



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta dopravní
Ústav letecké dopravy

*DETEKCE NESPRÁVNÉHO ZPĚTNÉHO POTVRZENÍ V ŘÍZENÍ LETOVÉHO
PROVOZU*

Incorrect Read-back Detection in Air Traffic Control

Bakalářská práce

Studijní program: Bakalářský

Studijní obor: Letecká doprava

Vedoucí práce: Ing. Terézia Pilmannová, MBA.

Adam Kalanin

Praha 2023

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta dopravní

děkan

Konviktská 20, 110 00 Praha 1



K621.....Ústav letecké dopravy

ZADÁNÍ BAKALÁŘSKÉ PRÁCE (PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení studenta (včetně titulů):

Adam Kalanin

Studijní program (obor/specializace) studenta:

bakalářský – LED – Letecká doprava

Název tématu (česky): **Detekce nesprávného zpětného potvrzení v řízení letového provozu**

Název tématu (anglicky): Incorrect Read-back Detection in Air Traffic Control

Zásady pro vypracování

Při zpracování bakalářské práce se řiďte následujícími pokyny:

- Cílem práce je navrhnout softwarový nástroj pro detekci nesprávného zpětného potvrzení (read-back) v rámci komunikace mezi řídicím letového provozu a posádkou a tento nástroj v simulacích otestovat.
- Analyzujte situace, které vedou k nesprávnému zpětnému potvrzení.
- Navrhněte nástroj pro detekci nesprávného potvrzení.
- Navržený nástroj ověřte v rámci simulace.
- Zhodnoťte výstupy experimentu s ohledem na četnost správné detekce chyby.
- Výsledky v kontextu dalšího rozvoje nástroje pro detekci nesprávného zpětného potvrzení diskutujte.



- Rozsah grafických prací: dle pokynů vedoucího závěrečné práce
- Rozsah průvodní zprávy: minimálně 35 stran textu (včetně obrázků, grafů a tabulek, které jsou součástí průvodní zprávy)
- Seznam odborné literatury: ICAO, Doc 4444 - Air Traffic Management (5th ed.), International Civil Aviation Organization, 2001
Jean-Paul Imbert, H.M Hodgetts, Robert Parise, François Vachon, Frédéric Dehais, et al.. Attentional costs and failures in air traffic control notifications. Ergonomics, Taylor & Francis, 2014,

Vedoucí bakalářské práce: **Ing. Terézia Pilmannová, MBA**

Datum zadání bakalářské práce: **8. října 2021**
(datum prvního zadání této práce, které musí být nejpozději 10 měsíců před datem prvního předpokládaného odevzdání této práce vyplývajícího ze standardní doby studia)

Datum odevzdání bakalářské práce: **30. listopadu 2023**
a) datum prvního předpokládaného odevzdání práce vyplývající ze standardní doby studia a z doporučeného časového plánu studia
b) v případě odkladu odevzdání práce následující datum odevzdání práce vyplývající z doporučeného časového plánu studia

doc. Ing. Jakub Kraus, Ph.D.
vedoucí
Ústavu letecké dopravy



prof. Ing. Ondřej Příbyl, Ph.D.
děkan fakulty

Potvrzuji převzetí zadání bakalářské práce.

Adam Kalanin
jméno a podpis studenta

V Praze dne..... 8. srpna 2023



Abstrakt

Zpětné potvrzení je postup, při kterém přijímací stanice (posádka letounu) opakuje přijatou zprávu nebo její příslušnou část zpět vysílací stanici (řídícímu letového provozu), aby získala potvrzení o správném příjmu. Neopravené chybné zpětné potvrzení (známé jako incorrect read-back) může vést k odchylce od zamýšleného povolení řídicího a nemusí být zjištěno, dokud řídicí nezaznamená odchylku na svém situačním zobrazení.

Cílem práce je navrhnout softwarový nástroj, který by za použití standardní frazeologie chybu odhalil a řídicího letového provozu upozornil. Pro účely vyhodnocení úspěšnosti navrženého systému pro automatické rozpoznání řeči (ASR) byl navržen experiment, jenž obsahoval dvacet vybraných letových povolení a jejich předem nahraných zpětných potvrzení. Osm z dvaceti připravených zpětných potvrzení obsahovalo nesprávnou informaci. Správnost zpětného potvrzení byla vyhodnocována jak systémem ASR, tak testovanými subjekty.

Výsledky získané z experimentu ukázaly, že první iterace systému ASR nijak upravená do leteckého prostředí dosahuje úspěšnosti vyhodnocení správnosti zpětného potvrzení 69,9 %, přičemž hlavně dlouhé a komplexní povolení se jevila jako problémová. Oproti tomu, detekce chybného potvrzení u subjektů činila 95,6 %. Zároveň byly zkoumány vybrané demografické znaky subjektů a jejich vliv na konečnou úspěšnost systému ASR, jenž se na úspěšnosti systému nepodepsaly.

Klíčová slova: letové povolení, rozpoznávání řeči, řídicí letového provozu, zpětné potvrzení



Abstract

Read-back is a procedure, in which a receiving station (flight crew) repeats the received message, or a specific part of it, back to the broadcasting station (Air Traffic Controller), in order to ensure that the message was received in the correct format. Incorrect read-back may lead to deviation from the intended clearance issued by the Air Traffic Controller (ATC) and may not be discovered until the deviation is visible on the radar screen of the Air Traffic Controller.

The goal of this thesis is to design a software tool, which would identify this error in standard phraseology of the read-back and will notify the Air Traffic Controller of such situation. An experiment was proposed for the purpose of evaluating the success rate of the designed automatic speech recognition (ASR) system. This experiment contained twenty selected phrases of ATC clearances and their pre-recorded read-backs. Eight of the twenty prepared read-backs intentionally contained incorrect information. The read-back accuracy was evaluated by both the ASR system, and the tested subjects.

Results obtained from the experiment showed that the first iteration of the ASR system, which has not been in any way modified to the specifics of aviation English, reached the success rate of read-back evaluation of 69,9 %. The lowest success rate was recorded for long and complex route clearances. On the other hand, the incorrect read-back detection by subjects reached 95,6 %. At the same time, the influence of demographic traits of subjects proved to be of little influence in this experiment.

Keywords: air traffic controller, clearance, read-back, speech recognition



Poděkování

Rád bych touto cestou poděkoval všem lidem, kteří byli ochotni se zapojit do experimentu související s touto prací. Velký dík pak patří Ing. Terézii Pilmannové, MBA. za cenné rady a zejména za její ochotu, bez níž bych se k dopsání této práce jen těžko dostal. Dále bych rád poděkoval Stanislavu Schmidtovi za vznesení myšlenky tématu a za aktivní výpomoc při nastavování a průběhu experimentu.

Osobní poděkování směřuji svým přátelům, kteří mi v nejtěžších chvílích byli oporou a podali pomocnou ruku. V neposlední řadě děkuji své rodině a blízkým za nepřetržitou podporu nejen během studia.



Čestné prohlášení

Prohlašuji, že jsem bakalářskou/diplomovou práci s názvem *Název práce* vypracoval/a samostatně a použil/a k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k bakalářské/diplomové práci.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

Praze dne 30.11.2023

Podpis



Obsah

Úvod	12
1. Teoretické základy práce	13
1.1. Úvod do problematiky	13
1.2. Systém Safety Nets.....	14
1.3. Standartní komunikace v leteckém provozu	14
1.4. Jazykové požadavky na řídicí letového provozu a piloty	18
1.5. Zpětné potvrzení letového povolení (read-back)	19
1.6. Lidský hlas a jeho analýza pro účely rozpoznání řeči	20
1.7. Rozpoznávání řeči (speech recognition)	21
1.8. Výzkumy a studie rozpoznání řeči v leteckém provozu	24
1.8.1. Projekt MALORCA.....	25
1.8.2. Projekt HAAWAI	27
1.9. Limitace současného stavu	29
1.10. Shrnutí kapitoly	29
2. Metodika	31
2.1. Příprava experimentu	31
2.2. Popis a průběh experimentu	32
2.3. Vybrané fráze pro experiment.....	33
2.3. Aplikace pro záznam vyhodnocení správnosti zpětného potvrzení subjekty	35
2.4. Program pro automatické rozpoznání řeči	36
2.4.1. Import Whisper	36
2.4.2. Import speech_recognition.....	36
2.4.3. Import threading	37
2.4.4. Vyhodnocování stejnosti frází	37
2.4.5. Dotazník demografických faktorů.....	38
2.5. Zpracování dat	38
3. Výsledky	41
3.1. Korelace počtu slov v letovém povolení s celkovým počtem chybných vyhodnocení správnosti zpětného potvrzení systémem ASR	41
3.2. Úspěšnost subjektů ve vyhodnocení správnosti zpětného potvrzení	42



3.3. Úspěšnost systému ASR ve vyhodnocování správnosti zpětného potvrzení	44
4. Diskuse	47
5. Závěr	50
Seznam použité literatury	51
Seznam příloh.....	55



Seznam obrázků

Obrázek č.1: ICAO fonetická abeceda [10].....	17
Obrázek č.2: ICAO fonetický přepis názvů čísel [10]	18
Obrázek č.3: Spektrogramy jednotlivých samohlásek s vyznačenými formanty [17].....	21
Obrázek č.4: Schéma jednotlivých postupů experimentu.....	32
Obrázek č.5: Aplikace pro záznam vyhodnocení správnosti jednotlivých read-backů subjektem	35
Obrázek č.6: Různé velikosti modelů Whisper a jejich požadavky na výpočetní techniku.....	36
Obrázek č.7: grafické znázornění vztahu počtu slov ve frázi s počtem celkových chyb vyhodnocení správnosti zpětného potvrzení systémem ASR	41
Obrázek č.8: Celkový počet chyb subjektů ve vyhodnocování správnosti zpětného vyhodnocení.....	43
Obrázek č.9: Relativní a procentuální úspěšnost jednotlivých subjektů ve vyhodnocování správnosti zpětného potvrzení.....	44
Obrázek č.10: Relativní a procentuální úspěšnost ASR ve vyhodnocování správnosti zpětného potvrzení u jednotlivých subjektů	45
Obrázek č.11: Celkový počet chyb systému ASR ve vyhodnocování jednotlivých frází	45



Seznam tabulek

Tabulka č.1: Přehled počtu slov v jednotlivých frázích, celkového počtu chyb ve vyhodnocení správnosti zpětného potvrzení a procentuální úspěšnosti programu ASR ve vyhodnocování jednotlivých frází.....42

Tabulka č.2: Přehled demografických proměnných a jejich vliv na procentuální spěšnost vyhodnocení správnosti zpětného potvrzení systémem ASR46



Seznam symbolů a zkratk

%	Procento
ABSR	Automatické rozeznání řeči
AI	Umělá inteligence
AJ	Anglický jazyk
ANSP	Poskytovatel leteckých navigačních služeb
ASR	Automatické rozeznání řeči
ATC	Řízení letového provozu
ATCO	Řídící letového provozu
ATM	Soubor palubních a pozemních funkcí pro zajištění bezpečného pohybu letadel
ATS	Letecké služby
CPDLC	Řídící-pilo datová komunikace
HMI	Ergonomie stroje a člověka
Hz	Jednotka frekvence, Hertz
ICAO	Mezinárodní organizace pro civilní letectví
ILS	Přistání pomocí přístrojů
SSR	Sekundární přehledový radar
RVSP	Snížení minimálního vertikálního rozstupu
RWY	Přistávací / vzletová dráha
ŘLP	Řízení letového provozu
TLR	Úroveň připravenosti technologie



Úvod

Letecká doprava se těší stále se zvyšující popularitě, například dne 26.11.2023 padl ve Spojených státech amerických rekord co do počtu přepravených osob během jediného dne [37]. Z tohoto důvodu se neustále zvyšují nároky na pracovní efektivitu řídicích letového provozu (ATCO). Tím pádem je silný tlak na vývoj technologií a postupů, jenž by pomohly se zvyšující se zátěží ATCO, čímž by byla udržena stávající úroveň bezpečnosti letového provozu.

Některé zdroje uvádějí, že až 70 % ze všech nehod, na kterých se podílelo lidské zapříčinění, je způsobeno právě chybou související s komunikací [9]. Motivací této práce tedy je nalézt dostupné řešení pro eliminaci incidentů pramenících z chybné komunikace.

V současné době zatím není implementován žádný systém, který by sloužil k detekci nesprávného zpětného potvrzení. Hlavním důvodem, proč zatím tyto systémy nejsou implementovány do provozu je, že tato technologie založená na systému automatického rozeznávání řeči (ASR) vyžaduje specifický model učení pro využití v leteckém prostředí, ve kterém se využívá speciální forma angličtiny, letecká frazeologie. Tyto modely zatím nejsou dostatečně robustní pro zajištění bezpečného chodu systému, jelikož vykazují nedostatečné úspěšnosti v přepisu mluveného slova.

Cílem práce je navrhnout softwarový nástroj pro detekci nesprávného zpětného potvrzení (read-back) v rámci komunikace mezi řídicím letového provozu a posádkou. Dále se tato práce zaměří na otestování tohoto nástroje v simulacích pomocí navrženého experimentu. V rámci testování bude porovnána úspěšnost ve vyhodnocení správnosti zpětného potvrzení navrženého systému ASR s úspěšností testovaných subjektů. Dále pak bude zkoumán vliv demografických znaků na úspěšnost navrženého systému.



1. Teoretické základy práce

1.1. Úvod do problematiky

V řízení a organizaci letového provozu (Air Traffic Management) je komunikace mezi řídicím letového provozu (ATCO) a pilotem letadla klíčovým prvkem ATM. Radiotelefonní spojení stále představuje primární prostředek pro komunikaci mezi pilotem a ATCO, a to i přes to, že v běžném provozu již začíná nabírat na popularitě elektronický systém pro vydání letových povolení a základní komunikaci zvaný Data Link (CPDLC). Například ve vzdušném prostoru operovaném MUAC (Maastricht Upper Area Control Centre), jedním z nejrušnějších vzdušných prostorů na světě, přijímá více než 50 % prolétajících provozu alespoň nějaké povolení skrze komunikační kanál CPDLC [1]. Výjimku z používání CPDLC jsou situace, kdy se letadla pohybují ve výškách pod 10 000 stop, a to z důvodu minimalizace času, který by pilot strávil sledováním obrazovky vzhledem k blízkosti terénu. V takových případech probíhá veškerá komunikace skrze radiotelefonní kanály [2]. Organizace zabývající se provozem letecké dopravy neustále vyvíjejí nové systémy a postupy pro zvýšení bezpečnosti a zefektivnění leteckého provozu, nicméně samotná radiokomunikace mezi pilotem a řídicím letového provozu doteď spoléhá pouze na postup zpětného potvrzení neboli „read-backu“. Tím je myšleno, že řídicí vydá směrem k pilotovi letové povolení a dotyčný pilot tuto informaci zopakuje.

V případě, kdy pilot nesprávně potvrdí řídicím vydané povolení, nebo jeho část, musí být tato chyba neprodleně napravena. Vzhledem k tomu, že provozní zátěž, které jsou ATCOs vystaveni neustále stoupá, zvyšuje se tak i riziko, že řídicí letového provozu nesprávné zpětné potvrzení nezaregistruje. To by pro letový provoz mohlo mít fatální následky. Než si ATCO tuto skutečnost uvědomí, letadlo uletí i několik mil. Odhalení tedy může nastat až ve chvíli, kdy již nemá žádný manévrovací prostor pro nápravu [3].

Tato situace představuje velké riziko pro bezpečnost letecké dopravy. Řídicí letového provozu, jsou vycvičeni na zvládnutí zvýšené zátěže, i krátkodobého přetížení. V současné době může být během jediného dne registrováno přes 35 000 pohybů ve vzdušném prostoru Evropské Unie (data z července 2019) [4]. To představuje při současném stavu počtu ATCOs dlouhodobě neudržitelnou situaci. Právě z důvodu přetížení je důležité, aby se řídicí letového provozu mohli spolehnout na systémy, které jim při práci asistují. Klíčovým prvkem těchto systémů je snižování provozní zátěže a zvyšování provozní bezpečnosti letecké dopravy. Soubor těchto



systémů se nazývá Safety Nets, který byl navržen v roce 1998 organizací EUROCONTROL. Ta neustále pracuje na jejich aktualizacích a vylepšení funkčnosti.

1.2. Systém Safety Nets

Safety Nets představuje soubor podmnožin systémů, které jsou v různých aspektech leteckého provozu aktivovány v situaci, kdy už některý předchozí prvek systému selhal. Primárním úkolem je takovouto situaci zachytit a minimalizovat její následky. Představují tedy pomyslnou záchrannou síť pro případ, že veškeré předchozí ochranné mechanismy selžou. Řídicímu poskytují důležité podklady pro odhalení a tím pádem i vyhodnocení a úspěšnému vyřešení jednotlivých situací.

Dobrym příkladem pro popsání úlohy Safety Nets pro ATM (Air Traffic Management) systémy je analogie k modelu Švýcarského sýra. Safety Nets by v něm byly znázorněny jako pomyslný poslední plátek sýru, který zachytí veškeré situace, které projdou „dírami“ předchozích vrstev mechanismu [5].

V reálném provozu se prvky systému Safety Nets uplatňují například při blížícím se konfliktu letadel, kdy je řídicí letového provozu upozorněn na možný konflikt. Jako konflikt se označuje situace, kdy je horizontální rozstup dvou letadel méně než 5 NM (9,26 km), nebo vertikální rozstup méně než 1000 ft (305 m), konkrétně v oblasti se sníženým minimálním vertikálním rozstupem (RVSM), dle přehledového zařízení ATS (Air Traffic Service).

Z výše zmíněných prvků systémů Safety Nets lze vyčíst, že existuje celá řada pomůcek pro detekci potenciálního konfliktu. V živém provozu se nicméně zatím nikde nepodařilo implementovat systém, který by řídicího upozornil na nesprávné zpětné potvrzení vydaného povolení. Jedním z hlavních důvodů je nákladnost modelování potřebných dat [6]. V rámci praktické části této práce je takovýto systém navržen a testován.

1.3. Standartní komunikace v leteckém provozu

V komunikaci mezi řídicím letového provozu a posádkou letadla představuje verbální komunikace základní prostředek úspěšné výměny informací [7]. Vzhledem rozmachu letecké dopravy během první poloviny 20. století a zavedení prvních mezinárodních linek, bylo nutné zajistit schopnost vzájemného dorozumění řídicích letového provozu s posádkami letadel. Z tohoto důvodu vydala ICAO (Mezinárodní organizace pro civilní letectví) v roce 1944 v rámci Chicagské úmluvy, konkrétně v ICAO Annex 10 (Vol. I, 5.2.1.1.2) doporučení pro zavedení



anglického jazyka jakožto globálního jazyka pro radiotelefonní komunikaci. Zároveň je v tomto doporučení ale uvedeno, konkrétně v části 5.2.1.1.1., že výběr jazyka pro vedení radiotelefonické komunikace je zúžen na anglický jazyk a jazyk běžně používaný na území pozemní stanice. Ve většině případů tedy úřední jazyk území, na kterém se daná stanice ATC nachází [7,8].

Od zavedení tohoto doporučení nicméně stále nevzešel jasný požadavek na vedení této komunikace v anglickém jazyce. Účastníci provozu tak stále mají možnost vést komunikaci v libovolném jazyce s tím, že pouze po vyžádání použití anglického jazyka jsou povinni tuto skutečnost respektovat a přejít do angličtiny. Tudiž se doteď objevují situace, kdy některá stanoviště ATS mají tendenci používat primárně svůj mateřský jazyk. V Evropě je takovouto zemí například Francie, kde radiotelefonní komunikace mezi francouzskými řídicími letového provozu a piloty je poměrně často vedena ve francouzském jazyce. Tato skutečnost může mít neblahý vliv zejména na ostatní, cizojazyčné posádky na stejné frekvenci, kteří této komunikaci nerozumí, což může ovlivnit jejich situační povědomí [9].

Pro účely sjednocení různých světových dialektů anglického jazyka, jakožto univerzálního jazyka v oblasti letecké dopravy, vytvořila Mezinárodní organizace civilního letectví zjednodušenou speciální formu angličtiny. Nazývá se letecká angličtina a je využívána výhradně v oblasti letectví. Základní potřebou pro ustanovení letecké angličtiny byla eliminace nedorozumění kvůli rozdílnému porozumění kontextu letového povolení ATCO a pilotem. Specifikem letecké angličtiny je, že je v ní implementována speciální frazeologie, která usnadňuje komunikaci, snižuje nároky na úroveň angličtiny zúčastněných subjektů a tím snižuje riziko špatného porozumění mezi piloty a řídicími letového provozu [9].

Rychlost komunikace by se měla pohybovat okolo 100 vyslovených slov za minutu. Tato rychlost byla stanovena, jelikož zajišťuje dobrou srozumitelnost slov, při zachování poměrně rychlého vydání letového povolení. V reálném provozu je však poměrně běžnou praxí tuto rychlost nedodržovat, a komunikace tak probíhá ve významně rychlejším tempu [9].

Dalším ze specifik letecké angličtiny je rozdílné pojmenování jednotlivých písmen v abecedě, kdy každé písmeno v abecedě je reprezentováno slovem, viz. obrázek č.1. V praxi se lze s tímto označením písmen nejčastěji setkat u volacích znaků letadel. Výjimkou pro používání běžného označení písmen jsou zkratky, které se v letecké dopravě používají na denní bázi. Příkladem může být QNH, ILS apod. [10].



U číselných znaků se lze setkat s rozdílnou výslovností oproti běžné angličtině. Například číslovka „3“ se čte jako „tree“ namísto „three“, dále pak číslovka „9“ je vyslovována „niner“ namísto „nine“. Dalším pravidlem je, že v číslovkách se každé jednotlivé číslo, které číslovku tvoří, čte samostatně.

Výjimku tvoří čísla, která jsou přesné stovky, či tisíce, viz obrázek č. 2.



Letter	Word	Pronunciation in CR	International pronunciation
A	Alfa	ælfə	AL FAH
B	Bravo	bra:'vəu	BRAH VOH
C	Charlie	ča:li	CHAR LEE
D	Delta	delta	DELL TAH
E	Echo	ekəu	ECK OH
F	Foxtrot	fokstrot	FOKS TROT
G	Golf	golf	GOLF
H	Hotel	həu'tel	HO TTEL
I	India	indiə	IN DEE AH
J	Juliett	džu:li'et	JEW LEE ETT
K	Kilo	ki:ləu	KEY LOH
L	Lima	li:mə	LEE MAH
M	Mike	maik	MIKE
N	November	nəu'vembə	NO VEM BER
O	Oscar	o:skə	OSS CAH
P	Papa	pə'pa	PAH PAH
Q	Quebec	kə'bek	KEH BECK
R	Romeo	rəumiəu	ROW ME OH
S	Sierra	si'ere	SEE AIR RAH
T	Tango	tæŋgəu	TANG GO
U	Uniform	ju:nifo:m	YOU NEE FORM
V	Victor	viktə	VIK TAH
W	Whiskey	wiski	WISS KEY
X	X-ray	eks'rei	ECKS RAY
Y	Yankee	jæŋki	YANG KEY
Z	Zulu	zu:lu:	ZOO LOO

Obrázek č.1: ICAO fonetická abeceda [10]

Number	Czech word	English word	Pronunciation	International pronunciation
0	NULA	ZERO	ziərəu, 'zi:rəu	ZE-RO
1	JEDNA	ONE	wan	WUN
2	DVA	TWO	tu	TOO
3	TŘI	THREE	tri	TREE
4	ČTYRY	FOUR	fo:r, 'faur	FOW-er
5	PĚT	FIVE	fajf	FIFE
6	ŠEST	SIX	siks	SIX
7	SEDUM	SEVEN	sevn	SEV-en
8	OSUM	EIGHT	eit	AIT
9	DEVĚT	NINER	najnr	NIN-er
desetinná čárka	ČÁRKA	DECIMAL	desəml	DAY-SEE-MAL
sto	STO	HUNDRED	handrid, handrəd	HUN-dred
tisíc	TISÍC	THOUSAND	tauznd	TOU-SAND

Obrázek č.2: ICAO fonetický přepis názvů čísel [10]

1.4. Jazykové požadavky na řídící letového provozu a piloty

Jelikož je anglický jazyk doporučeným jazykem pro vedení komunikace v leteckém provozu, je samozřejmé, že uchazeči o post pilota a řídícího letového provozu musí splnit základní podmínky, které zaručují subjektu certifikaci k výkonu povolání.

Úspěšní držitelé licence pro letecký personál by dle ICAO Annex I měli splňovat tyto jazykové podmínky:

- a) Efektivní vyjadřování a porozumění při hlasové komunikaci (telefon/radiotelefon) a komunikaci tváří v tvář.
- b) Přesná a srozumitelná komunikace o běžných, konkrétních a pracovních tématech.
- c) Používání vhodných komunikačních strategií k výměně zpráv, rozpoznat a vyřešit nedorozumění (např. zkontrolovat, potvrdit nebo ujasnit přijatou informaci) v obecném nebo pracovním kontextu.



- d) Úspěšně a s relativní lehkostí zvládnout lingvistické překážky způsobené komplikacemi, či neočekávaným průběhem událostí, které se vyskytují zejména v kontextu rutinních pracovních situací, či komunikačních úkolů, se kterými se jinak běžně setkávají.
- e) Používat dialekt, či akcent, který je srozumitelný pro leteckou komunitu. [8]

K tomu, aby kandidát na pozici komerčního pilota, či řídicího letového provozu mohl obdržet platnou licenci, je nutné, aby na škále ICAO pro úroveň anglického jazyka dosáhl alespoň na čtvrtou (operativní) ze šesti kategorií. Úrovně 4 až 6 na stupnici anglického jazyka dle ICAO zaručují dokonalé osvojení si letecké angličtiny a frazeologie s ní spojené, jelikož právě letecká angličtina představuje zásadní část komunikace v letecké dopravě.

1.5. Zpětné potvrzení letového povolení (read-back)

Zpětné potvrzení je definováno organizací ICAO jako procedura, při které přijímací stanice (posádka letadla) zopakuje obdrženou zprávu (nebo její relevantní část) zpět ke stanici vysílající dané letové povolení (stanoviště ATC), a to za účelem potvrzení správného přijetí vydaného povolení. [7]

Mezinárodní organizace civilního letectví tento postup popisuje jako uzavřenou smyčku potvrzení nebo opravy letového povolení v těchto čtyřech krocích:

- 1) Odesílatel předá zprávu.
- 2) Příjemce aktivně naslouchá přijaté zprávě.
- 3) Příjemce zopakuje obsah zprávy zpět směrem k odesílateli.
- 4) Odesílatel aktivně naslouchá a kontroluje správnost zpětného potvrzení. [12]

V praxi to znamená, že pro každé povolení, které řídicí letového provozu směrem k posádce vydá, se musí od dané posádky letadla k řídicímu letového provozu stejná informace vrátit. Při radiotelefonické komunikaci mezi ATCO a posádkou letadla je dle ICAO stanoveno, které části vydaného povolení musí být zpětně potvrzeny. Jsou to:

- a) Povolení k využití letové tratě.
- b) Povolení a instrukce k vjezdu na, přistání na, vzletu z, zastavení na linii pro vyčkávání, přejetí přes, či provedení back-track na RWY.



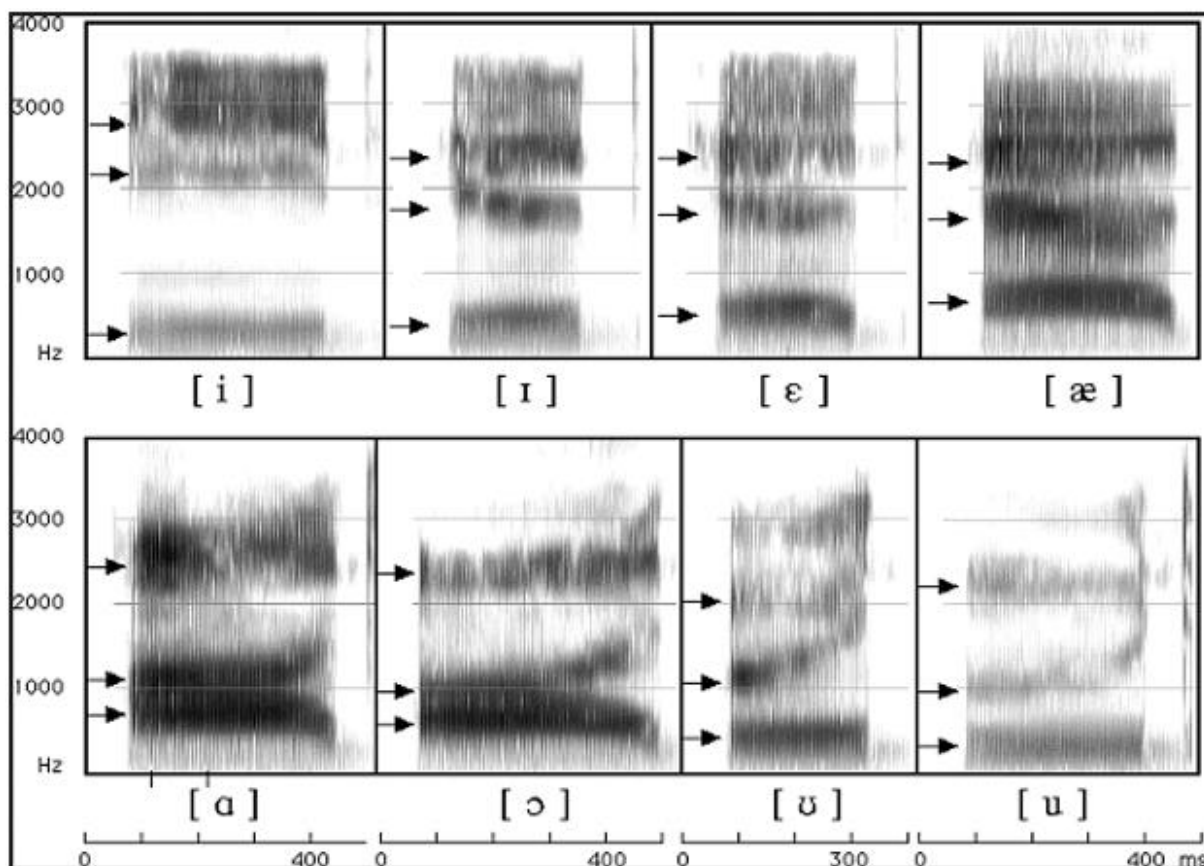
- c) RWY v provozu, nastavení výškoměru, kódy SSR, instrukce k hladině letu, směrové instrukce a instrukce k rychlosti letadla. Dále pak v případě, že je ATCO vydána nebo je obsažena v ATIS, informace o přechodných hladinách. [13]

Řídící letového provozu je povinen poslouchat a vyhodnotit správnost zpětného potvrzení pro zajištění, že jím vydané letové povolení bylo posádkou letadla správně přijato a pochopeno. V případech, kdy se zpětné potvrzení neslučuje s vydaným povolením, které bylo řídicím letového provozu vydáno, musí ATCO učinit nápravné kroky a letové povolení ihned zopakovat. V opačném případě hrozí, že dojde k porušení např. minimálních rozestupů dvou letadel.

V případech datové komunikace skrze systém CPDLC, zpětné potvrzení není vyžadováno, pokud není příslušným poskytovatelem ATS (Air Traffic Services) uvedeno jinak.

1.6. Lidský hlas a jeho analýza pro účely rozpoznání řeči

Mechanismus vzniku lidského hlasu lze rozdělit do tří částí lidského těla – plíce, hlasivky v hrtanu a artikulační ústrojí. Zdrojem hlasu je výdechový proud z plic, který je následně modulován kmitáním hlasivek a dává tak vznik primárnímu hrtanovému tónu. [14]. Pro účely zajištění funkčnosti systémů pro rozpoznání řeči je klíčová rezonance uvnitř vokálního traktu, zejména konkrétně rozdíly v rezonanci pro jednotlivé samohlásky. Tyto formy rezonance se nazývají formanty, či formantové frekvence a představují významné oblasti koncentrace energie ve zvukovém spektru. Z toho vyplývá, že jednotlivé formanty představují jakási maxima lokální energie ve spektru, díky čemuž je možné identifikovat, viz obrázek č.4. Formanty jsou označeny písmenem „F“ a číslovkou. Formant s nejnižší frekvencí je označován F1, další pak F2 atd. Základní frekvence tónu hlasu je F0, nicméně ta i přes to, že má stejné označení, formantem není. [15,16]



Obrázek č.3: Spektrogramy jednotlivých samohlásek s vyznačenými formanty [17]

K identifikaci jedné samohlásky je zapotřebí lokalizovat alespoň 2 formanty. Pro identifikaci samohlásky se hodnota frekvence (Hz) formantu s menší frekvencí odečte od formantu s vyšší frekvencí. Výsledný rozdíl ve frekvencích poté koresponduje s určitou samohláskou, kde každá má své rozmezí hodnot, které ji reprezentují.

Vyhodnocení formantů je univerzálním prostředkem pro účely získání samohlásky či fonému a to, ačkoliv se výška hlasu, intonace, či výslovnost na základě různých demografických vlivů u každého člověka liší. Jednotlivé frekvence formantů se na základě těchto vlivů mohou lišit, nicméně rozdíl ve frekvencích 2 formantů zůstává, v rámci určité tolerance, stejný. Rozdíly mohou vznikat na základě rozdílné výslovnosti jednotlivých fonémů u různých jazyků [18].

1.7. Rozpoznávání řeči (speech recognition)

Rozpoznávání řeči, nebo také automatické rozpoznání řeči (ASR), je mezi-oborová disciplína spadající pod informatiku a počítačovou lingvistiku. Jedná se o schopnost určitého programu přenést lidskou řeč do psaného formátu. Často je tato disciplína zaměňována za rozpoznání



hlasu, což je obor zaměřující se na identifikaci jednotlivých hlasů, ale obsahem mluveného slova se nezabývá [19].

Typický přínos využití této technologie nastává v situacích, kdy ruce nebo zrak jsou zaměřeny na jinou činnost a ovládání hlasem tak umožňuje běh souběžných činností (multi-tasking). Tím se zkracuje doba pro dokončení prováděného úkonu. Další možnosti použití lze nalézt například v informačních službách přes telefon, kdy je intuitivnější informaci ústně sdělit než mačkáním tlačítka pro výběr požadované operace. S rozpoznáváním řeči se lze také setkat u různých pomůcek pro zdravotně postižené.

V ideálním případě by byl rozpoznávač řeči schopen přesného vyhodnocení slov od jakéhokoliv uživatele v jakémkoliv prostředí. V reálném světě je ale výkon ovlivněn nepřeberným množstvím vnějších vlivů. Obsáhlý slovník, velké množství uživatelů, kontinuální řeč vs. izolovaná slova a zejména hlučné okolní prostředí, to vše jsou faktory, které zásadně ovlivňují funkčnost algoritmu [20].

První pokusy o vytvoření zařízení pro rozpoznávání řeči se datují do 50. let 20. století. Ve výzkumných laboratořích telekomunikační společnosti Bell kolektiv vědců Stephen Balashek, R. Biddulph a K. H. Davis vytvořili analogový systém s názvem „Audrey“, který byl schopen rozpoznávat jednociferné číslovky. Systém byl funkčně omezen pouze na jednoho uživatele a fungoval na principu identifikování frekvencí formantů samohlásek specifických pro jednotlivé číslovky [21].

Od tohoto průlomu udělala technologie rozpoznávání řeči významný posun, a to zejména s vývojem informačních technologií. Právě díky těm se rozpoznávání řeči přesunulo do digitálního formátu. Mikrofon tedy převede mechanické vlnění (hlas) do elektrických signálů, které jsou počítačem převedeny do digitální podoby. Následně software pro rozpoznávání řeči tyto signály analyzuje za účelem identifikace jednotlivých fonémů (nejmenší jednotka řeči). Každý foném se skládá alespoň ze dvou formantů. Na závěr pak algoritmus tyto signály shlukuje a spáruje je s vhodným přepisem, který dané signály reprezentuje [22].

Dalším významným krokem dopředu byla možnost využití umělé inteligence (AI) a to za pomoci technologie neurálních vláken. Tento pojem si lze představit jako model interpretující určité vlastnosti lidského nervového systému. Neurální síť se skládá, stejně jako nervová soustava, z velkého počtu nervových zakončení (neuronů), která jsou organizována



v jednotlivých vrstvách a navzájem propojena vlákny s rozdílnými silami. Architektura neurálních sítí rozděluje vrstvy do tří kategorií: vstupní vrstva, ukryté vrstvy a výstupní vrstva. Informace je vložena do vstupní vrstvy, odkud putuje dál do první ukryté vrstvy. Ukrytých vrstev může být neomezený počet, a právě v nich probíhá veškerý proces vyhodnocení dat. Data zpracována touto vlákny propojenou sítí jsou předána na výstupní vrstvu, která obsahuje jediný neuron s výstupními daty [23].

Interakce mezi jednotlivými neurony funguje na principu ohodnocení vazeb, kdy data přijímaná vláknem vystupujícího z předchozího neuronu jsou na vstupu následujícího neuronu vynásobena hodnotou, kterou je toto vlákno ohodnoceno. Tyto hodnoty od každého vstupujícího vlákna do neuronu jsou posléze sečteny. Každý neuron má vlastní prahovou hodnotu, které dokud nedosáhne, tak neposílá signál do navazujících vláken a neuronů. V momentu, kdy součet hodnot překročí danou prahovou hodnotu, neuron vysílá signály všemi vystupujícími vlákny do následných neuronů [24].

Na základě audio dat, které má umělá inteligence k dispozici jsou analyzovány veškeré fonémy, které jsou počítačem izolovány. Tato data se nazývají modely, a některé z nich mohou obsahovat až miliony hodin záznamů. Umělá inteligence vyhodnocuje, s jakými dalšími fonémy jsou nejčastěji spárovány, či v jakém kontextu jsou užity pro kompletaci co nejpresnějšího přepisu slov. Tato skutečnost je hlavní výhodou v komplexních prepisech reálných konverzací, ale v jiných případech může být naopak zdrojem nepřesností. Například v komunikaci mezi pilotem a řídicím letového provozu se oproti běžným konverzacím vyskytují nestandardní kombinace slov a některá slova se také vyslovují jinak, než je běžné. Tyto skutečnosti mohou ve výsledku mít vliv na přesnost speech recognition programu, jenž pro svoji funkčnost využívá AI. Oproti tomu, pokud by se vytvořil model, který by obsahoval pouze audio záznamy z leteckého prostředí, na výslednou úspěšnost přepisu by to mohlo mít výrazně pozitivní vliv. V současné době žádný takový model není komerčně ani volně dostupný. Na základě v současné době prováděných studií je pravděpodobné, že některé organizace věnující se provozu letecké dopravy právě takovému modelem pro cvičení umělé inteligence disponují, či jej vyvíjejí.

Nejčastěji se lze setkat se speech recognition programy zejména v oblasti virtuálních asistentů. Asi nejznámějším představitelem je virtuální asistentka Siri od společnosti Apple. Následuje například virtuální asistentka Alexa od společnosti Amazon, či Google Assistant od společnosti Google. Jejich funkčnost je založena na reakci na nějaký „trigger phrase“. To



znamená, že reagují na specifickou frázi, která je „probudí“ a od té doby naslouchají pro instrukce vydané uživatelem. [22]

V leteckém odvětví se se systémy využívajícími technologii rozpoznání řeči lze setkat například v moderních typech stíhacích letounů. Jeden z pokusů spojených s implementací speech recognition v leteckých kokpitech probíhal na v tuzemsku dobře známých stíhacích letounech JAS-39 Gripen. Výstupem této studie, provedené na Královském technickém institutu ve Stockholmu v roce 2001 bylo, že gravitační přetížení mělo významný vliv na úspěšnost transkripce. Dalším faktorem ovlivňujícím úspěšnost transkripce byl hlasitý okolní zvuk a zejména pak dýchání pilotů. Prvotní úspěšnost programu se pohybovala pod 15 %, nicméně po předělání modelu pro zachycení a vyhodnocování získaných dat se finální úspěšnost pohybovala okolo 80 %. Pro transkripci mluveného slova při gravitačním přetížení o hodnotě 4G se konečná úspěšnost přepisu pohybovala mezi 60 % a 70 % [20]. Současné moderní stíhací letouny, jako například Lockheed Martin F-35, či Eurofighter Tycoon jsou již plně vybaveny systémy pro rozpoznávání řeči, díky kterým mohou piloti za pomoci hlasu řídit určité ovládací prvky letounu.

1.8. Výzkumy a studie rozpoznání řeči v leteckém provozu

Automatické rozpoznání řeči je vnímáno jako technologie s velkým potenciálem pro zvýšení bezpečnosti a efektivity v leteckém provozu [25]. Největším limitujícím faktorem pro implementaci systémů pro rozpoznání řeči je, že se stále jedná o relativně novou technologii, zejména pak z hlediska jejího univerzálního využití [26]. Civilní letectví je samo o sobě, alespoň co se rychlosti implementace nových procesů týče, poměrně neprogresivním odvětvím průmyslu. Zároveň je na evropském území přísná regulace ze strany Evropské komise, která již od prvotní fáze návrhu vyžaduje od každé nové technologie velmi rigorózní analýzu bezpečnosti, za účelem prokázání spolehlivosti a funkčnosti konečného produktu. Výsledkem je, že ačkoliv již jsou dostupná komerční řešení pro spolehlivý přepis mluveného slova, v leteckém provozu se zatím nepoužívají [27].

Projekt SESAR PJ16 CWP HMI, jenž probíhal od listopadu 2016 do prosince 2019 měl za cíl analyzovat možnosti pro snížení provozních nákladů poskytovatelů leteckých navigačních služeb (ANSP). Z velké části se tento projekt zaměřil na interakci mezi řídícími letového provozu a přístroji, které využívají pro svou práci, ve zkratce Human Machine Interface (HMI). Tato studie se dotkla i technologie automatického rozpoznání řeči, kdy byly vytvořeny prototypy těchto systémů, nicméně testování probíhalo pouze v laboratorním prostředí, a ne



v reálném provozu. Závěrem studie bylo, že tato technologie je ve fázi připravenosti (TLR) 4 [28]. TLR je metoda pro ohodnocení vyspělosti technologie během výzkumné a vývojové fáze, která zaručuje konzistentní a univerzální škálu. TLRs mohou nabývat hodnot od 1 do 9, kde hodnota 9 představuje nejvíce vyspělou technologii [29].

Studie [26] (navazující na výše zmíněný projekt SESAR a zejména pak na projekt PJ.10-W2-96 [30]) v listopadu 2023 se zaměřila na automatizovaný přepis informací na jednotlivých formulářích letadel viditelných na přehledovém zařízení, a to ve chvíli kdy ATCO vydá pokyn ke změně směru, rychlosti, letové hladiny apod. Hodnota úspěšnosti vyhodnocení transkripce činila 96,9 %, nicméně jakmile jediné slovo z fráze bylo vyhodnoceno s nízkou mírou jistoty správnosti, formulář nemohl být přepsán. Proto byl implementován algoritmus, který povoloval určité míry tolerance při nejistém vyhodnocení, což mělo významně pozitivní vliv na finální počet automaticky aktualizovaných formulářů (92,1 %), ačkoliv se zvýšil počet falešných pozitiv, tedy nesprávného přepsání formuláře (3,4 %). V simulovaném scénáři, který trval přibližně 14 hodin, se díky využití automatizovaného vyplňování formulářů skrze technologii rozpoznání řeči významně snížila doba, kterou ATCO strávil manuálním vyplňováním formulářů. ATCO bez použití ASR technologie strávil přepisováním formulářů 12 700 sekund, tedy přibližně 25 % doby trvání celé simulace. Při využití technologie implementující ASR byla tato hodnota snížena na 405 sekund, což představuje přibližně 31x snížení časové zátěže. Hypotéza z roku 1990 předpovídala, že implementace ASR technologie by v budoucnu mohla mít efekt na redukci výskytu chyb způsobených lidským faktorem, což se přeneso do zvýšení bezpečnosti letecké dopravy [31]. Tuto předpověď potvrzují i výsledky výše zmíněné studie [26], která uvádí, že míra chybovosti pro nesprávné přepsání formuláře činí s využitím ASR systému 3,4 %. Oproti tomu, ve scénáři bez použití ASR systému, se chybovost rovnala 11,6 %. Na základě této studie tedy systémy ASR představují velký přínos pro snížení zátěže řídicích letového provozu a zejména pro zvýšení bezpečnosti letecké dopravy.

1.8.1. Projekt MALORCA

Hlavním cílem projektu MALORCA (Machine Learning of Speech Recognition Models for Controller Assistance) bylo zjištění možnosti využití automatizovaného přepisu informací do leteckých stripů, čímž by se řídicím letového provozu výrazným způsobem snížil čas strávený touto činností. Společně s tím byla zkoumána možnost využití strojového učení bez dohledu, která by výrazně snížila náklady spojené s uvedením takovéto technologie do provozu. [32].



ABSR (Assistent Based Speech recognition) je potenciálním řešením pro snížení pracovní zátěže řídicích letového provozu. Jedná se o kombinaci asistenčního systému a systému rozpoznání řeči. Generátor hypotéz (hypotheses generator) využívá výstupu asistenčního systému. Generátor neví přesně jaké pokyny řídicí zadá, ale ví, které z nich jsou v dané situaci pravděpodobnější. Hypotézy se zadávají do bloku automatického rozpoznávání řeči (automatic speech recognition block). Extrahovaná letová povolení jsou systémem rozpoznání řeči (ASR) odeslána zpět do asistenčního systému, je zkontrolována jejich věrohodnost a pravděpodobnost v daném kontextu. Systém Automatického rozpoznání řeči (ASR) vyhodnotí nejbližší shodu vydaného povolení s asistentem dodanými očekávanými hypotézami. Pokud vyhodnotí, že jsou si tyto dvě dodaná letová povolení (první povolení vydané jako hypotéza asistentem, druhé povolení vydané ATCO) dostatečně podobné, doplní tato data do náležité kolonky v příslušném stripu. [32]

Pro tuto studii bylo zaznamenáno více než 100 hodin radarových dat a audiozáznamů z komunikace řídicích letového provozu s piloty pro přibližovací oblasti Vídně a Prahy. 20 % dat z této komunikace bylo ručně přepsáno a byly identifikovány pokyny, které jsou relevantní pro závěrečnou analýzu. Pro účely přepisu relevantních ATC informací jsou z dat odfiltrovány fráze jako například „good morning“, které pro analytické potřeby nejsou relevantní. Celkem byly přepsány čtyři hodiny (čistého času bez úseků ticha) hlasových záznamů pro Vídeň i Prahu. [32]

Počáteční míra úspěšnosti rozpoznávání letových povolení činila pro Prahu a Vídeň 80 %, respektive 60 %. Rozdíly v úspěšnosti jsou přiřazené k rozdílné kvalitě získaných audio dat, kdy data z Prahy byla kvalitnější. Do modelu ABSR bylo posléze ručně přepsáno 25 % z nevyhodnocených dat za účelem vylepšení modelu ABSR. Výsledná úspěšnost systému zaznamenala výrazné zlepšení. Postupně se v následujících měsících přepsalo vždy dalších 25 % z nepřepsaných dat (celkem tedy 25 %, 50 %, 75 %, 100 %) a to za účelem emulace efektu strojového učení na měsíční bázi. Konečná úspěšnost systému se nakonec zvýšila na 92 % pro Prahu, respektive na 83 % pro Vídeň. [32]

K dosažení těchto výsledků je zapotřebí plnohodnotného systému rozpoznávání řeči založeném na asistentovi. Běžný systém rozpoznávání řeči zaznamenává hlasový signál pořízený mikrofonom a transformuje jej jako posloupnost slov v textovém formátu. Rozpoznávání řeči na bázi asistenta (ABSR) využívá výstupu k předpovědi souboru pokynů řídicího, které jsou v aktuální situaci na základě radarových dat realistické. Tento přístup



výrazně zmenšuje interní data, kterými musí systém procházet a zároveň zajišťuje kontrolu věrohodnosti výstupů. [32]

Projekt MALORCA byl první ukázkou potenciálu (polo)automatické adaptace nástrojů pro asistenci řídicím letového provozu na základě učení se z velkého množství zvukových a radarových dat denně zaznamenávaných v operačních místnostech. K ověření nebo přizpůsobení tohoto přístupu jiným letištím, než jsou Vídeň a Praha je zapotřebí dalšího výzkumu. MALORCA se opírá o méně než 50 hodin dat v doméně (letiště), což je velmi omezené množství. Nicméně i z takto malého množství analyzovaných dat provozu letišť Prahy a Vídně je zřejmé, že strojové učení bez dohledu je schopno znatelně zlepšit míru rozpoznávání povelů, a že automatické učení z radarových dat a hlasových záznamů může snížit náklady na sběr dat, urychlit vývoj a snížit potřebu manuální adaptace. Vzhledem k tomu, že je možné dosáhnout míry rozpoznání vydaných povolení větší než 95 % [33], se situace, kdy ATCO musí ručně opravit výstup systému ABSR vyskytne v průměru v jednom z dvaceti případů, což má za následek snížení zátěže a zvýšení produktivity ATCO. [32]

MALORCA ukázala, že s pomocí nesledovaného strojového učení je možné automatizovat systém rozpoznání řeči, s minimální potřebou další podpory od externích činitelů (expertů). MALORCA vyvinula nový přístup k automatickému rozpoznávání řeči, a to především díky algoritmům, které posuzují zvukové nahrávky v daných kontextech a vyhodnocují věrohodnost a pravděpodobnost využití daných pokynů v těchto kontextech. Algoritmy strojového učení můžeme považovat za instrumentální část tohoto projektu a jedním z hlavních důvodů pro dosažení pozitivních výsledků. Zároveň se díky strojovému učení výrazně snížila potřeba pro manuální přepis zvukových nahrávek a dá se tedy považovat za mnohem více efektivní. Tento faktor naznačuje, že implementace podobných systémů do reálných center ATS by byla velmi silně doporučována. [32]

1.8.2. Projekt HAAWAI

Podpora automatizace vyvinutá v první vlně programu SESAR zahrnuje využití automatického rozpoznávání řeči, které snižuje množství ručních vstupů dat ze strany řídicích letového provozu. Výkon automatického rozpoznávání řeči v tomto programu byl omezený, např. z důvodu neschopnosti rozlišit přízvuky řídicích, deviací od standardních frází a omezeného výkonu rozpoznávání v reálném čase. Předchozí výzkum financovaný projektem MALORCA ukázal, že lze dosáhnout uspokojivého výkonu pomocí nových přístupů strojového učení založených na datech. Cílem projektu HAAWAI (Highly Automated Air Traffic Controller



Workstations with Artificial Intelligence Integration), založeného na výsledcích projektu MALORCA, byl výzkum a vývoj spolehlivého a přizpůsobivého řešení pro automatický přepis hlasových pokynů vydávaných řídicími letového provozu i piloty [34].

Projekt využíval strojové učení spolu s velmi rozsáhlými daty, na základě kterých byla vyvinuta nová sada modelů rozpoznávání řeči v komplexním prostředí ATS terminálního prostoru Londýn a vzdušného prostoru nad Islandem. K tomuto projektu byly využity hlasové nahrávky reálné komunikace mezi piloty a řídicími [34].

Cílem projektu HAAWAIi bylo provést ověřovací zkoušky v náročných prostředích, tj. přímé propojení s reálnými údaji z operační místnosti. Hlavní potenciální využití tohoto projektu je v oblasti detekce chyb pilotů při zpětném potvrzení. Dílčím cílem projektu HAAWAIi bylo tedy zvýšit validitu modelů rozpoznávání řeči. Digitalizace hlasových projevů řídicích a pilotů může být využita pro širokou škálu kvalitativních znaků souvisejících s bezpečností a výkonností ATCO, včetně automatického vyplňování informací do elektronických letových stripů [34].

V rámci projektu byl vyvinut nový software pro rozpoznávání řeči na základě modelu využívajícího technologii neurálních sítí. Projekt vychází z velmi rozsáhlého sběru dat. Partneři projektu (Londýn a Island) nahráli přes 500 hodin hlasových nahrávek komunikací mezi piloty a řídicími. 40 hodin z těchto nahrávek bylo manuálně přepsáno. Po dodání pouze jedné hodiny přepsaných dat do modelu rozpoznávače řeči HAAWAIi se frekvence rozpoznávaných slov zdvojnásobila. Po trénování systému na veškerých přepsaných datech činila úspěšnost rozpoznání slov přes 95 % pro řídicí a přes 90 % pro piloty [34].

Strojové učení bylo rovněž využito pro vytvoření asistenta pro detekci chyb ve zpětném potvrzení (příklad chyby ve zpětném čtení: řídicí povolí pilotovi vystoupat do 7 000 stop a pilot povolení zopakuje jako povolení stoupat do 8 000 stop). Toto by mohlo způsobit riziko kolize v případě, že by chyba zůstala ATCO nedetekovaná. Asistent vygeneruje výstrahu, když taková chyba nastane. Asistent pro detekci chyb byl vyhodnocen v laboratorním prostředí pěti řídicími z Islandu, kteří zjistili, že počet chyb detekovaných při testování byl přes 80 % z přepisů reálných dat a počet falešně identifikovaných chyb (tedy falešné negativy) byl pod 20 % [34].

Snížením množství zátěže opuštěním manuálního zapisování informací může efektivní systém rozpoznávání řeči výrazně zvýšit bezpečnost letového provozu, a to potenciálním zvýšením přesnosti řídicích letového provozu, kteří tak budou mít více prostoru pro aktivní sledování provozu. Rozpoznávání řeči by se také mohlo využívat pro podporu simulací práce na



stanovištích ATC, čímž by se výrazně zjednodušil a zefektivnil proces výcviku, který by mohl potenciálně probíhat na dálku [34].

Prototyp systému fungoval v terminální oblasti Londýna, nejrůšnějšího vzdušného prostoru v Evropě, a v islandském vzdušném prostoru, který pokrývá přes 5 milionů km čtverečních. Systém byl schopen detekovat hlasový vstup pilotů s chybovostí v rozpoznávání slov pod úrovni 10 %, a to navzdory rozmanitým přízvukům pilotů z celého světa a velmi rušné radio kanály [34].

1.9 Limitace současného stavu

Agenturami řízené projekty jako SESAR financují z fondů Evropské unie výzkumné studie, které mají za cíl nalézt nejefektivnější cestu pro zavedení používání systémů ASR v reálném provozu. Jen náklady na výše popsané projekty MALORCA a HAWAII činily více než 2,5 mil. EUR [32, 34]. Tudiž je patrné, že existuje značná motivace pro vytvoření a následnou implementaci systémů ASR mezi prvky Safety Nets. Ekonomická stránka a zejména vysoká nákladnost vývoje těchto systémů však lze považovat za limitující faktor.

Z komerčně dostupných produktů spojujících systémy ASR a letectví existuje na trhu zařízení od společnosti Appareo, jenž zapisuje komunikaci mezi ATCO a pilotem. Tento produkt je zaměřen zejména na piloty, kterým nabízí možnost zobrazit na tabletu informace získané komunikací s ATCO v textové podobě. Jeho limitací je, že není certifikovaný úřady pro civilní letectví, tudiž v této chvíli představuje pouze doplněk při komunikaci [35].

V současné době nejsou k dispozici žádná certifikovaná komerční řešení pro implementaci systémů ASR do leteckého provozu. Motivací této práce bylo ověření možnosti vytvořit takovýto systém za použití co nejméně nákladných řešení.

1.10. Shrnutí kapitoly

Radiotelefonní komunikace navzdory zvyšující se popularitě datové komunikace CPDLC (Controller Pilot Data Link Communication) stále představuje významnou část komunikace mezi řídicími letového provozu (ATCO) a piloty. V současné době se díky značným pokrokům ve vývoji strojového učení a pokrokům ve výkonu výpočetní techniky intenzivně vyvíjejí systémy takzvaných asistentů. Tito asistenti představují soubor různých systémů, které navzájem spolupracují za účelem zajištění co nejširšího záběru asistence. Potřeba pro tyto asistenty vzniká vysokou poptávkou po letecké dopravě, čímž vzrostlo pracovní vytížení ATCO. Právě použití asistentů by mělo vést ke zvýšení produktivity ATCO (např.



automatizovaným přepisem dat z hlasové komunikace do leteckých stripů), ale zejména ke zvýšení bezpečnosti letecké dopravy.

Jedním z těchto systémů je systém pro automatické rozpoznání řeči (ASR), jenž slouží k transkripci mluveného slova do textového formátu. Na základě této transkripce pak mohou další asistenti s touto informací pracovat, v kontextu této práce například kontrolovat správnost zpětného potvrzení pilotem. Tato automatizovaná detekce v reálném provozu upozorní ATCO na nesprávné zpětné potvrzení ihned, tudíž by se tím eliminovala situace, kdy ATCO chybu ve zpětném potvrzení nezaregistruje a chybu objeví až na základě přehledových dat v situaci, kdy již nebude možné zabránit incidentu, např. porušení minimálního rozestupu letadel.

Systémy ASR se v běžném životě využívají stále více, nicméně do letecké dopravy, konkrétně do živého provozu, tyto systémy zatím implementovány nebyly. Jedním z důvodů jsou vysoké náklady na vytváření univerzálních modelů pro rozpoznání řeči v letecké dopravě, kdy letecká angličtina představuje specifickou formu jazyka, jenž často skladbou slov nedává v běžné angličtině smysl. Tento problém se snaží řešit projekt SESAR, konkrétně projekty MALORCA a navazujícím projektem HAAWAI, které navrhují nekontrolované strojové učení za účely zdokonalování těchto modelů pro rozpoznání řeči. Takovýto přístup zajišťuje autonomní strojové učení, ke kterému není potřeba tisíce hodin práce expertů pro zpracování dat. Dalším faktorem je pak stále nedostatečná úspěšnost ASR systému v přepisu hlasové komunikace, nicméně po implementaci a úpravě modelů se tato úspěšnost blíží kýžené hodnotě.



2. Metodika

Pro účely práce byly navrženy vlastní metody měření. Byl vytvořen a proveden experiment, který je popsán v následující kapitole.

Primárním cílem experimentu bylo porovnat úspěšnosti různých systémů pro automatické rozpoznání řeči (ASR) při analýze správnosti zpětného potvrzení při radiotelefonní komunikaci pilotů s řídicím letového provozu (ATCO). Vedlejším cílem bylo porovnání úspěšnosti v detekci nesprávného zpětného potvrzení těchto programů s vyhodnocením uskutečněným řídicím letového provozu. Dalším vedlejším cílem bylo nalezení korelace mezi délkou letového povolení a vlivem demografických proměnných na úspěšnosti vyhodnocení systémem ASR. Při hlubším zkoumání se nicméně nepodařilo nalézt žádné volně či komerčně dostupné systémy, které by byly po spuštění automaticky funkční. Z tohoto důvodu se z počátku tato práce zaměřila na vytvoření takového programu, který bude schopný přepisu komunikace a vyhodnocení stejnosti zpětného potvrzení, a to v programovacím jazyce Python [Příloha 1].

Experiment nezkoumá přesnost přepisu mluveného slova systémem ASR, zaměřuje se pouze na schopnost systému vyhodnotit, zda se shoduje obsah vydaného letového povolení a jeho zpětného potvrzení. Systém rovněž vydá upozornění v případě, že se obsah frází liší.

2.1. Příprava experimentu

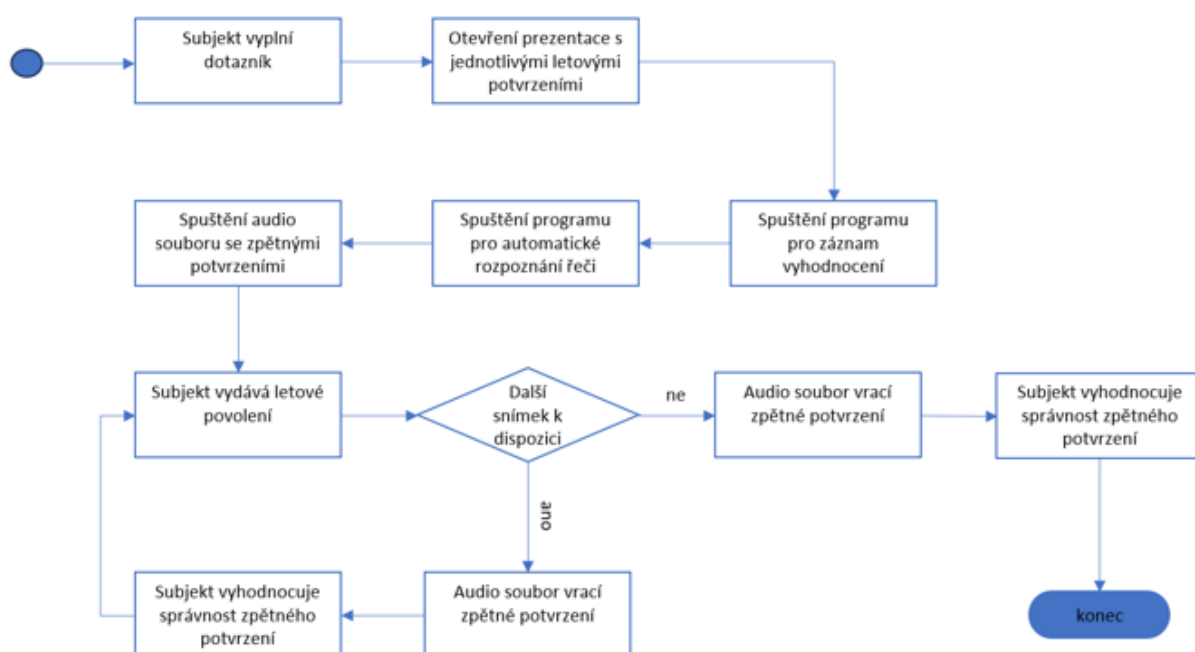
Pro účely zjednodušení sběru dat vyhodnocení správnosti zpětného potvrzení subjekty byl v programovacím jazyce Python vytvořen program, ve kterém za pomoci tlačítek Ano / Ne testovaný subjekt zaznamenával svou odpověď vyhodnocení správnosti zpětného potvrzení [Příloha 2]. Pro zajištění stejnosti dat bylo pomocí programu Audacity předem nahráno a následně editováno dvacet vybraných frází, které představovaly v tomto scénáři zpětné potvrzení pilota. Testovanému subjektu, tedy pomyslnému řídicímu letového provozu byla poskytnuta PowerPoint prezentace, která obsahovala dvacet vybraných letových povolení. Tyto fráze se lišily jak samotnou délkou, tak i komplexností jednotlivých vydaných letových povolení, a to za účelem nalezení korelace délky fráze s mírou úspěšných vyhodnocení programem ASR. Zakomponovány byly i specificky vyslovované číslovky (např. 3 = „tree“, 9 = „niner“), které by pro tuto iteraci systému mohly potenciálně představovat překážku v úspěšnosti správného rozeznání. Pro sběr dat byla do obou vytvořených programů Python implementována funkce exportu výsledků do .csv souborů, ze kterých byla data převedena do tabulkové a grafické podoby. Za účelem analýzy demografických vlivů na úspěšnost vyhodnocení správnosti zpětného potvrzení softwarem byl testovaným subjektům předložen

dotazník, který sledoval tři různé proměnné: pohlaví subjektu, rodný jazyk a úroveň anglického jazyka. [Příloha 3]

Samotný průběh experimentu je popsán v následující kapitole 2.2 Popis a průběh experimentu.

2.2. Popis a průběh experimentu

Samotný experiment probíhal od 1.10.2023 do 18.11.2023 a účastnilo se jej 50 dobrovolníků. Dobrovolníci se skládali zejména ze studentů fakulty dopravní spadající pod České vysoké učení technické v Praze, z nichž většina byla aktivními uživateli simulátoru řízení letového provozu (ATC), jenž se na fakultě nachází. Silné zastoupení mezi subjekty měli také řidiči letového provozu ve výcviku (ŘLP), či profesionální pseudopiloti zaměstnaní řízením letového provozu České republiky (ŘLP).



Obrázek č.4: Schéma jednotlivých postupů experimentu

Cvičení se v reálném čase vždy účastnil jen jeden subjekt – řidiči letového provozu (ATCO). ATCO byl umístěn na pomyslné stanoviště, které se nacházelo na dedikovaném počítači na simulátoru ATC na Fakultě dopravní ČVUT. Na počítač byly nainstalovány veškeré potřebné programy a subjektu byla předána sluchátka s mikrofonom. Úkolem subjektu bylo ústně vydat



letové povolení vypsané v PowerPoint prezentaci a vyčkat než se mu z předem nahraného audio souboru vrátí do sluchátek zpětné potvrzení od „pilota“. Následně ATCO (subjekt) porovná, zda se informace obsažené ve zpětném potvrzení shodovaly s informacemi, které pilotovi předal. Toto vyhodnocení provedené ATCO bylo pro každou frázi zaznamenáno skrze Python aplikaci, ve které ATCO zaznamená své subjektivní hodnocení správnosti zpětného potvrzení. Každý subjekt (ATCO) takto postupně vydal dvacet vypsanych letových povolení a vyhodnotil správnost příslušného zpětného potvrzení. V průběhu celého experimentu pak na pozadí běží systém automatického rozpoznání řeči (ASR), který přepisuje vyslovené fráze do textu a následně porovnává, zda se shodují. Detailní průběh experimentu je znázorněn v obrázku č.4.

2.3. Vybrané fráze pro experiment

Pro účely analýzy úspěšnosti vyhodnocení zpětného potvrzení bylo vytvořeno dvacet párů frází, které postupně ATCO (subjekt) přečetl a pilot (předem připravená nahrávka autorem) zopakoval. Párové fráze, tedy letové povolení a zpětné potvrzení mohly být identické, či se od sebe v některých detailech odlišovat. Z těchto dvaceti párů frází obsahovalo dvanáct správné zpětné potvrzení, v osmi případech se naopak zpětné potvrzení od vydaného povolení lišilo. Jednotlivé fráze jsou rozepsány zde, žlutě jsou pak označeny deviate zpětného potvrzení od vydaného povolení ATCO:

1. Climb to Flight level 270.
Climb to Flight level 270.
2. Turn left heading 090.
Turn left heading 090.
3. Charlie Sierra Alpha 1923 increase speed 360 knots.
Charlie Sierra Alpha 1923 increase speed 360 knots.
4. QNH1014.
QNH1014.
5. Tango Victor Sierra 1448 descend 5000 feet, QNH 1012.
Tango Victor Sierra 1448 descend 5000 feet, QNH 10112.
6. Dynasty 62 climb to flight level 300, maintain speed 420 knots.
Dynasty 62 climb to flight level 300, maintain speed 420 knots.
7. Descend to flight level 160.
Descend to flight level 070.
8. Alpha November Alpha 976 turn right heading 180, maintain speed 250.

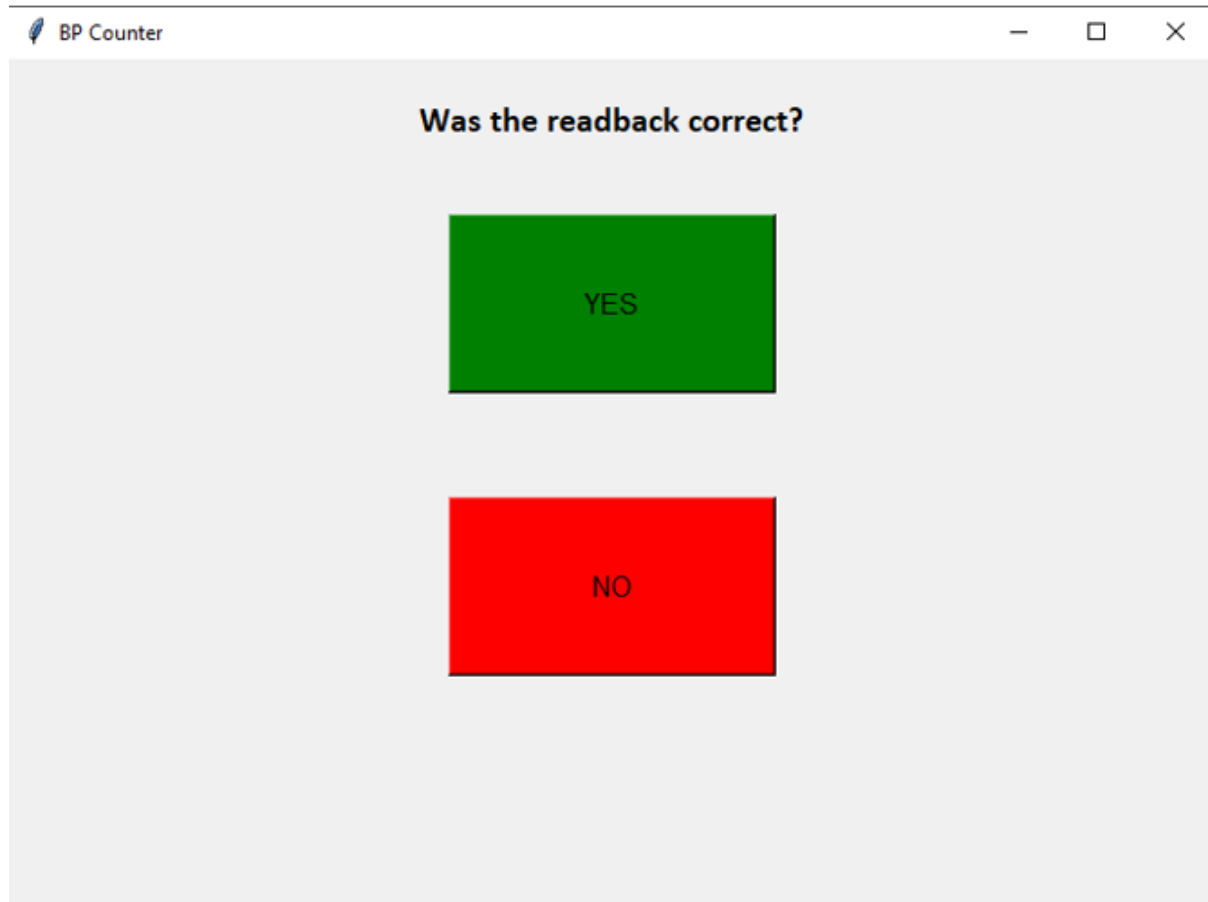


Alpha November Alpha 976 turn right heading 180, maintain speed 250.

9. Reduce speed 190 knots descend 5000 feet QNH 1009.
Reduce speed 200 knots descend 5000 feet QNH 1010.
10. Reduce speed 140.
Reduce speed 140.
11. Quebec Tango Romeo 555 continue heading 310, climb to flight level 310.
Quebec Tango Romeo 555 continue heading 310, climb to flight level 310.
12. Kilo Lima Mike 321 turn right heading 030, reduce speed 220 knots, descend 5000 ft, QNH 1011.
Kilo Lima Mike 321 turn right heading 010, reduce speed 220 knots, descend 4000 ft, QNH 1011.
13. Kilo Lima Mike 321 turn right heading 010, reduce speed 220 descend 4000 ft, QNH 1011.
Kilo Lima Mike 321 turn right heading 010, reduce speed 220 descend 4000 ft, QNH 1011.
14. Climb to flight level 220.
Climb to flight level 120.
15. Foxtrot India Lima 342 climb to flight level 290, increase speed 400 knots.
Foxtrot India Lima 342 climb to flight level 290, increase speed 410 knots.
16. Descend 5000 feet.
Descend 5000 feet.
17. Uniform Alpha Echo 777.
Uniform Alpha Echo 777.
18. Delta Yankee November 738 climb flight level 120 maintain heading 120 increase speed 300 knots.
Delta Yankee November 738 climb flight level 120 maintain heading 120 increase speed 280 knots.
19. Romeo Yankee November 738 climb to flight level 120 maintain heading 120 increase speed 480 knots.
Romeo Yankee November 738 climb to flight level 120 maintain heading 120 increase speed 280 knots.
20. Reduce speed 160.
Reduce speed 160

2.3. Aplikace pro záznam vyhodnocení správnosti zpětného potvrzení subjekty

Pro účely záznamu vyhodnocení správnosti zpětného potvrzení subjekty byla v programovacím jazyce Python vytvořena aplikace, obsahující dvě tlačítka – „Yes“ a „No“. Tato aplikace byla vytvořena s využitím modulu tkinter, jenž po spuštění skriptu otevře program v novém okně s modelovatelným uživatelským rozhraním. Po každém kliknutí na



Obrázek č.5: Aplikace pro záznam vyhodnocení správnosti jednotlivých read-backů subjektem

jedno z tlačítek se pomocí modulu csv exportují data do textového .csv souboru, ve kterém jsou posléze data po úpravě vyhodnocována. Po skončení každého měření je vhodné program vypnout a před novým měřením znovu zapnout, jelikož každé nové spuštění programu slouží jako oddělovač dat, kdy se na nový řádek vypíše fráze „new measurement“. Limitace tohoto programu spočívají v tom, že pokud subjekt provede takzvané dvojí kliknutí, kdy místo jednoho kliknutí na tlačítko subjekt klikne v rychlém sledu 2x po sobě, program zaregistruje obě kliknutí a tím se naruší integrita dat pro daný subjekt.

2.4. Program pro automatické rozpoznání řeči

Za účelem automatického vyhodnocení správnosti zpětného potvrzení byl v programovacím jazyce Python navržen a vytvořen program pro rozpoznání řeči. Skript využívá několika importovatelných modulů, které umožňují využívání různých funkcí, bez nutnosti je programovat od nuly.

2.4.1. Import Whisper

Whisper je model strojového učení vytvořený společností OpenAI pro účely automatického rozpoznávání řeči a její transkripce do textu. Tento model byl vydán v září roku 2022, tudíž se jedná o velmi mladou technologii. Model je postavený na architektuře a vztahu enkodér-dekodér. Enkodér přijímá vstup, který následně zpracuje a převede do kýženého formátu. Takto zpracovaný vstup poté vstupuje do dekodéru, kde je dekódován na nejvíce pravděpodobnou podobu textu, za pomoci neuronových sítí [36]. Model Whisper disponuje řadou různých verzí pod-modelů, ze kterých výběr modelu má vliv na přesnost transkripce, ale také na požadavky výpočetní techniky (místo na disku, výkon procesoru nebo grafické karty)

Size	Parameters	English-only model	Multilingual model	Required VRAM	Relative speed
tiny	39 M	<code>tiny.en</code>	<code>tiny</code>	~1 GB	~32x
base	74 M	<code>base.en</code>	<code>base</code>	~1 GB	~16x
small	244 M	<code>small.en</code>	<code>small</code>	~2 GB	~6x
medium	769 M	<code>medium.en</code>	<code>medium</code>	~5 GB	~2x
large	1550 M	N/A	<code>large</code>	~10 GB	1x

při zachování real-time rychlosti. Vzhledem k tomu, že dedikovaný počítač pro toto měření nedisponoval high-end komponenty, byla pro účely měření zvolena verze „base“, tedy základní typ modelu.

Obrázek č.6: Různé velikosti modelů Whisper a jejich požadavky na výpočetní techniku

2.4.2. Import `speech_recognition`

Tento modul zastává hlavní pozici pro příjem signálu z mikrofonu. Pro zajištění funkčnosti programu vytvořeného pro tento experiment byly vytvořeny 2 různé instance mikrofonů: mikrofon 1 = subjekt (řídící letového provozu), mikrofon 2 = autor (pilot) a to ve formě .wav souboru obsahující nahrávku zpětných potvrzení přenesené do mluvené formy. Hlavní překážkou ve vytvoření funkčního programu bylo přepínání mezi těmito vstupy. Nakonec byl



zvolen systém podobný klasickému „push to talk“, kdy se před požadovanou změnou aktivního vstupu na druhý mikrofon vždy 1x stiskne vybraná klávesa na klávesnici. Vlastnosti mikrofonů se dají různými způsoby optimalizovat, například lze upravit citlivost mikrofonu, minimalizování okolního šumu, vzorkovací frekvence vstupujícího signálu atp.

2.4.3. Import threading

Aby systém ASR mohl probíhat v reálném čase, bylo potřeba umožnit, aby paralelně s procesem monitorování a přepínání aktivního vstupu běželo také samotné přepisování mluveného slova do textu. Za tímto účelem byl importován modul threading, který umožňuje paralelní průběh funkcí v rozdílných vláknech. Pro program vytvořený pro tento experiment to znamená, že vlákno s funkcemi zabývající se analýzou vstupních dat a jejich následným přepisem si bere data z fronty, kterou zásobuje vlákno zabývající se získáváním zvukových dat. V situaci, kdy je fronta prázdná, tak vedlejší vlákno čeká 60 vteřin, zda se fronta nenaplní novými daty. Po uplynutí se toto vedlejší vlákno ukončí. To znamená, že vlákno vykonávající přepis mluveného slova do textu nemůže fungovat bez hlavního vlákna, které jej zásobuje daty.

2.4.4. Vyhodnocování stejnosti frází

Pro vyhodnocování správnosti zpětného potvrzení bylo nutné spárovat párové fráze a následně je porovnat. K zajištění správného spárování bylo provedeno indexování iterací, kdy první letové povolení má index 0 a první zpětné potvrzení index 1. Poslední letové povolení a poslední zpětné potvrzení mají index 38, respektive 39. Vzhledem k tomu, že po každém povolení následovalo zpětné potvrzení, bylo možné jednotlivá povolení a zpětná potvrzení rozdělit pomocí funkce modulo 2. Modulo 2 je funkce, která přiřazený index dělí číslem 2 a vypíše zbytek po vydělení, tzn. např. pro index 2 funkce modulo 2 vypíše výsledek 0, naopak pro číslo 3 by byl výsledek 1, jelikož po dělení zůstane právě zbytek 1. Sudé indexy a počáteční index 0 náleží vstupu ATCO, lichá čísla pak náleží vstupu pilota. Pro zvýšení úspěšnosti vyhodnocování správného zpětného potvrzení byla čísla izolována od textu, a pro účely tohoto experimentu byla vyhodnocována pouze stejnost číselných hodnot obsažených v letových povoleních / zpětném potvrzení.

Postup programu při analýze, vypsání a porovnání textu je následující, viz [příloha 1]:

- Spuštění „while“ cyklu, index (idx) = 0
- Analýza a následné vypsání prvního textu, idx = 0.



- Index se zvedne o 1, $idx = 1$
- Podmínka „if / else“, zda je hodnota indexu sudá / lichá (momentálně činí hodnota indexu 1), text se přepíše do proměnné „prev text“.
- Analýza a následné vypsání druhého textu, $idx = 1$.
- Index se zvedne o 1, $idx = 2$
- Podmínka „if / else“, zda je hodnota indexu sudá / lichá (momentálně činí hodnota indexu 2), hodnota je tedy lichá, text se přepíše do proměnné „text“.
- Probíhá izolace slov a písmen z textů za využití funkce „re“ do čtyř separátních proměnných: „prev_not_numbers“, „prev_numbers“, „not_numbers“, „numbers“ dle toho, zda text náleží proměnné „prev_text“, či „text“. Dále je vytisknuta hodnota „True / False“ na základě toho, zda se proměnná „prev_numbers“ rovná hodnotě proměnné „numbers“, tedy zda jsou číselné hodnoty v těchto dvou frázích identické. To samé ve stejný čas probíhá i pro proměnné obsahující izolovaná slova.

Takto algoritmus pro vyhodnocení frází pokračuje až do hodnoty indexu $idx = 40$.

2.4.5. Dotazník demografických faktorů

Za účelem sledování demografických vlivů na úspěšnost systému ASR byl vytvořen krátký dotazník se třemi otázkami. První otázka se ptala subjektu na pohlaví. Pro odpověď bylo možné zvolit jednu ze dvou variant – muž / žena. Takto omezený soubor odpovědí byl zvolen z důvodu získání nejkonzistentnějšího rozdělení úrovně výšky tónu hlasu. Další otázka se subjektu ptala na mateřský jazyk. Každý jazyk má svá specifika, která se následně přenášejí do mluvené formy, což ve výsledku ovlivňuje přízvuk a výslovnost při komunikaci v cizím jazyce. Cílem tedy bylo zjistit, zda se tyto deviace jazyka projeví na přesnosti systému ASR. Poslední otázka se dotazovala subjektu na jeho úroveň anglického jazyka dle mezinárodní stupnice (A1, A2, B1, B2, C1, C2). Rozdílné úrovně AJ by mohly mít vliv na správnost výslovnosti slov, či rychlost čtení. Tato otázka tedy slouží jako doplněk k otázce na mateřský jazyk.

2.5. Zpracování dat

Po sesbírání veškerých dat vyhodnocení správnosti zpětného potvrzení vytvořeným softwarem, dat správnosti zpětného potvrzení jednotlivými subjekty a dat z dotazníků následovalo vyhodnocení jednotlivých částí. Správnost jednotlivých prepisů programu ASR a jeho vyhodnocení správnosti zpětného potvrzení bylo stanoveno na základě porovnání s originálními transkripty použitých frází. Jak již bylo zmíněno v kapitole 2.4., pro účely tohoto



experimentu se z dat získaných systémem ASR vyhodnocovaly jen číselné hodnoty při detekci nesprávného zpětného potvrzení. Důvod pro toto rozhodnutí se vztahuje již k návrhu použitých frází, jenž se v párech liší právě jen v číselných hodnotách, a to z důvodu zjednodušení procesu zpracování dat. Jednou z komplikací při zpracování dat ASR systému se prokázala takzvaná falešná negativita, kdy daný pár frází byl vyhodnocen, jakožto odlišný, ale ne z důvodu správného přepsání mluvených slov. Tento problém se vyskytoval v situacích, kdy dle originálního přepisu frází bylo vráceno nesprávné zpětné potvrzení. Systém tedy správně vyhodnotil nesprávnost zpětného potvrzení, nicméně po analýze daného transkriptu vyšlo najevo, že došlo k chybné transkripci. Další překážkou pak byly situace, kdy přepis čísel ve frázi odpovídal správným hodnotám, nicméně v rozdílném formátu, např. QNH 1014 bylo přepsáno na QNH 1,0,1,4. V takových případech systém vyhodnotil, že zpětné potvrzení nebylo správné, i když kontext zprávy správný byl. V takovýchto situacích byl zápis výsledku vyhodnocení přepsán podle toho, zda kontext obsahu vyhodnocovaných vstupů byl totožný. Data byla následně přenesena z .csv souboru do tabulky v programu Excel, ve které bylo dvacet sloupců, každý reprezentující jeden pár frází. V řádcích (1 až 50) pak byla data vyhodnocení jednotlivých subjektů, k nimž se přidala také data z formuláře vytvořeného pro toto měření na portálu Google Forms. Tato data každému subjektu přiřazovala pohlaví, mateřský jazyk a úroveň anglického jazyka. V tabulce byly do každé buňky doplněny hodnoty dle toho, zda systém vyhodnotil, že zpětné potvrzení odpovídalo vydanému povolení (do tabulky doplněna 1). V případě, že systém vyhodnotil, že se tyto fráze neshodují, byla do tabulky zapsána 0. Následně bylo porovnáno, zda v příslušné buňce tabulky byla zapsána správná hodnota. To bylo provedeno pomocí porovnání dodané hodnoty (0 nebo 1) s hodnotou, jenž byla pro danou frázi správná. Jednotlivé fráze (1 až 20) nabývaly hodnot 0 a 1 dle toho, zda ve scénáři měly správné zpětné potvrzení či nikoliv. Tedy, v případě, že se letové povolení shodovalo se zpětným potvrzením, je správná hodnota 1. V případě, že zpětné potvrzení neodpovídalo vydanému povolení, hodnota je 0. Následně bylo uplatněno podmíněné formátování, kdy byla buňka zvýrazněna pokud se hodnota v buňce nerovnála správné hodnotě pro danou frázi. Pro získání počtu zvýrazněných buněk byla použita makro funkce „get.cell“, která slouží k identifikaci zvýrazněných buněk. To znamená, že vedle původní tabulky vznikla zcela nová tabulka, ve které buňky zvýrazněné v původní tabulce byly číselně odlišeny od zbytku. Pro získání počtu nesprávných vyhodnocení systémem ASR byly pak v každém sloupci a řádku takto označené buňky spočítány pomocí funkce „countif“. Pomocná tabulka byla následně skryta a výsledek této funkce byl pak prezentován jako součást původní tabulky.



V každém řádku se také hodnotila celková (0 až 20) a procentuální (0 až 100 %) úspěšnost programu ASR ve vyhodnocování letového povolení a jeho zpětného potvrzení jednotlivých subjektů. Po zapsání dat všech uživatelů se dále vypočítala celková (0 až 50) a procentuální (0 až 100 %) úspěšnost vyhodnocení dle jednotlivých frází do odpovídajícího sloupce dané fráze. U každé fráze bylo také zaznamenáno, z kolika slov se skládá, což bylo následně použito pro grafické znázornění vztahu délky fráze s úspěšností jejího vyhodnocení.

Pro zpracování dat vyhodnocení správnosti zpětného potvrzení subjekty byla data obsažená textovém .csv souboru přenesena do souboru Excel, kde byla vepsána do tabulky, ve které bylo 20 sloupců, každý reprezentující jeden pár frází. V řádcích (1 až 50) pak byla data vyhodnocení jednotlivých subjektů. V každém řádku se také hodnotila celková (0 až 20) a procentuální (0 až 100 %) úspěšnost subjektu ve vyhodnocování letového povolení a jeho zpětného potvrzení jednotlivých subjektů. Po zapsání dat všech uživatelů se také vypočítala celková (0 až 50) a procentuální (0 až 100 %) úspěšnost vyhodnocení dle jednotlivých frází do odpovídajícího sloupce. Pro získání počtu chybných vyhodnocení subjekty byly použity stejné funkce jako pro výše popsanou tabulku vyhodnocení úspěšnosti systému ASR

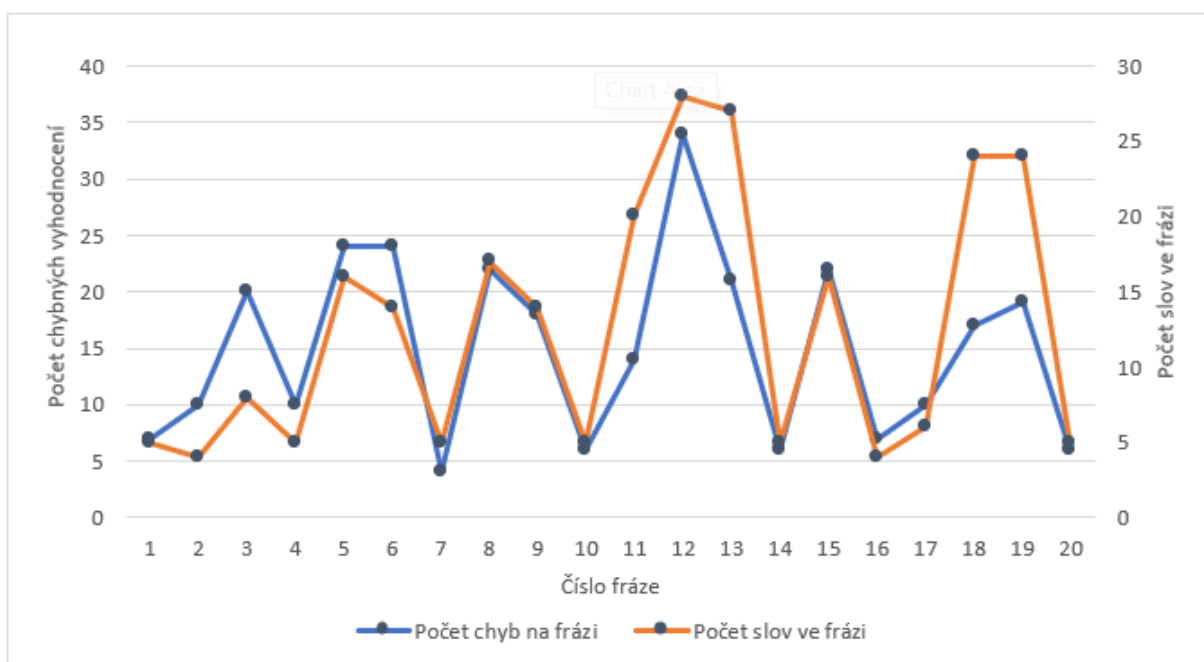
Odpovědi subjektů na jednotlivé otázky dotazníku byly v programu Excel filtrovány dle kategorií (pohlaví, mateřský jazyk a úroveň anglického jazyka). Tyto demografické znaky získané z dotazníku byly vypsány ke každému řádku měření. Pomocí formátování a filtrování pak bylo možné tyto informace analyzovat. K filtrování byla použita funkce „průřez“. Následně byla získána úspěšnost programu ASR ve vyhodnocení správnosti zpětného potvrzení dle jednotlivých demografických znaků.

V programu Scilab se z dat získaných programem Excel vyhodnotil korelační koeficient mezi délkou jednotlivých povolení a úspěšností jejich správného vyhodnocení zpětného potvrzení. Toho bylo dosaženo přenesením dat z tabulky reprezentující počet chyb (0 až 50) ve vyhodnocení pro každé z dvaceti povolení do programu Scilab. Těchto dvacet hodnot bylo přiřazeno do vektoru x. Do vektoru y následně byla uložena data počtu slov obsažených v jednotlivých povoleních. Výstupem kalkulace je Pearsonův korelační koeficient, tedy ukazatel lineárního vztahu mezi těmito dvěma proměnnými.

3. Výsledky

3.1. Korelace počtu slov v letovém povolení s celkovým počtem chybných vyhodnocení správnosti zpětného potvrzení systémem ASR

Hodnota Pearsonova korelačního koeficientu $r = 0,792$ představuje hodnotu na rozmezí silné a velmi silné pozitivní korelace. Tedy, délka jednotlivých frází má silný, až velmi silný vliv na úspěšnost vyhodnocení správnosti zpětného potvrzení systémem ASR. Tato silná korelace je také velmi dobře viditelná na grafu č.1, kde linie pro obě proměnné viditelně kopírují podobné trendy. Z grafu je patrné, že systém ASR měl nejvíce problémů se správným vyhodnocením správnosti letového povolení a zpětného potvrzení č.12., jenž zároveň byly nejkompexnějšími z celého souboru. Zajímavé je, že fráze č.13 obsahovala pouze o jedno slovo méně než fráze č.12, a kromě toho se tyto fráze lišily jen drobnými změnami v hodnotách obsažených číselných informací. Přes to je možné pozorovat viditelný rozdíl v úspěšnosti těchto dvou souborů frází (34 chyb z 50 u páru č. 12 oproti 21 chybám z 50 u páru č.13) viz tabulka č.1. Rozdíl v úspěšnosti transkripce a vyhodnocení pro tyto dva páry frází činí 36 %. Nabízí se tedy hypotéza, že významný vliv na úspěšnost transkripce a vyhodnocení správnosti zpětného potvrzení má také obsah jednotlivých povolení / zpětných potvrzení.



Obrázek č.7: grafické znázornění vztahu počtu slov ve frázi s počtem celkových chyb vyhodnocení správnosti zpětného potvrzení systémem ASR



Tabulka č.1: Přehled počtu slov v jednotlivých frázích, celkového počtu chyb ve vyhodnocení správnosti zpětného potvrzení a procentuální úspěšnosti programu ASR ve vyhodnocování jednotlivých frází

Počet chyb na frázi (-)	Procentuální úspěšnost (%)	Počet slov ve frázi (-)
7	86	5
10	80	4
20	60	8
10	80	5
24	52	16
24	52	14
4	92	5
22	56	17
18	64	14
6	88	5
14	72	20
34	32	28
21	58	27
6	88	5
22	56	16
7	86	4
10	80	6
17	66	24
19	62	24
6	88	5

3.2. Úspěšnost subjektů ve vyhodnocení správnosti zpětného potvrzení

V kontextu výsledků vyhodnocení správnosti stejnosti zpětného potvrzení jednotlivými subjekty lze z vyhodnocených dat vyčíst, že z celkového počtu padesáti subjektů, bylo pouze dvaadvacet schopno identifikovat veškerá chybná zpětná potvrzení, viz obrázek č.9. Pokud bychom k tomuto výsledku přičetli i subjekty, které nesprávným způsobem vyhodnotili pouze zpětné potvrzení č.5, které se dle grafu č.9 jeví jako nejproblematičtější, počet stoprocentně úspěšných subjektů by vzrostl na pětatřicet.

Fráze č. 5:

Tango Victor Sierra 1448 descend 5000 feet, QNH 1012.

Tango Victor Sierra 1448 descend 5000 feet, QNH 10112.

Složitostí v rozpoznání rozdílné hodnoty obsažené v tomto zpětném potvrzení mohl být způsob, jakým byla chybná číselná hodnota prezentována: One-Zero-One-Twelve namísto One-Zero-One-One-Two.

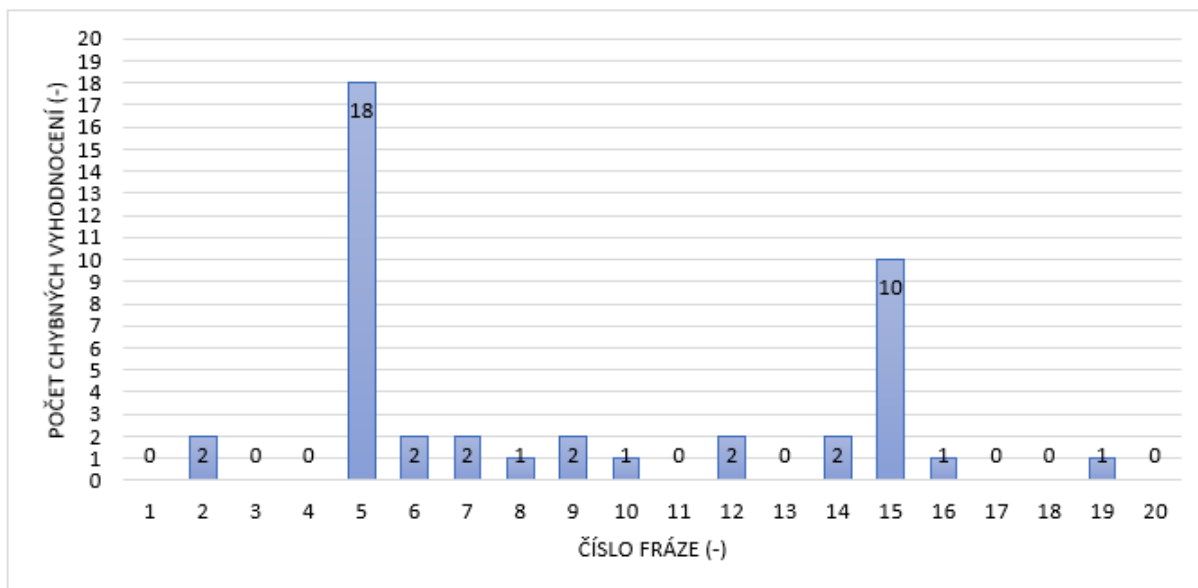
Průměrná úspěšnost jednotlivých subjektů činí 95,6 %, což představuje poměrně vysokou míru úspěšnosti identifikace nesprávného zpětného potvrzení. Pokud bychom z výsledků odebrali data fráze č.5, úspěšnost by dosahovala 97,4 %. Tyto hodnoty by však stále mohly potenciálně vést k velkému množství konfliktních situací v reálném provozu.

Kromě fráze č.5 ze získaných dat vyčnívá fráze č.15. Vysvětlení hypotéz, proč zrovna tyto fráze činili subjektům největší problémy jsou rozebrány v diskusi.

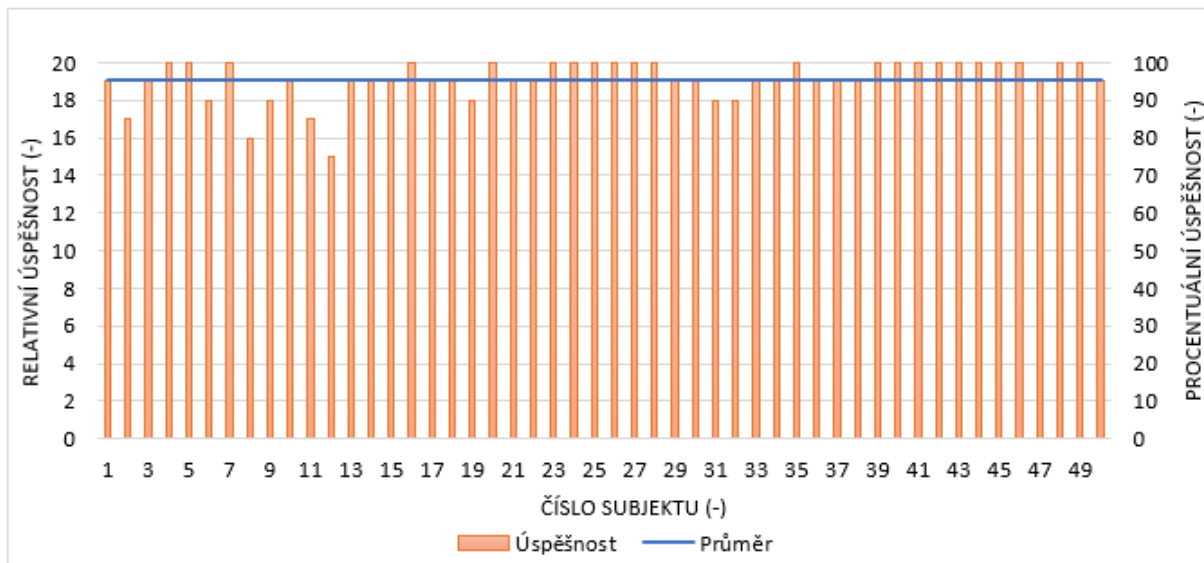
Fráze č.15:

Foxtrot India Lima 342 climb to flight level 290, increase speed 400 knots.

Foxtrot India Lima 342 climb to flight level 290, increase speed 410 knots.



Obrázek č.8: Celkový počet chyb subjektů ve vyhodnocování správnosti zpětného vyhodnocení



Obrázek č.9: Relativní a procentuální úspěšnost jednotlivých subjektů ve vyhodnocování správnosti zpětného potvrzení

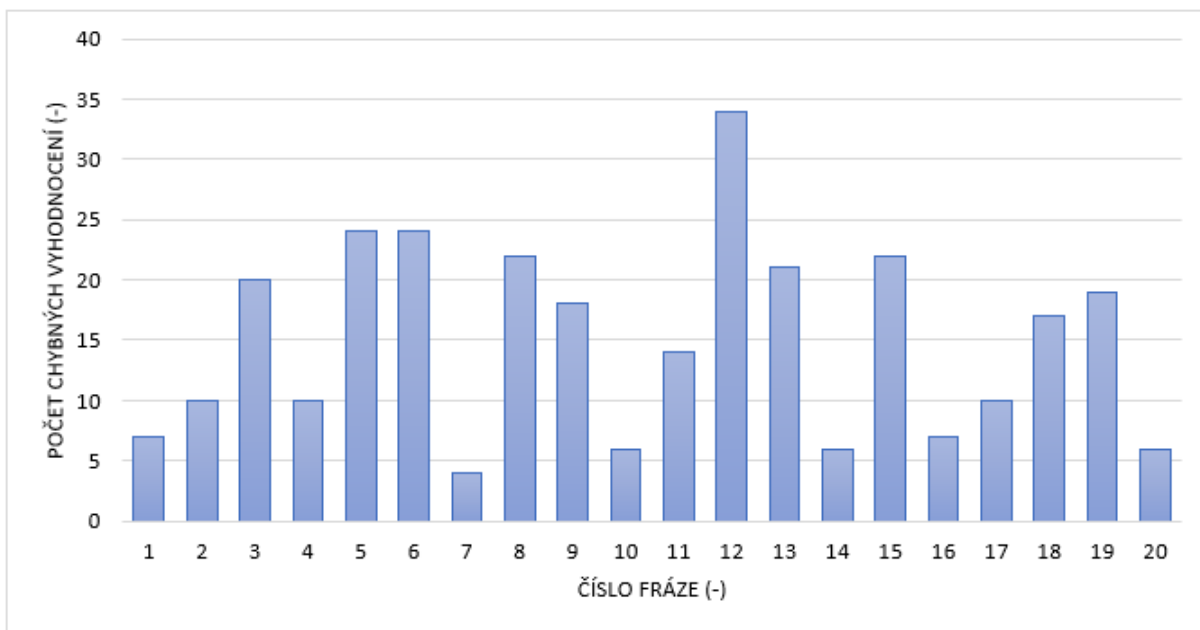
3.3. Úspěšnost systému ASR ve vyhodnocování správnosti zpětného potvrzení

Z výsledných dat úspěšnosti vyhodnocení systémem ASR vyplývá úspěšnost rozeznání správnosti zpětného potvrzení pro padesát různých subjektů 69,9 %. Jak je čitelné z obrázku č. 11, nejvyšší dosažená úspěšnost činila ve dvou případech 90 %. Pouze u čtrnácti z padesáti subjektů úspěšnost vyhodnocení systému ASR dosáhla alespoň 80 %.

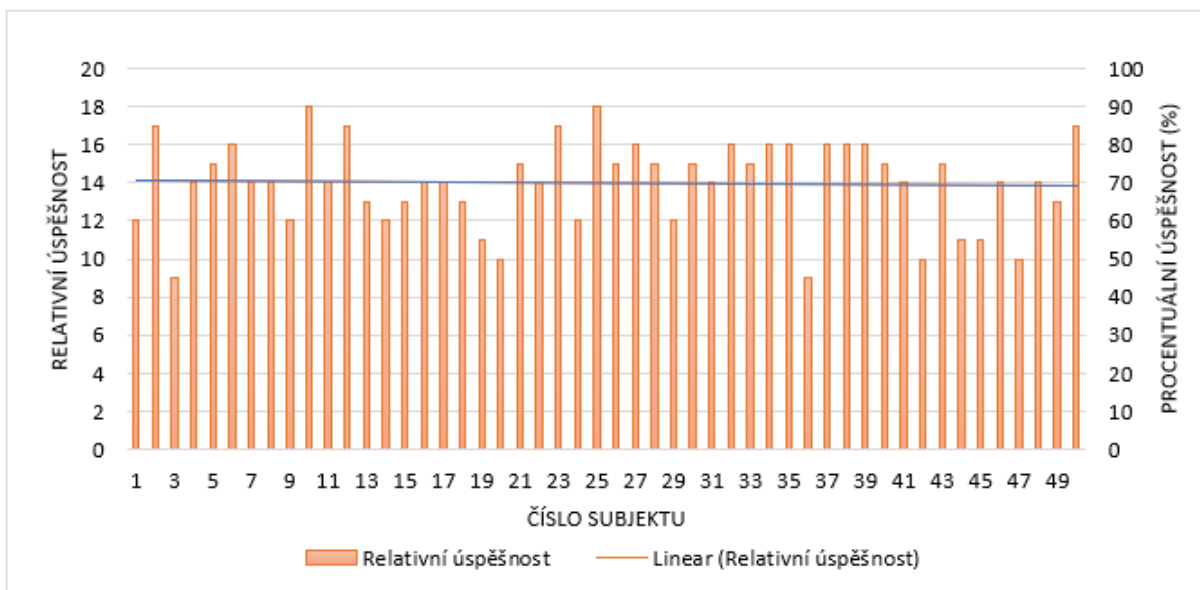
Co se týče jednotlivých frází, je z obrázku č.10 patrné, že některé fráze systémem ASR pravidelně vyhodnocoval úspěšněji než ostatní. Nejhůře z tohoto porovnání vychází fráze č.12, která byla vyhodnocena správně pouze v šestnácti z padesáti případů. Jak je viditelné v kapitole 3.1., jedná se o vůbec nejdelší frázi ze souboru. Fráze č. 12:

Kilo Lima Mike 321 turn right heading 030, reduce speed 220 knots, descend 5000 ft, QNH 1011.

Kilo Lima Mike 321 turn right heading 010, reduce speed 220 knots, descend 4000 ft, QNH 1011.



Obrázek č.10: Relativní a procentuální úspěšnost ASR ve vyhodnocování správnosti zpětného potvrzení u jednotlivých subjektů



Obrázek č.11: Celkový počet chyb systému ASR ve vyhodnocování jednotlivých frází

Veškeré procentuální úspěšnosti systému ASR jsou v tabulce č.2 rozděleny společně s počtem zástupců jednotlivých kategorií. Díky tomu lze zhodnotit, zda by tyto výsledky mohly být využity pro účely statistické analýzy vlivu demografických faktorů na úspěšnost systému



ASR. Z tabulky je patrné, že se experimentu zúčastnila skupina subjektů vykazující nízkou variací demografických proměnných. Rozdělení subjektů dle pohlaví nepřináší do analýzy výsledků žádnou relevantní informaci o vlivu tohoto demografického znaku na úspěšnost systému ASR. Důvodem je viditelně nízká účast příslušníků ženského pohlaví na experimentu. Další otázka se týkala rodného jazyka účastníků. Osmatřicet z padesáti subjektů uvedlo jako rodný jazyk češtinu, což představuje 78 % všech účastníků. Druhou nejpočetnější skupinu představoval slovenský jazyk, jenž čítá sedm subjektů. Zbytek odpovědí byl spojen do jednotlivé kategorie ostatní, jelikož žádná z odpovědí nebyla zaznamenána více než jednou. Pod ostatní spadají mateřské jazyky angličtina, němčina, norština, ruština a srbština. Vzhledem k rozložení dat pro tento demografický znak, pro porovnání vlivu mateřského jazyka nebyl získán dostatečný počet mateřských jazyků jiných než čeština. Poslední otázka se subjektů dotazovala na jejich úroveň anglického jazyka. Jednačtyřicet z padesáti subjektů uvedlo možnost pokročilý (B2) nebo velmi pokročilý (C1) pro jejich úroveň anglického jazyka. Vzhledem k tomu, že tyto 2 úrovně jsou si poměrně blízké a rozdíl v úspěšnosti systému ASR pro ně činí 1,19 %, lze považovat vliv rozdílu těchto úrovní jazyka za zanedbatelný. Pro analýzu vlivu ostatních úrovní nebyl získán potřebný počet dat.

Tabulka č.2: Přehled demografických proměnných a jejich vliv na procentuální úspěšnost vyhodnocení správnosti zpětného potvrzení systémem ASR

Úspěšnost dle demografických proměnných		
Dle pohlaví	Počet	Úspěšnost (%)
muž	47	67,79
žena	3	71,67
Dle rodného jazyka		
čeština	38	69,21
slovenština	7	72,14
ostatní	5	72,00
Dle úrovně AJ		
pokročilý začátečník (A2)	3	76,67
mírně pokročilý (B1)	5	69,72
pokročilý (B2)	23	69,69
velmi pokročilý (C1)	18	70,98
rodilý mluvčí	1	75,00



4. Diskuse

Výsledky provedeného experimentu přináší příslib, že technologie automatického rozpoznání řeči (ASR) může být v budoucnu významným přínosem pro zvýšení bezpečnosti letového provozu. I s omezenými finančními zdroji v rámci závěrečné práce lze dosáhnout vzhledem k rychlému pokroku výpočetní techniky a rozvoji umělé inteligence obstojné úspěšnosti přepisu. Program dosáhl bez jakéhokoliv modelování do letecké komunikace průměrné úspěšnosti rozpoznání nesprávného zpětného potvrzení (read-backu) 69,9 %. V porovnání s průměrnou hodnotou (95,6 %) vyhodnocení správnosti zpětného potvrzení člověkem je tato úspěšnost stále nízká. Na základě výsledků studií zmíněných v první části této práce však lze očekávat, že po vytvoření specifického modelu pro leteckou komunikaci by úspěšnost systému mohla dosáhnout podobných hodnot. Pro hodnotu úspěšnosti rozpoznání zpětného potvrzení uživateli je nutné zohlednit, že žádný ze subjektů nebyl certifikovaný řídicí letového provozu (ATCO), tudíž lze očekávat, že tato hodnota by v takovém případě byla vyšší než získaných 95,6 %.

Dotazník demografických znaků splnil svůj účel pro rozdělení dat dle určitých kvalit, nicméně se projevilo, že se experimentu zúčastnilo jen padesát subjektů. Nejvíce je tento fakt znát v počtu žen, které experiment absolvovaly, a to pouze tři, oproti sedmačtyřiceti mužům. Podobná situace se vyskytla také u mateřského jazyka, kde pouze dvanáct účastníků uvedlo rodný jazyk jiný než český. U otázky na úroveň angličtiny pak jednačtyřicet z padesáti subjektů uvedlo úroveň B2, či C1. Ty představují dvě po sobě jdoucí úrovně angličtiny, tudíž mezi nimi není dramatický rozdíl. Průměrný subjekt účastnící se experimentu byl tedy muž s češtinou jako mateřským jazykem a angličtinou na úrovni B2 / C1. Z tohoto důvodu jsou data poskytnutá dotazníkem nedostatečná pro vyhodnocení, zda nějakým způsobem ovlivňují úspěšnost systému ASR. Jedním z pozorovaných potenciálních vlivů mateřského jazyka na výsledky měření je častý výskyt cizojazyčných slov, či dokonce znaků z různých abeced. Tyto situace pravidelně nastávaly i přes to, že subjekt mluvil anglickým jazykem. Až v průběhu měření byla zjištěna možnost využití pouze anglického modelu namísto vícejazyčného (viz obrázek č.6), který by mohl úspěšnost v transkripci mluveného slova zvýšit.

Pro účely využití ve fázi výcviku řídicích letového provozu (ATCO) není vhodné, že systém nedokáže rozpoznat, zda subjekt vyslovil číslovky dle frazeologie, např. 110 = „One-One-Zero“ a ne „One-Hundred-Ten“. To by mohlo mít za následky nežádoucí návyky ATCO ve výcviku, které by se mohly přenést do živého provozu. Další limitací je pak falešná negativita, kdy



systém vyhodnotí správné zpětné potvrzení jako nesprávné, čímž ATCO ztrácí drahocenný čas, jelikož je nucen nesprávnou informací zkontrolovat, případně opravit.

Velký vliv na úspěšnost systému ASR měla délka vydaných letových povolení. Pearsonův korelační koeficient o hodnotě 0,792 představuje hodnotu na pomezí silné až velmi silné korelace. Jedním z faktorů, proč délka fráze mohla mít vliv na úspěšnost ASR systému může být, že během delší doby snímání zvuku mikrofonu, mohl okolní ruch narušit čistotu přijímaného signálu. Další faktor, který má vliv na úspěšnost ASR systému je skladba slov letového povolení. To je například velmi dobře pozorovatelné na úspěšnosti systému ve vyhodnocování frází č.12 a č.13., kdy fráze č.12 byla ve čtyřiatřiceti případech z padesáti vyhodnocena nesprávně, fráze č. 13 však pouze v jednadvaceti případech z padesáti.

Fráze č.12:

Kilo Lima Mike 321 turn right heading 030, reduce speed 220 knots, descend 5000 ft, QNH 1011.

Kilo Lima Mike 321 turn right heading 010, reduce speed 220 knots, descend 4000 ft, QNH 1011.

Fráze č.13:

Kilo Lima Mike 321 turn right heading 010, reduce speed 220 descend 4000 ft, QNH 1011.

Kilo Lima Mike 321 turn right heading 010, reduce speed 220 descend 4000 ft, QNH 1011.

Jak je patrné, tyto fráze se navzájem liší jen ve třech slovech, přesto je u nich výsledná úspěšnost vyhodnocení systémem ASR rozdílná o 36 %. V průběhu experimentu se ukázalo, že některá slova představují pro systém rozpoznání řeči větší problém než ostatní. V tomto případě zde byla poměrně problematická číslovka „3“ (vyslovováno „tree“), a dále pak slovo „knots“ následující za číslovkou „0“ („zero“).

Nejčastější chyby pozorované v přepisu zaznamenaným systémem ASR:

Přidání extra nuly na konci čísla, zejména pokud následovalo slovo „knots“, např. z „220 knots“ vzniklo „2200 not“.

Po číslovce „9“ („niner“) často vypsáno „or“.

Místo číslovky „2“ často vypsáno rozmezí (2-0 místo 220, tedy něco ve smyslu 2 až 0).

Časté záměny „2“ a „to“.

Nekonzistentní vypisování čísel, kdy jednotlivé číslovky byly odděleny mezerou, či čárkou.



Místo „increase“ systém vypíše „in 3“.

U subjektu s němčinou jako mateřským jazykem bylo místo číslovky „9“ vypisováno slovo „einer“.

„5th level“ místo „flight level“.

Místo „foxtrot“ často „4th std“.

Návrh experimentu byl pro uživatele relativně složitý a bylo nutné jej pro zajištění konzistentnosti dat upravit. Pro snížení zátěže subjektů autor převzal úlohu manuálního přepínání aktivního vstupu mikrofonu. Nutnost přepínání mezi jednotlivými mikrofony pramenila z použití audio souboru, jakožto zdroje zpětných potvrzení. Jelikož audio soubor představoval kontinuální zvukový signál, vytvořený systém ASR nebyl schopen vyhodnotit kdy přepnout aktivní mikrofon zpět na vstup testovaného subjektu. Systém byl také velmi náchylný na zvýšenou hladinu okolního zvuku, a to i přes implementaci různých funkcí pro jeho potlačení. Roli v úspěšnosti systému může hrát například i kvalita použitého mikrofonu, či zvukové karty instalované v počítači. Tyto vlivy by bylo vhodné dále zkoumat v případných dalších studiích zabývajících se touto problematikou.

Na tento experiment a zejména pak na zkoumání možnosti využití vytvořeného systému ASR se do budoucna nabízí výzkumný projekt, jenž by se zabýval modelováním systému pro použití specificky v leteckém prostředí. Pokud by se takovýto model podařilo (třeba na Fakultě dopravní ČVUT) vytvořit, otevřelo by se množství oblastí pro výzkum potenciální implementace systému ASR v živém provozu. Jednou z dalších výzev, alespoň pro autora této práce, je zajištění kontinuálního snímání hlasových vstupů na frekvenci – tedy od ATCO a pilotů. Výsledky provedeného experimentu přináší příslib, že technologie automatického rozpoznání řeči (ASR) může být v budoucnu významným přínosem pro zvýšení bezpečnosti letového provozu - to najzajímavější na tom je to, že je to možné dokázat s omezenými náklady vrámci závěrečné práce



5. Závěr

Bakalářská práce je zaměřena na vytvoření a následné ověření funkčnosti systému pro automatické rozpoznání řeči, díky kterému by bylo možné detekovat nesprávné zpětné potvrzení. Pro účely práce byl navržen koncept experimentu, který byl následně během měsíců října a listopadu vykonán. Experiment se odvíjel od testování funkčnosti vytvořeného systému společně s vyhodnocením vlivu demografických faktorů (získaných formou dotazníku) na jeho úspěšnost. Do experimentu se zapojilo padesát subjektů, jenž byli vyzváni k vydání dvaceti letových povolení vytvořených pro tyto účely a následnému ověření správnosti zpětného potvrzení. Zpětné potvrzení bylo předem nahráno autorem této práce do audio souboru a po každém vydaném letovém povolení subjektem přehráno. Jednotlivá vyhodnocení správnosti zpětných potvrzení subjekty byla zaznamenána autorem vytvořeným programem. Po nasbírání veškerých potřebných dat do textových souborů .csv, byla tato data přenesena do aplikace Excel a následně dána do tabulkové podoby. Proběhla analýza úspěšnosti systému ASR pro každý jeden subjekt, ze které byla získána celková procentuální úspěšnost vyhodnocení zpětného potvrzení systémem ASR. Tento výsledek se ukázal jako nedostatečný pro možnost uvedení systému do živého provozu letecké dopravy.

Statisticky významným ukazatelem přesnosti vyhodnocení správnosti zpětného potvrzení systémem ASR se ukázala být délka jednotlivých vydaných letových povolení. Tato korelace byla vyjádřena pomocí Pearsonova korelačního koeficientu.

Vytvořený systém ASR je v současné iteraci limitován nespecifičností použitého modelu. Pro transkripci nejpravděpodobnějšího textu specifického pro hlasovou komunikaci v leteckém provozu je nutné vytvořit unikátní model zohledňující leteckou frazeologii.

Navzdory tomu systémy ASR představují do budoucna příslib zvýšení efektivity práce ATCO, a to ať už detekcí nesprávného zpětného potvrzení, ale také například integrací různých systémů, které pak automaticky zajišťují přepis informací z hlasové komunikace do leteckých stripů. Pokud tedy budou pro systémy ASR vytvořeny modely specifické pro leteckou komunikaci, je jen otázkou času, kdy se tyto systémy přenesou z testovacích místností do živého provozu a budou sloužit jako relevantní nástroj pro zvýšení efektivity, ale hlavně bezpečnosti letecké dopravy.



Seznam použité literatury

- [1] EUROCONTROL. *Controller-pilot datalink communications at our Maastrich UAC*. Online. Dostupné z: <https://www.eurocontrol.int/service/controller-pilot-datalink-communications-our-maastricht-uac>.
- [2] PINSKA-CHAUVIN, Ella; HELMKE, Hartmut; DOKIC, Jelena; HARTIKAINEN, Petri; OHNEISER, Oliver et al. Ensuring Safety for Artificial-Intelligence-Based Automatic Speech Recognition in Air Traffic Control Environment. Online. *Aerospace*. 2023, roč. 10, č. 11. ISSN 2226-4310. Dostupné z: <https://doi.org/10.3390/aerospace10110941>. [cit. 2023-11-22].
- [3] HELMKE, Hartmut; ONDŘEJ, Karel; SHETTY, Shruthi a ARILÍUSSON, Hörður. *Readback Error Detection by Automatic Speech Recognition and Understanding: Results of HAAWAll project for Isavia's Enroute Airspace*. 2022.
- [4] EUROCONTROL. *Comprehensive Assessment: European Aviation*. XXVII, 2022. 2022. Dostupné také z: <https://www.eurocontrol.int/sites/default/files/2022-07/eurocontrol-comprehensive-air-traffic-assessment-20220707.pdf>.
- [5] REASON, James. The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* [online]. 1990, **327**(1241), 475-484 [cit. 2023-04-10]. ISSN 0080-4622. Dostupné z: doi:10.1098/rstb.1990.0090
- [6] *Machine Learning of Speech Recognition Models for Controller Assistance: MALORCA*. Project Results. 2018. Dostupné také z: <https://cordis.europa.eu/project/id/698824/results>.
- [7] INTERNATIONAL CIVIL AVIATION ORGANIZATION. Annex 10, Volume II, *Aeronautical Telecommunications*. Sixth Edition. 2001. Dostupné také z: https://www.icao.int/Meetings/anconf12/Document%20Archive/AN10_V2_cons%5B1%5D.pdf.
- [8] INTERNATIONAL CIVIL AVIATION ORGANIZATION. *Status of English Language Standard for Use in Civil Aviation*. I. 2003. Dostupné také z: <https://www.icao.int/sam/documents/2003/raac8/raac8ip18.pdf>.
- [9] SKUHROVÁ, Aneta. *Vliv intonace a výslovnosti na porozumění v radiokomunikaci*. Online, vedoucí Ing. Lenka Hanáková, Ph.D. Praha: České vysoké učení technické v Praze, 2023.



- [10] AIR NAVIGATION SERVICES OF THE CR. *VFR Manual: VFR-ENR-6: Phraseology*. 15 JUN 23. 2015. Dostupné také z: https://aim.rlp.cz/vfrmanual/actual/enr_6_en.html.
- [11] INTERNATIONAL CIVIL AVIATION ORGANIZATION. *Annex 1 To The Convention on International Civil Aviation: Personnel Licensing*. XII. 2018. Dostupné také z: https://www.icao.int/APAC/Meetings/2019%20COSCAP%20SEAEASA%20PEL/AN01_cons.2019_compressed.pdf.
- [12] MCMILLAN, David. *Miscommunications in Air Traffic Control*. Online, Master's Thesis. Queensland: Queensland University of Technology, 1998. Dostupné z: <http://www.aerolingo.com/docs/Miscommunications%20in%20Air%20Traffic%20Control.pdf>.
- [13] INTERNATIONAL CIVIL AVIATION ORGANIZATION. Annex 11, *Air Traffic Services*. XIII. 2001. Dostupné také z: <https://skyrise.aero/wp-content/uploads/2017/03/ICAO-Annex-11-Air-traffic-services.pdf>.
- [14] ŠVEC, Jan. *Studium mechanicko-akustických vlastností zdroje lidského hlasu*. Dizertační práce. Olomouc: Univerzita Palackého v Olomouci, 1996.
- [15] CATFORD, J. C. *A practical introduction to phonetics*. 2nd edition. Oxford: Clarendon, 1992. ISBN 01-982-4217-4.
- [16] REUTER, Christoph. The role of formant positions and micro-modulations in blending and partial masking of musical instruments. Online. *The Journal of the Acoustical Society of America*. 2009, roč. 126, č. 4_Supplement, s. 2237-2237. ISSN 0001-4966. Dostupné z: <https://doi.org/10.1121/1.3249188>.
- [17] THE EDUCATIONAL UNIVERSITY OF HONG KONG. *PRAAT Manual*. Online. 2023. Dostupné z: https://corpus.eduhk.hk/english_pronunciation/index.php/2-2-formants-of-vowels/.
- [18] K. GAIKWAD, Santosh a W. GAWALI, Bharti. *A Review on Speech Recognition Technique*. Volume 10. 2010.
- [19] IBM. *What is Speech Recognition*. Online.
- [20] ENGLUND, Christine. *Speech recognition in the JAS 39 Gripen aircraft - adaptation to speech at different G-loads*. Online, Master Thesis. Stockholm: Royal Institute of Technology, 2004. Dostupné z: <https://web.archive.org/web/20081002002102/http://www.speech.kth.se/prod/publications/files/1664.pdf>.
- [21] JUANG, B.H. a RABINER, Lawrence R. *Automatic Speech Recognition – A Brief History of the Technology Development*. 2004. Dostupné také z: https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf.



- [22] PERACCINI, Roberto. *From AUDREY to Siri. Is speech recognition a solved problem?* 2012. Dostupné také z: <https://www1.icsi.berkeley.edu/pubs/speech/audreytosiri12.pdf>.
- [23] *Neural Networks and Deep Learning*. Online. Self-published, 2019. Dostupné z: <https://www.ise.ncsu.edu/fuzzy-neural/wp-content/uploads/sites/9/2022/08/neuralnetworksanddeeplearning.pdf>.
- [24] JANSSON, Peter A. Neural Networks: An Overview. Online. *Analytical Chemistry*. 1991, roč. 63, č. 6. ISSN 0003-2700. Dostupné z: <https://doi.org/10.1021/ac00006a739>.
- [25] HELMKE, Hartmut; OHNEISER, Oliver; MUHLHAUSEN, Thorsten a WIES, Matthias. Reducing controller workload with automatic speech recognition. Online. *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. 2016, s. 1-10. ISBN 978-1-5090-2523-7. Dostupné z: <https://doi.org/10.1109/DASC.2016.7778024>.
- [26] PINSKA-CHAUVIN, Ella; HELMKE, Hartmut; DOKIC, Jelena; HARTIKAINEN, Petri; OHNEISER, Oliver et al. Ensuring Safety for Artificial-Intelligence-Based Automatic Speech Recognition in Air Traffic Control Environment. Online. *Aerospace*. 2023, roč. 10, č. 11. ISSN 2226-4310. Dostupné z: <https://doi.org/10.3390/aerospace10110941>.
- [27] European Commission. Commission Implementing Regulation (EU) 2017/373 of 1 March 2017 Laying down Common Requirements for Providers of Air Traffic Management/Air Navigation Services and Other Air Traffic Management Network Functions and Their Oversight Repealing Regulation (EC) No 482/2008, Implementing Regulations (EU) No 1034/2011, (EU) No 1035/2011 and (EU) 2016/1377 and Amending Regulation (EU) No 677/2011. 2017. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0373>
- [28] *Controller Working Position / Human Machine Interface - CWP/HMI*. 2020. Dostupné také z: <https://doi.org/10.3030/734141>.
- [29] European Space Agency. *Technology Readiness Levels Handbook for Space Applications*. September 2008. TEC-SHS/5551/MG/ap. Available online: https://connectivity.esa.int/sites/default/files/TRL_Handbook.pdf
- [30] EUROCONTROL. *Digital Network Management Services: More efficient network operations*. Studie. 2023. Dostupné také z: <https://cordis.europa.eu/project/id/874463>.
- [31] KARLSSON, Joakim. *Automatic Speech Recognition in Air Traffic Control: a Human Factors Perspective*. NASA, Langley Research Center, Joint University Program for Air Transportation Research, 1989-1990, 1990. Dostupné také z: <https://ntrs.nasa.gov/citations/19910009712>.



- [32] *Machine Learning of Speech Recognition Models for Controller Assistance: MALORCA*. Project Results. 2.00. 2018. Dostupné také z: <https://cordis.europa.eu/project/id/698824/results>.
- [33] HELMKE, Hartmut; RATAJ, Jürgen; MÜHLHAUSEN, Thorsten; OHNEISER, Oliver; EHR, Heiko et al. *Assistant-Based Speech Recognition for ATM Applications*. ATM 2015. Saarland University. German Aerospace Center (DLR), Institut of Flight Guidance, Braunschweig, Germany, 2015. Dostupné také z: https://www.malorca-project.de/wp/wp-content/uploads/Paper363-Helmke-Et-Al_AcListant-FAA-Eurocontrol-V-1.02.pdf.
- [34] SESAR. *HAAWAIi - Highly Automated Air Traffic Controller Workstations with Artificial Intelligence Integration: Exploratory research PROJECT*. 2022.
- [35] APPAERO. *Aviation Speech Recognition Software*. Online. Dostupné z: <https://appareo.com/aviation/aviation-speech-recognition-system/>.
- [36] ZHANG, Aston; LIPTON, Zachary C.; LI, Mu a SMOLA, Alexander J. *Dive into Deep Learning*. Online. Cambridge University Press, 2023. Dostupné také z: <https://D2L.ai>.
- [37] ARATANI, Lori. *TSA broke record on Sunday for number of passengers screened: The Sunday after Thanksgiving lived up to its billing as the busiest travel day of the year*. 2023. Dostupné také z: <https://www.washingtonpost.com/transportation/2023/11/27/tsa-thanksgiving-passenger-screening-record/>.



Seznam příloh

Příloha 1 – Skript pro transkripci mluveného slova a vyhodnocení správnosti zpětného potvrzení

Příloha 2 – Skript aplikace pro zaznamenávání vyhodnocení správnosti zpětného potvrzení pro subjekty

Příloha 3 – Dotazník demografických znaků, webový odkaz

Příloha 4 – Data z měření v programu Excel