2)

var1 = var1 \* 1000

var2 = var2 \* 1000

newAudio = AudioSegment.from\_wav("bos.wav")

newAudio = newAudio[var1:var2]

output = "./rozsekane/ATCtoPilot/{}.wav".format(x)

newAudio.export(output, format="wav")

x=0

with open('stampsfromatted.txt') as f:

for line in f:

if line.startswith('AP1'):

linez = next(f)

x=x+1

var1, var2 = linez.split()

var1 = float(var1)

var2 = float(var2)

print(var1, var2)

var1 = var1 \* 1000

var2 = var2 \* 1000

newAudio = AudioSegment.from\_wav("bos.wav")

newAudio = newAudio[var1:var2]

output = "./rozsekane/PilottoATC/{}.wav".format(x)

newAudio.export(output, format="wav")

## Generovanie grafov

Po rozdelení segmentov do patričných priečinkov, bolo získaných vyše 1000 segmentov z jednej nepretržitej komunikácie. 1000 segmentov, je dostatočná vzorka na zistenie bližších detailov. Tie boli zistené znovu pomocou pythonu v kombinácii saplikáciou Praat. Praat je aplikácia, ktorá je určená na analýzu hovoreného slova a dokáže zo vzorky audio súboru analyzovať reč a následne vygeneruje konkrétne dáta, ktoré sa viažu s danou vzorkou reči. Pre analýzu ATC komunikácie sú obzvlášť potrebné dáta o intenzite a základnej frekvencii hlasu F0, ktoré sa aj pomocou pythonu s Praatom podarilo extrahovať. Analýza samotných vzoriek bola obsiahnutá pomocou Praat knižnice pre python, parselmouth, tak, že cez cyklus for, boli postupne analyzované všetky segmenty komunikácie v obidvoch smeroch. Každý segment bol individuálne postupne načítaný a preskúmaný. Ku každému segmentu, boli vygenerované grafy analýz pomocou knižnice matplotlib. Grafy boli vygenerované pre amplitúdu hlasu v čase, intenzitu hlasu v čase a základnú frekvenciu F0 v čase pre každý konkrétny segment komunikácie. Pre porovnanie a určenie podobností a rozličností dát, boli taktiež vytvorené boxploty, pre porovnanie všetkých údajov jednej strany komunikácie. Grafy boli uložené vo formáte .png a pomenované podľa príslušných segmentov komunikácie.

Pri porovnaní boxplotov v prílohe A, si je možné všimnúť že medián hodnôt pri intenzite je pri komunikácii veža -> pilot: 67,8 a pri pilot -> veža je to 70,3

Čo sa týka kontúr F0, tak je rozdiel značne väčší veža -> pilot: 92,7 a pri pilot -> veža je to 113. Z týchto grafov, ktorých na porovnanie je vyše tisíc, sa dá potvrdiť že kontúry F0,bývajú z pravidla vyššie pri smere od pilota k veži a to značne. Pre vysoký počet grafov, je v prílohe zobrazené iba porovnanie prvého segmentu každej strany komunikácie. Zvyšné grafy sa nachádzajú v pripnutom .zip súbore.

Po ich porovnaní, je možné potvrdiť, čo majú navzájom spoločné a odlišné a na základe toho, sa dostať ku konkrétnym dátam, na základe ktorých sa ďalej v bakalárskej práci, bude môcť vytvoriť program, ktorý bude vyhodnocovať smer komunikácie nie podľa prepisu, ale práve podľa prozodických a spektrálnych vlastností.

# Záver

Ako prvý krok bolo obsiahnutie dát s ktorými s následne bude pracovať. K dispozícii som mal viacero nahrávok nepretržitej komunikácie z rôznych letísk a aj ich prepis do textového dokumentu.

Aby mohlo dôjsť k automatickému rozpoznávaniu, bolo najprv nutné rozdeliť danú nahrávku na samostatné súbory a takisto ich uložiť podľa smeru komunikácie spolu s vzostupným číslovaním pre prehľadné dohľadanie podľa prepisu. Následne už rozdelené nahrávky boli analyzované pomocou skriptu v programovacom jazyku python a takisto boli extrahované dáta o amplitúde v čase, intenzite hlasu v čase a v neposlednom rade aj údaje o základnej frekvencii F0 v čase danej nahrávky. Z grafov boli vytvorené boxploty, ktoré boli následne porovnané, pre získanie obrazu o spektrálnych a prozodických vlastnostiach oboch strán komunikácie. Po ich získaní boli dáta porovnané podľa toho, čo majú spoločné a čo odlišné v rámci rozdelenia podľa smeru komunikácie a na základe toho, bolo možné rozlíšiť, aké hodnoty údajov sprevádzajú komunikáciu v určitom smere. V bakalárskej práci ďalej na základe týchto zozbieraných dát bude identifikácia smeru komunikácie, bez prepisu, len na základe jej prozodických a spektrálnych vlastností.

Literatúra

1. BARSHI, Immanuel a Candace FARRIS. Misunderstandings in ATC communication : language, cognition, and experimental methodology. 2016. Abingdon: Routledge, 2016. ISBN 978-0-7546-9933-0.
2. MARY, Leena. Extraction of Prosody for Automatic Speaker, Language, Emotion and Speech Recognition. 2nd ed. Cham: Springer, 2019. ISBN 978-3-319-91171-7.
3. HOFBAUER, Konrad, Stefan PETRIK a Horst HERING. The ATCOSIM Corpus of Non-Prompted Clean Air Traffic Control Speech.
4. NOLAN, Michael S. *Fundamentals of Air Traffic Control*. 5th ed. Cliffton Park (New York): Cengage, c2011. ISBN 978-1-4354-8272-2.

Zoznam symbolov a skratiek

Zkratky:

LPCC Linear Prediction Cepstrum Coefficient

WAV Waveform Audio File Format

Symboly:

*U* napětí (V)

*I* proud (A)

Pozn.: V této části by měly být uvedeny všechny zkratky použité v textu a všechny symboly použité v rovnicích.

Zoznam príloh

[Příloha A - Naměřené hodnoty 23](#_Toc56778062)

1. Namerané hodnoty
   1. Namerané hodnoty – boxplotyChart, box and whisker chart

      Description automatically generatedChart

      Description automatically generated with low confidenceChart, box and whisker chart

      Description automatically generatedChart, box and whisker chart

      Description automatically generated