UNIVERZITA KARLOVA

Filozofická fakulta

Ústav translatologie

Diplomová práce

Bc. Veronika Žáčiková

Přesnost automatických titulků generovaných platformou Zoom – je na ně spoleh?

The accuracy of automated captions generated by Zoom – are they reliable?

Praha, 2024

Vedoucí práce: Mgr. Kateřina Ešnerová

Poděkování

Děkuji své vedoucí Mgr. Kateřině Ešnerové za věcné poznatky, technickou výpomoc a odborné vedení – praktická část této práce by bez její pomoci nebyla možná. Dále děkuji PhDr. Janě Pokojové, Mgr. Danielu Dolenskému, Mgr. Miguelovi Cuencovi, PhD. a MgA. Michaelovi Baughovi za spolupráci při přípravě experimentu a během jeho uskutečnění. Neskutečně si vážím vřelých gest a ochoty, kterou mi všichni prokázali. A v neposlední řadě chci samozřejmě poděkovat celému svému okolí, které mě podporovalo.

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 29. července 2024 ……

Bc. Veronika Žáčiková

Abstrakt

Tato diplomová práce se zabývá automatickými titulky na videokonferenční platformě Zoom. Klade si za cíl prozkoumat, nakolik automatické titulky mohou pomáhat tlumočníkům, a zda jsou dostatečně přesné, aby se na ně tlumočníci při tlumočení videokonferencí mohli spolehnout. V teoretické části se práce zabývá tématy automatických titulků, technologie ASR, simultánního tlumočení, tlumočení videokonferencí a simultánního tlumočení s textem. Zvláštní pozornost je věnovaná kognitivní zátěži při tlumočení. Teoretická část je zakončena shrnutím dosavadního výzkumu na poli tlumočení s automatickými titulky – jaký dopad mají automatické titulky na kognitivní zátěž tlumočníka a jak jim mohou pomoci. Cílem empirické části je ověřit přesnost automatických titulků na Zoomu ve třech jazycích (čeština, angličtina, španělština). K tomu využíváme metriku NER. Věnujeme se jak celkové přesnosti, tak přesnosti v předem určených kategoriích (jména, zkratky, akronymy, čísla, termíny, výčty a negace), u kterých se domníváme, že automatické titulky mohou být tlumočníkům k užitku. Ukázalo se, že celková přesnost je napříč všemi třemi jazyky poměrně vysoká (průměrně 95 %) a v určitých kategoriích dosahují automatické titulky stabilní vysoké přesnosti (např. domácí jména, nižší řády čísel či termíny). Automatické titulky v každém z jazyků ale mají svá specifika, na které je potřeba si dávat pozor (např. španělské titulky neobsahují interpunkci, znaménka, ani velká písmena, zatímco české titulky nezapisují čísla pomocí číslice, nýbrž slovně).

Klíčová slova

Automatické titulky, Zoom, simultánní tlumočení s textem, CAI, ASR, remote tlumočení

Abstract

This master's thesis deals with automated captions on the videoconference platform Zoom. It aims to examine how automated captions can help interpreters and whether they are accurate enough for the remote interprets to rely on. The theoretical part delves into topics such as automated captions, ASR technology, simultaneous interpreting, remote inrepreting, and simultaneous intepreting with text. Special attention is given to the cognitive load that intepreters experience while intepreting. The theoretical part concludes by summarizing existing research to date in the field of intepreting with automated captions – how automated captions impact intepreters' cognitive load, but also in which ways they can be helpful. The emphirical part aims

to determine the accuracy of Zoom's automated captions in three different languages (Czech, English, and Spanish) using the NER metrics. The research focuses on overall accuracy as well as the accuracy in pre-established categories (names, abbreviations, acronyms, numbers, terms, ennumerations and negations), in which the author deems automated captions might be helpful to interpreters. The results show that the overall accuracy in all examined languages is high (averaging 95%) and the automated captions also steadily achieve high accuracy in certain categories (f.e. local names, lower numbers and terms). However, the automated captions show different negative tendencies in different languages (f.e. Spanish captions do not contain punctuation, signs and capital letters; Czech captions use words to transcribe numbers instead of digits).

Keywords

Automated captions, Zoom, simultaneous interpreting with text, CAI, ASR, remote interpreting

OBSAH

1 **ÚVOD** **7**

2 **TEORETICKÁ ČÁST** **8**

2.1 **Titulky** 9

2.1.1 Inkluze 10

2.2 **Simultánní tlumočení** 12

2.2.1 Tlumočení na dálku a tlumočení videokonferencí 12

2.2.2 Simultánní tlumočení s textem 17

2.2.3 Kognitivní zátěž u tlumočení 18

2.3 **CAI** 21

2.3.1 Obeznámenost tlumočníků s CAI nástroji a jejich využívání 23

2.4 **ASR** 26

2.4.1 Historie 27

2.4.2 Přesnost a faktory ovlivňující přesnost 27

2.4.3 Metriky pro výpočet přesnosti automatických titulků 30

2.5 **Tlumočení a ASR** 35

2.5.1 Kognitivní zátěž u tlumočení při využívání ASR 35

2.6 **V čem mohou automatické titulky tlumočníkům pomoci** 40

2.6.1 Jména a názvy 40

2.6.2 Zkratky a akronymy 43

2.6.3 Čísla 43

2.6.4 Termíny 44

2.6.5 Výčty 46

2.6.6 Negace 47

3 **EMPIRICKÁ ČÁST** **48**

3.1 **Metodologie** 48

3.1.1 Příprava experimentu a účastníci 48

3.1.2 Tvorba projevu 49

3.1.3 Vyhodnocení 54

3.2 **Výsledky experimentu** 57

3.2.1 Celková přesnost 57

3.2.2 Výsledky v rámci jednotlivých kategoriích 60

3.3 **Shrnutí výsledků a závěrečná diskuze** 76

3.3.1 Srovnání s předchozími výzkumy, limitace našeho výzkumu a budoucí výzkum 79

4 **ZÁVĚR** **81**

5 **SEZNAM POUŽITÉ LITERATURY** **82**

6 **SEZNAM OBRÁZKŮ A GRAFŮ** **94**

7 **PŘÍLOHY** **96**

1 ÚVOD

Čím dál tím více se v online prostoru objevují automatické titulky (na platformách jako YouTube, Instagram a další) generované technologií ASR (*Automatic Speech Recognition*, automatické rozpoznávaní řeči). Ty slouží nejen jako nástroj inkluze, ale v teorii mohou posloužit také tlumočníkům při tlumočení videokonferencí. Mnozí tlumočníci se ale bojí se na ně spolehnout, jiní je ignorují kompletně. Důvody pro tuto nedůvěru jsou mnohé, jedním z nich je například obava z nedostatečné přesnosti automatických titulků. V této práci se soustředíme na platformu Zoom, přičemž se pokusíme zodpovědět, zda je tato nedůvěra odůvodněná či nikoli pomocí vytvoření experimentálního projevu ve třech různých jazycích a následného zhodnocení přesnosti automatických titulků při jejich přepisu.

V teoretické části nejdříve rozebereme pozitivní dopad automatických titulků, hlavně pro osoby s poruchou či vadou sluchu. Dále se budeme věnovat typům tlumočení, která jsou relevantní pro téma naší práce (simultánní tlumočení, tlumočení videokonferencí a simultánní tlumočení s textem) a jejich kognitivní zátěži. Ve třetí a čtvrté kapitole vysvětlujeme, co to jsou CAI nástroje, jak se dělí, kam spadá ASR, co to ASR je, jak funguje, jeho limitace a používané metriky k vyhodnocení jeho přesnosti. V neposlední řadě pak na základě poznatků o tlumočení a ASR probíráme, jak fungují společně: jaký dopad mají automatické titulky tvořené ASR na kognitivní zátěž tlumočníků a v čem jim mohou pomoci.

V empirické části se věnujeme experimentu, ve kterém zkoumáme přesnost ASR ve třech jazycích (čeština, angličtina, španělština). Nahlížíme na přesnost celkově, z hlediska čtenáře, ale také z hlediska tlumočníka v rámci předem vymezených kategorií (jména, zkratky, akronymy, čísla, termíny, výčty a negace). K vyhodnocení používáme metriku NER, která nám umožňuje přiřazovat různým chybám různý stupeň závažnosti.

V diskuzi poté shrnujeme tendence automatických titulků v jednotlivých jazycích, srovnáváme vlastní výsledky s předchozím výzkumem a také uvažujeme nad dalšími důležitými aspekty automatických titulků, pokud je mají tlumočníci v budoucnu využívat.

2 TEORETICKÁ ČÁST

V rámci teoretické části definujeme klíčové termíny a snažíme se uvést čtenáře do kontextu tlumočení a ASR. Online tlumočení videokonferencí se v dnešní době stává čím dál tím častěji poptávanou službou (Rayaa & Martin, 2022), přičemž se k jeho výkonu používá hned několika různých platforem (Interprefy, QuaQua a další). Jednou z takových platforem je i Zoom, po období pandemie v České republice nejvíce rozšířená platforma na konání online konferencí (Kasík 2023). Zoom nabízí službu tzv. automatických titulků, které jsou vytvářeny za použití technologie ASR (*Automatic Speech Recognition*). Takové titulky napomáhají přístupnosti audiovizuálního obsahu pro osoby s poruchou či vadou sluchu, a proto v naší práci rozebíráme také téma inkluze.

Z hlediska tlumočení se tato práce zaměřuje pouze na simultánní tlumočení videokonferencí. Během tohoto typu tlumočení mohou automatické titulky pomáhat tlumočníkům právě již během výkonu tlumočení. Konsekutivní tlumočení videokonferencí bylo vynecháno, protože k němu tlumočník nemůže využívat titulků tolik jako při simultánním tlumočení (ačkoli Wang & Wang, 2019, dokazují, že to není nemožné). Tlumočení videokonferencí s automatickými titulky lze považovat za tlumočení na dálku s textem, a proto se v rámci kapitoly tlumočení věnujeme definici disciplín simultánního tlumočení, simultánního tlumočení na dálku (a simultánního tlumočení videokonferencí jako jeho podtypu), simultánního tlumočení s textem a jednotlivé kognitivní zátěži zmíněných disciplín.

Těžiště práce tkví v experimentu zkoumajícím přesnost automatických titulků, a proto považujeme za důležité v této části práce vysvětlit, jak ASR zapadá do počítačem podporovaného tlumočení (CAI), jak ASR funguje, jaké metriky se používají k výpočtu jeho přesnosti a na jaké problémy a limity technologie ASR naráží.

V závěru teoretické části práce propojíme jednotlivé poznatky z oblasti tlumočení a technologie ASR, abychom shrnuli, v čem může ASR pomáhat tlumočníkům.

2.1 Titulky

V této kapitole vymezíme základní názvosloví ohledně titulků a také se vyjádříme k jejich historii a přínosu.

Pro termín titulky se v americké angličtině používá dvojí označení: *subtitles* a *captions*.

Jako *captions* se označuje text, který:

1. zachycuje okolní zvukovou složku (řeč, ale také zvuky jako smích, hudbu apod.),
2. zachycuje, pokud je to potřeba, kdo mluví. (Spina, 2021)

Captions se tedy primárně zaměřují na potřeby neslyšících.

Jako *subtitles* se označuje text, který je překladem z výchozího jazyka, a kromě samotné řeči výše zmíněné informace neobsahuje. *Subtitles* totiž většinou nepředpokládají, že by jejich uživatelé byli sluchově postižení, a proto v sobě nezahrnují popis jiných zvukových složek ani označení mluvčího. V britské angličtině toto rozlišení neexistuje a oběma typům se říká *subtitles,* někdy s upřesněním SDH (*subtitles for the d/Deaf and hard of hearing,* titulky pro osoby neslyšící a nedoslýchavé). Spina (2021) rozlišuje mezi *interlingual subtitles* (mezijazykové titulky) a *intralingual subtitles* (= *captions,* vnitrojazykové titulky).

V češtině podobné (jazykové) rozlišení nenajdeme. Občas se setkáme s upřesněním titulků pro neslyšící a nedoslýchavé, případně s názvem „skryté titulky“. (Pošta, 2023)

Obdobně, španělština také nerozlišuje mezi *captions* a *subtitles*, obojí nazývá jako *subtítulos* a případně přidává v případě titulků pro neslyšící upřesnění *subtítulos SDH* vycházející z angličtiny.

Tradiční titulky vytvářejí titulkáři, kteří zvuk videa nejdříve přepíší a poté mu přiřadí časové značky tak, aby zobrazovaný text odpovídal tomu, co se zrovna říká. Text se pak dále může překládat do dalších jazyků. Tento typ titulků lze ale vytvořit pouze ex post.

Jedním z prvních řešení, jak vytvářet titulky v reálnem čase, byla profese stenografů a tzv. respekearů, profese také zvaná *live subtitling* (Romero-Fresco & Martínez, 2015) či *voice writing.* (TimeLine Media, 2022)

Stenografové přepisují zvuk pomocí speciálního zařízení a programu, který dekóduje jejich vlastní zkratky. Dodnes je asi nejlépe známe ze zapisování soudních procesů. Respeakeři používají technologii ASR (viz kapitola 2.4 *ASR*), která přepisuje to, co říkají, do textu a je trénovaná přímo na nich a jejich hlasu (ibid.). V mnohém se respeakeři dokonce podobají tlumočníkům (musí opakovat řečenou informaci, rozhodovat, co je v rámci zachování vyhovujícího odstupu od řečníka podstatné apod.), s tím rozdílem, že pracují v rámci jednoho jazyka.

Automatické titulky jsou takové titulky, které jsou vytvářeny právě technologií ASR. Je v nich ale vynechán mezikrok respeakerů, kteří by řeč pro ASR zjednodušili či předfiltrovali.

Co se týče označení těchto titulků, to je v angličtině různorodé: *automatic subtitles* (Romero-Fresco & Martínez, 2015), *automated captions* (Zoom, 2024a) či *live captions.* (Google Chrome, 2024a)

V českém kontextu se pro tyto titulky vžil název automatické titulky, a proto toto označení budeme v naší práci používat i my.

Ve španělštině se používá dvojího označení vycházejícího z angličtiny: *subtítulos automáticos* (Google Chrome, 2024b) či *subtítulos automatizados,* (Zoom, 2023a)

2.1.1 **Inkluze**

Ve 21. století pozorujeme rostoucí zájem společnosti o tzv. inkluzi. Komise evropských společenství (dnes Evropská komise) definovala sociální inkluzi na začátku tohoto tisíciletí následovně:

„*Sociální inkluze je proces, který zajišťuje, že osoby, které jsou ohroženy chudobou a sociální exkluzí, dostanou příležitosti a zdroje potřebné k tomu, aby se plně účastnily ekonomického, sociálního, politického a kulturního života a měly životní úroveň takovou, jaká se ve společnosti, ve které žijí, považuje za normální. Zajišťuje, že tyto osoby mají vyšší účast na rozhodovacích procesech, které ovlivňují jejich život, a přístup k základním právům.“* (Commission of the European Communities, 2003, vl. překlad)

Co se týče oblasti zábavního průmyslu, jakýkoli audiovizuální obsah byl doteď vysoce nepřístupný jedincům s poruchou či vadou sluchu. Způsobů, jak to změnit je několik: během pandemie covidu jsme například viděli nárůst tlumočení televizních zpráv do znakového jazyka (Bačáková, 2023). Dalším způsobem ale mohou být právě titulky.

O tom, nakolik jsou titulky inkluzivní, rozhoduje několik faktorů: Spina (2021) zmiňuje například zda jsou *hard-coded* nebo *soft-coded*. *Hard-coded*, tzv. „vypálené“ do obrazu, jsou titulky bez možnosti úpravy, anglicky také zvané *open captions*. Jejich výhodou je, že je nikdo nemusí zapínat či vypínat, mohou tak ale působit jako rušivý podnět. *Soft-coded* titulky lze zapnout a vypnout, navíc mohou být upravitelné z hlediska velikosti, fontu či barvy, anglicky jsou také zvané *closed captions.*

Prvním televizním programem, který pravidelně obsahoval titulky pro osoby s poruchou či vadou sluchu, byl *The French Chef* v USA v roce 1972, tehdy se jednalo o tzv. vypálené titulky. (AI Media, 2023b) Vzestup automatických titulků vidíme až s růstem internetu a streamovacích platforem, např. automatické titulky na YouTube byly zavedeny roku 2009. (YouTube, 2023) Mnohdy se ale zavedení titulků musí asociace pro sluchově postižené domáhat soudně, jako to například bylo v případě Netflixu. (Mullin, 2012) Přitom osoby s poruchou či vadou sluchu tvoří až 20 % světové populace. (World Health Organization, 2019)

Nejsou ale jediní, kteří titulky využívají. Spina (2021) mluví také o skupinách lidí, kteří lépe zpracovávají informace skrze text, dále pak lidí, kteří si z různých důvodů nemohou při sledování videa zapnout zvuk (poslouchají v knihovně, nemají sluchátka apod.), případně lidí, jejichž mateřský jazyk není jazykem mluvčího, nebo když má mluvčí specifický přízvuk. Průzkum společnosti Verizon Media z roku 2019 ukázal, že 80 % spotřebitelů v rozmezí 18–54 let, kteří využívají titulky pro neslyšící a nedoslýchavé, netrpí vadou či poruchou sluchu (ibid.).

Díky automatickým titulkům tak mohou mít všechny tyto skupiny lidí přístup k audiovizuálnímu obsahu, který jim byl doteď nepřístupný.

2.2 Simultánní tlumočení

V této kapitole probereme simultánní tlumočení a jeho typy, které jsou relevantní pro téma naší práce: tlumočení videokonferencí (protože platformy, na kterých se takto tlumočí, mohou nabízet funkci automatických titulků) a simultánní tlumočení s textem (protože v momentě, kdy tlumočníci začnou automatické titulky využívat, se jedná o tento typ tlumočení). Zaměříme se také na platformu probíranou v této práci – Zoom. V neposlední řadě se soustředíme také na kognitivní zátěž těchto tlumočení, konkrétně specifika probíraných typů tlumočení a typických problémů s nimi spojených.

Čeňková (2008) o simultánním tlumočení píše:

„*(...) [P]odstatou simultánního tlumočení je souběžnost řečových činností a rozdvojení pozornosti tlumočníka na poslech originálu a jeho převod do cílového jazyka.“*

Počátek simultánního tlumočení tak, jak ho známe dnes, se datuje do roku 1927. Díky mikrofonu a sluchátkům mohli tlumočníci slyšet výchozí projev a zároveň ho tlumočit pro více posluchačů najednou (ibid.). To vedlo ke vzniku tzv. kabinového tlumočení. Tlumočník při něm sedí v kabině a mívá výhled na řečníka či publikum. Zpravidla se provádí ve dvou, kdy se tlumočníci střídají. Dobrou praxí je, že netlumočící tlumočník i nadále poslouchá a pomáhá kolegovi ať už psychicky, nebo pomocí zápisu projevu (o tom, v čem tlumočníci potřebují pomoc, pojednáváme v kapitole 2.2.3 *Kognitivní zátěž u tlumočení*).

2.2.1 Tlumočení na dálku a tlumočení videokonferencí

Podobně jako před sto lety, nástup nových technologií a posléze internetu ovlivnil i tlumočnickou profesi. V tomto případě znamenal, že tlumočníci, řečníci, či posluchači se již nutně nemusejí nacházet na stejném místě. Nicméně, na tom, jak tento typ tlumočení nazývat, se výzkumníci neshodnou. Mnoho výzkumníků zavádí vlastní termíny pro jednotlivé podtypy tlumočení, a proto považujeme za důležité se této problematice věnovat o něco hlouběji.

Například Čeňková definuje tlumočení na dálku, neboli *remote interpreting*, jako situaci, „*kdy jsou všichni účastníci spolu ve stejném sále, ale tlumočnické kabiny jsou umístěny jinde (...) a nemají na řečníka ani na posluchače, pro které tlumočí, přímý výhled“.* Dále Čeňková rozlišuje podtyp tohoto tlumočení, „*kdy se jak účastníci, tak tlumočníci nacházejí v různých*

místech světa a jsou propojeni díky satelitnímu přenosu obrazu i zvuku“. To nazývá tlumočením videokonferencí. Hlavním znakem je, že jsou všichni účastníci ve stejné situaci, tedy že se vidí a komunikují pouze za pomoci videopřenosu.

Pöchhacker (2004) popisuje tlumočení na dálku jako obojí: jedná se o typ tlumočení, při kterém tlumočník není na stejném místě jako řečník, posluchač, či oba dva. Původně se podle něj jako tlumočení na dálku označovalo telefonní tlumočení, které bylo hlavně bilaterální (do obou jazyků) a konsekutivní (řečník a tlumočník nemluví souběžně). S nástupem videokonferencí pak mluví o tzv. tele-tlumočení (*tele-interpreting*).

Braun (2015) mluví o tlumočení na dálku podobně jako Čeňková, tedy o situacích, kdy jsou účastníci spolu a tlumočník je mimo. Vyděluje poté „telekonferenční tlumočení“ (*teleconference interpreting*), při kterém se všichni participanti nachází na jiných místech, a dále rozlišuje telefonní a videokonferenční tlumočení samostatně jako podtypy tohoto tlumočení.

Pro účely naší práce proto přejímáme přesnější název tlumočení videokonferencí, protože nejpřesněji popisuje situace, kterým se v práci věnujeme. Jedná se tedy o situace, kdy se řečník, tlumočník i posluchač nenachází na stejném místě a jsou spolu spojeni přes komunikační platformu pomocí audio a videopřenosu. Pokud mluvíme o tlumočení na dálku, myslíme tím obecnější pojetí, o kterém mluví Pöchhacker (2004).

Pro tlumočení videokonferencí je příznačné, že se řečník, posluchač i tlumočník připojí online a každý využívá vlastních zařízení. Pro tlumočníka to v mnohých případech znamená, že je sám sobě technikem. To zahrnuje zajištění počítače či notebooku, mikrofonu a sluchátek, zároveň tlumočník musí najít tiché místo na výkon práce a sám musí zkontrolovat, že technika a internet funguje. (Ešnerová & Tite, 2020)

Důležité je, aby se tlumočení odehrávalo na platformě, která umožňuje funkci simultánního tlumočení. Akhulkova et al. (2022) třídí platformy, které se používají k tlumočení, na tyto:

- Videokonferenční platformy bez funkce simultánního tlumočení: řadí se sem např. Skype, u kterého je ale potřeba vymyslet alternativní řešení, pokud má být tlumočení skutečně simultánní. Tím může být například oddělený hovor na aplikaci WhatsApp, či propojení s tlumočnickou platformou.

- Videokonferenční platformy s funkcí simultánního tlumočení: do této kategorie se řadí platformy jako Zoom, Webex, Google Meet a MS Teams. Všechny tyto platformy nabízí funkce simultánního tlumočení, ačkoli v mnohých ohledech omezené – ne všechny nabízí pilotáž, či speciální rozhraní pro tlumočníky.
- Samostatné specializované platformy pro simultánní tlumočení: tyto jsou navržené přímo pro tlumočníky. Disponují tak speciálním rozhraním, ve kterém spolu tlumočníci mohou komunikovat, indikovat předání slova apod.
- Platformy pro simultánní tlumočení s virtuální kabinou: tyto jsou velmi podobné předchozím, akorát s tím rozdílem, že nejsou napojené na konkrétní platformu, na které by se akce musela odehrávat. Tlumočníci tak získají výhody specializovaného rozhraní, zatímco klienti mohou využívat videokonferenční platformy, na kterou jsou zvyklí.

Co se týče výhod a nevýhod tlumočení videokonferencí, Rayya & Martin (2022) provedli dotazníkové šetření mezi 37 tlumočníky, ze kterého vyplývá, že mezi hlavními výhodami tlumočení videokonferencí figuruje flexibilita práce z domova, ušetření času za dopravu a peněz pro organizátory, kratší práce, příležitost pracovat na předtím nedostupných akcích a větší publikum. Kasík (2023) pak z rozhovorů s 8 tlumočníky a výzkumníky ještě doplňuje např. snížení emisí CO_2, příjemnější (domácí) prostředí, zvýšenou dostupnost menších jazyků a možnost říct si o vyšší honorář, pokud má tlumočník pro danou oblast vzácnou jazykovou kombinaci. Dále tlumočníci mluví o větší konkurenceschopnosti, protože týmy tlumočníků mohou být poskládané z tlumočníků po celém světě. Protože velké procento tlumočníků jsou ženy, uváděly v rozhovorech také možnost se starat lépe o domácnost a mít svůj životní styl pod kontrolou.

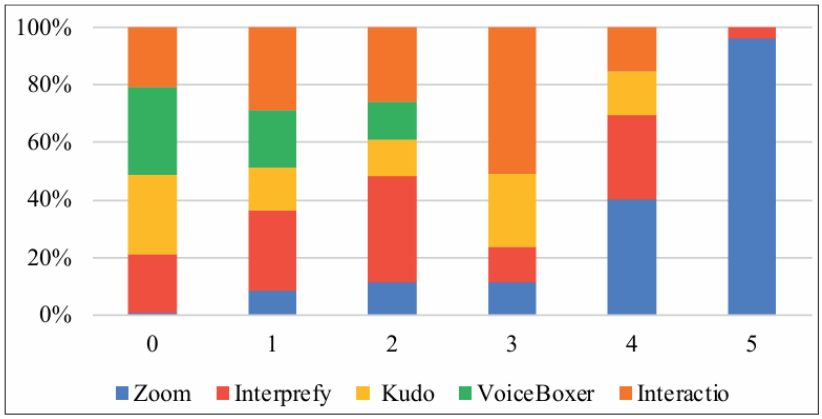
Jakožto nevýhody tlumočníci vidí právě nedostatek technické podpory, což může být na jednu stranu nedostatečně kvalitní technika (špatná kamera či mikrofon řečníka apod.), ale i to, že platformy, na kterých se videokonference konají, nemívají speciální rozhraní určené pro tlumočníky např. k předání slova (jako je tomu např. u Zoomu), a proto se tlumočníci musí i nadále scházet osobně, přicházet s alternativními řešeními, případně se spokojit s nepřítomností pomáhajícího kolegy. Ohledně tohoto zmiňují jako nevýhodu i nedostatek interakce s kolegy, jakožto i s řečníkem a posluchači. Dále jako nevýhody vidí vyšší kognitivní zátěž, možné

problémy se sluchem do budoucna, nejistá autorská práva u nahrávání tlumočení, osamocenost, neprofesionální prostředí a zhoršující se podmínky do budoucna kvůli větší konkurenci. Kasík (2023) na základě rozhovorů doplňuje náklady na potřebné vybavení a obavu z toho, že klienti nabydou dojmu, že tlumočení obstarává „nějaká krabička“, protože nevidí tlumočníky fyzicky na místě. Globalizaci lze dle tlumočníků v Kasíkově šetření považovat také za teoreticky negativní, protože kvůli časovému posunu může práce přijít např. v noci, což není komfortní, a zároveň zahraniční tlumočníci získávají přístup na domácí trh, což zhoršuje situaci domácích tlumočníků.

2.2.1.1 Platforma Zoom

Zoom je platforma pro videokonference, na které se uživatelé mohou propojit pomocí videa i zvuku. Navíc mohou sdílet obrazovku s dokumenty, prezentací apod. Zoom lze otevřít ve webovém rozhraní, které podporuje většinu funkcí, ale aby měl uživatel přístup ke všem nabízeným funkcím, je potřeba si stáhnout aplikaci a přihlásit se.

Zoom se rozšířil hlavně během pandemie Covidu-19 (Molla, 2020), pro školy a studenty nabídl dokonce neomezené trvání schůzek zdarma (Zoom, 2023b). Mnoho uživatelů si tak na Zoom navyklo a využívají ho i nadále, což doplňují také zkušenosti tlumočníků: podle výzkumu Rayaa & Martina (2022) většina tlumočníků, kteří tlumočí online, nejčastěji využívají právě platformu Zoom. To podporuje také výzkum Kasíka (2023) provedený na českém trhu, kdy 47 % z 60 dotazovaných tlumočníků používá k tlumočení na dálku nejčastěji právě Zoom, zatímco jiné platformy obdržely méně než 15 %.



Obrázek 1: Rayaa & Martin (2022): které platformy využíváte při tlumočení videokonferencí nejvíce? 0 = nikdy jsem na této platformě nepracoval/a; 5 = na této platformě pracuji nejčastěji.

Pro tlumočníky je Zoom důležitý v tom, že poskytuje funkci simultánního tlumočení. V praxi to funguje tak, že hostitel musí tlumočení umožnit před konáním schůzky (tato funkce je dostupná pouze, pokud má hostitel placenou verzi). Tlumočník je pak pozván e-mailem, popřípadě je mu role tlumočníka přiřazena během schůze a při tlumočení má velmi podobné podmínky jako při kabinovém tlumočení: může se vypínat a zapínat, slyší řečníka ve sluchátkách a mluví do vlastního mikrofonu. Host si může mezi jazyky přepínat a slyší pouze jeden zvolený jazyk, zatímco video zůstává na původním řečníkovi. (Zoom, 2023c) Nevýhodou je to, že se tlumočníci při tlumočení nevidí navzájem, nemohou si v aplikaci pomáhat (kromě využívání funkce soukromého chatu, která je ale vysoce riskantní, protože probíhá ve stejném okně jako chat se všemi účastníky setkání) a nemohou si naznačit předání slova.

Kromě funkce tlumočení Zoom nabízí také automatické titulky, jejichž přesnost je hlavním tématem této práce. Ty fungují na základě technologie ASR, o které pohovoříme v samostatné kapitole. Pro tlumočníky je ale důležité, že pokud jsou automatické titulky zapnuty, vidí je i oni během tlumočení (a nezaznamenávají jejich tlumočení, nýbrž výchozí projev).

Obrázek 2: Zobrazení automatických titulků v aplikaci Zoom. (Ball, 2022)

Aktuálně (červen 2024) poskytuje Zoom automatické titulky pro následující jazyky: arabština, kantonská čínština, čínština (zjednodušená a tradiční), čeština, dánština, nizozemština, angličtina,

estonština, finština, francouzština, kanadská francouzština, němčina, hebrejština, hindština, maďarština, indonéština, italština, japonština, korejština, malajština, perština, polština, portugalština, rumunština, ruština, španělština, švédština, jazyk tagalog, tamilština, jazyk telugu, thajština, turečtina, ukrajinština a vietnamština. (Zoom, 2024b) Navíc, pro účty, které si platí plány Workplace Business Plus nebo Workplace Enterprise Plus, Zoom nabízí také možnost překladu automatických titulků mezi všemi uvedenými jazyky. Pokud je výchozí jazyk jeden z výše zmíněných, je překlad navíc možný také do bengálštiny, řečtiny, norštiny a velštiny (ale nikoli naopak). (Zoom, 2024c)

Co se týče procentuální přesnosti automatických titulků, sama platforma žádné konkrétní číslo nezmiňuje. Doporučuje ale, že pokud je třeba dosáhnout určité úrovně přesnosti, měl by organizátor zvážit přiřazení funkce titulkáře člověku. (Zoom, 2024a)

Existují ale výsledky nezávislých studií z univerzit, kde se Zoom využíval jako nástroj ke konání distanční výuky. Přesnost Zoom titulků se podle nich různí. Univerzita v Denveru mluví o 90% přesnosti, aniž by poskytla jakékoli údaje k tomu, jak a na čem přesnost zkoumala. (University of Denver, 2021) Obdobně, kalifornská Univerzita Santa Barbara zmiňuje přesnost 80 % bez poskytnutí detailů, jak k tomuto číslu došla. (University of California, Santa Barbara, 2021)

2.2.2 Simultánní tlumočení s textem

Tlumočení s textem se řadí jako subdisciplína simultánního tlumočení. Čeňková (2008) ho definuje jako činnost, „*při které tlumočník [musí] roztrojit svou pozornost mezi poslech originálu, vizuální sledování písemného podkladu a vlastní tlumočení*“. L. Zou (2022) zmiňuje typy textů, které tlumočníci mohou mít k dispozici dopředu: může se jednat o přepis projevu, prezentaci či dokumenty a další podklady. Je ale nutno poznamenat z praxe, že ne vždy dostanou tlumočníci text s dostatečným předstihem, aby si ho mohli prostudovat.

V mnohém se simultánní tlumočení s textem podobá tlumočení z listu, kdy má tlumočník list před sebou a tlumočí ho nejenom ze psané formy do mluvené, ale zároveň z výchozího jazyka do cílového. Existují mezi nimi ale dva základní rozdíly:

1. Rychlost čtení není závislá na tlumočníkovi, nýbrž řečníkovi. Ten rozhoduje, jak rychle musí tlumočník zpracovávat informace. (Chmiel & Lijewska, 2019)

2. Psaný text je během simultánního tlumočení s textem pouze sekundárním zdrojem informací, tím hlavním je zvukový vstup. (Pöchhacker, 2004)

I kvůli těmto hlediskům se simultánní tlumočení s textem řadí mezi jednu z nejtěžších tlumočnických disciplín. (Gile 2009)

2.2.3 Kognitivní zátěž u tlumočení

Gile (2009) ve svém modelu úsilí popisuje, čemu tlumočník v rámci výkonu simultánního tlumočení věnuje kapacitu (úsilí). Rozlišuje čtyři možná úsilí: úsilí poslechu (poslech a analýza řečeného), úsilí memorizace (zapamatování si řečeného), úsilí produkce (reformulace řečeného do jiného jazyka) a úsilí koordinace (rozdělení dostupných kognitivních prostředků do jednotlivých úsilí). Jeho model vypadá takto:

Simultánní tlumočení = úsilí poslechu + úsilí memorizace + úsilí produkce + úsilí koordinace

Aby tlumočení probíhalo správně, nesmí celkové požadavky jednotlivých úsilí překročit dostupnou kapacitu na zpracování. Gile pracuje s tzv. *tightrope hypothesis*, neboli hypotézou, že tlumočníci většinově pracují právě na pokraji překročení této kapacity.

Pokud k jejímu překročení dojde, jedná se o tzv. saturaci neboli zahlcení. Případně může také dojít ke špatnému rozložení kapacity: pokud tlumočník věnuje 80 % dostupné kapacity na elegantní reformulaci řečeného, nejspíše nebude mít dostatek volné kapacity na poslech a analýzu nového segmentu. To pak vede k tzv. individuálním deficitům v rámci daných úsilí. Jak saturace, tak individuální deficit vede ke špatnému přetlumočení, ačkoli ne nutně tam, kde k jednotlivým problémům došlo. V některých případech je problematický segment přetlumočen správně, ale nedostatek kapacity se projeví později, na jinak jednoduchém segmentu (tzv. deficitní řetězec).

K saturaci či individuálnímu deficitu může dojít mimo jiné, pokud se ve výchozím projevu objeví tzv. spouštěče problémů (vl. překlad; orig. *problem triggers*). Mankauskienė (2018) na základě předchozích studií rozděluje spouštěče problémů v tlumočení do následujících kategorií:

- Problémy spojené s řečníkem: přízvuk, nerodilý mluvčí, vysoká rychlost projevu, monotónní intonace, čtený projev

- Problémy spojené s výchozím textem:
    - Lexikální: vlastní jména, čísla, zkratky, technické termíny
    - Syntaktické: idiomy, kolokace, syntaktické rozdíly mezi jazyky, lexikální nahuštěnost, dlouhé věty, dlouhá souvětí, výčty
    - Sémiologické: metafory, vtipy, sarkasmus
- Problémy spojené s tlumočníkem: zkušenost, znalosti, komunikační schopnosti, únava apod.
- Problémy spojené s technikou: potíže s tlumočnickým vybavením, ruch okolí, nedostatečný výhled na řečníka apod.

V rámci případové studie Mankauskienė zkoumá, které spouštěče problémů jsou nejzávažnější. Studie byla provedena na profesionálních tlumočnících mezi angličtinou a litevštinou na třech konferencích organizovaných Evropským Parlamentem. Ze zjištění vyplývá, že nejzávažnějšími spouštěči problémů byly rychlý projev, silný přízvuk (zvláště pokud se jednalo o nerodilého mluvčího) a lexikální problémy (čísla, jména, příjmení, zkratky, místní názvy a názvy institucí).

2.2.3.1 Kognitivní zátěž u tlumočení videokonferencí

Podle Moser-Mercer (2005) každá změna v prostředí, ve kterém tlumočník tlumočí, vede ke změně kognitivní zátěže. Mění se tím totiž také způsoby, jakými tlumočník vnímá zvukový i vizuální vstup. Dle experimentu Moser-Mercer (2003) tlumočníci, kteří pracovali na dálku, pociťovali během výkonu tlumočení více stresu a únavy. Mj. popisovali nepříjemné pocity bezmoci, protože měli pocit, že situaci nemají pod kontrolou, což poté mělo dopad na jejich krátkodobou a dlouhodobou paměť. Obdobně shrnují výzkum v oblasti kognitivní zátěže u tlumočení na dálku i Rayaa & Martin (2022). Vyzdvihují hlavně odtrženost od místa konání a kolegy (se kterým si musí volat nebo chatovat, což je další souběžná činnost navíc), ale i stres vyvolaný obavami o internetové připojení.

Detailně se této problematice věnovali Roziner & Shlesinger (2010), kteří zkoumali rozdíly mezi tlumočením na místě a tlumočením na dálku. Dle jejich závěrů je hlavní přidaný stresový faktor to, že tlumočníci neviděli pořádně na řečníka a publikum a postrádali přímou zpětnou vazbu. Hlavním rozdílem však bylo sebehodnocení tlumočníků a subjektivní pocity – tlumočníci na dálku si více stěžovali na bolest hlavy, únavu očí a problémy se soustředěním. Svůj výkon

hodnotili negativněji, než je hodnotil panel objektivních porotců. Kvalita tlumočení při tlumočení na dálku se tedy oproti tlumočení na místě nezhoršila, ačkoli tlumočníci samotní měli pocit, že ano.

Ačkoli se tedy tlumočnická obec shoduje, že tlumočení na dálku představuje vyšší kognitivní zátěž, není prokázáno, že by to mělo přímý dopad na kvalitu tlumočení. Moser-Mercer (2003) však na základě výsledků doporučuje častější střídání, aby se přílišné únavě předešlo.

2.2.3.2 Kognitivní zátěž u simultánního tlumočení s textem

Co se týče simultánního tlumočení s textem, přidává se dle Gila k výše zmíněným úsilí ještě úsilí čtení. Nový model tedy vypadá následovně:

Simultánní tlumočení s textem = úsilí čtení + úsilí poslechu + úsilí memorizace + úsilí produkce + úsilí koordinace

Gile zmiňuje dvě výhody simultánního tlumočení s textem:

1. Tlumočníci slyší projev, ačkoli aspekty čtených projevů jako pauzy a intonace se mohou od spontánních projevů lišit.
2. Díky textu se snižují nároky na paměť, případně se snižuje i úsilí vynaložené na poslech a analýzu, pokud má mluvčí specifický přízvuk, nebo se tlumočí ve špatných zvukových podmínkách.

Jako nevýhody zmiňuje následující:

1. Text bývá nahuštěný informacemi a psaný jazyk se liší od mluveného, což naopak zvyšuje úsilí vynaložené na analýzu.
2. Zvyšuje se riziko jazykových interferencí.
3. Zvyšuje se kognitivní zátěž, protože tlumočník musí sledovat mluvený i psaný projev. Obzvláště začínající tlumočníci se snaží přetlumočit vše i v případech, že nestíhají řečníka. Navíc se řečník může od psaného projevu odklonit, čehož si tlumočník musí být vědom.

Ohledně přidané kognitivní zátěže s Gilem souhlasí také Seeber (2017), který ve svém modelu predikuje, že přidání textu k simultánnímu tlumočení zvyšuje kognitivní nároky na tlumočníka.

2.3 CAI

CAI, neboli počítačem podporované tlumočení (*Computer Assisted Interpreting*), je tlumočení, při kterém tlumočníci využívají počítačového softwaru. Nejedná se však o strojové tlumočení. Rozdíl mezi počítačem podporovaným tlumočením a strojovým tlumočením je ten, že při CAI počítač může tlumočníkovi pomoci (např. přepsat výchozí řeč, přeložit ji a překlad nabídnout tlumočníkovi k využití), zatímco při strojovém tlumočení zachází stroje ještě dále a produkují řeč v cílovém jazyce pomocí technologie syntézy hlasu. (Wang & Wang, 2019) O automatickém či strojovém tlumočení mluvil již Pochhacker (2004). Nyní, v roce 2024, se na strojovém tlumočení pracuje (mj. také pod názvy *speech-to-speech translation* nebo *AI interpreting*). (Fantinuoli, 2024)

V této kapitole se pokusíme vymezit, co to CAI nástroje jsou a jak mezi ně spadá automatické rozpoznání řeči. Neexistuje totiž jedna přijímaná definice CAI nástrojů. Guo et al. (2022) nabízí komplexní přehled většiny dosud vydaných empirických studií na téma CAI nástrojů a srovnává jednotlivé definice. Zde nabízíme výběr z prací, ze kterých jsme také čerpali:

- Fantinuoli (2018) definuje CAI nástroje jako počítačové programy navržené, aby pomáhaly tlumočníkům v alespoň jednom z jednotlivých podprocesů tlumočení. Rozlišuje mezi první a druhou generací. První generace se zaměřuje na zpracování multilingválních glosářů v rozhraní, které tlumočníci mohou využít, ale nenabízí pokročilé vyhledávání. Zahrnuje tam Interplex či Terminus.

  Druhá generace pak nabízí pokročilejší funkce, např. možnost organizovat texty, vyhledávat v korpusu či dalších zdrojích apod. Sem patří např. InterpretBank či Interpreter's Help (konkrétně jejich AI funkci Boothmate, která je ale k červnu 2024 pozastavena). Mluví také o budoucí generaci CAI nástrojů, které budou využívat AI a NLP (*Natural language processing*).
- Prandi (2020): CAI nástroje jsou softwarová řešení navržená pro tlumočníky, které jim pomáhají s terminologií a organizací znalostí.
- Corpas Pastor et al. (2018): nástroje k organizaci terminologie, k tlumočnickému zápisu, zařízení, která převádí hlas do textu, případně převádí jednotky (měnu apod.) atd.

- Wang & Wang (2019): mezi CAI nástroje patří i takové, které nejsou navržené pro tlumočníky (jako internet, Microsoft Word či Excel).

Tuto absenci shody na definici CAI nástrojů Guo et al. shrnují v podobě tří hlavních otázek, v nichž se různé definice liší:

1. Musí být CAI nástroje navržené pro tlumočníky?
2. Patří počítačem podporované nástroje k výuce tlumočení (CAIT), audio a video konferenční software a vyhledávače mezi CAI nástroje?
3. Je ASR a AST (*Automatic Speech-to-text Translation*, automatický překlad řeči do textu) CAI nástroj?

Na základě těchto otázek autoři přicházejí s vlastní definicí CAI nástrojů:

- CAI nástroj může být jakýkoli počítačový software, mobilní aplikace, či digitální zařízení,
- CAI nástroje se využívají během tlumočení, nepočítají se ty, které se používají pouze před či po samotném tlumočení,
- CAI nástroje mají snižovat kognitivní stres tlumočníků a zvyšovat jejich celkovou kapacitu ke zpracování informací.

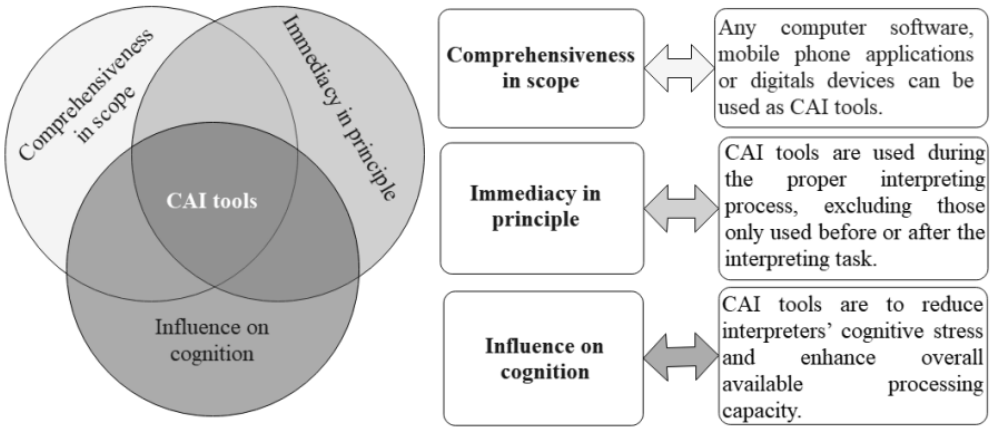
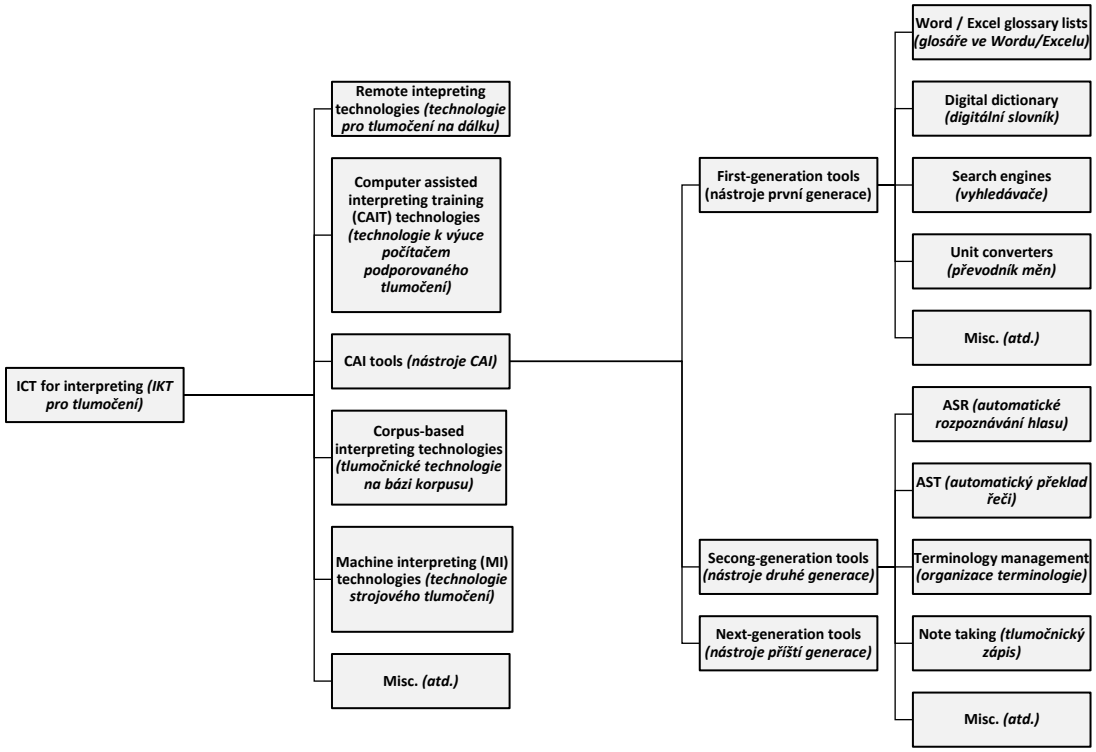


Figure 2. Main Features of CAI Tools and Explanations

Obrázek 3: V čem se musí jednotlivé funkce nástroje protínat, aby byl považován za CAI nástroj. (Guo et al., 2022)

Na základě těchto tří kritérií přichází s následujícím rozdělením CAI nástrojů:



Obrázek 4: Kategorizace CAI nástrojů. (Guo et al., 2022; vl. překlad)

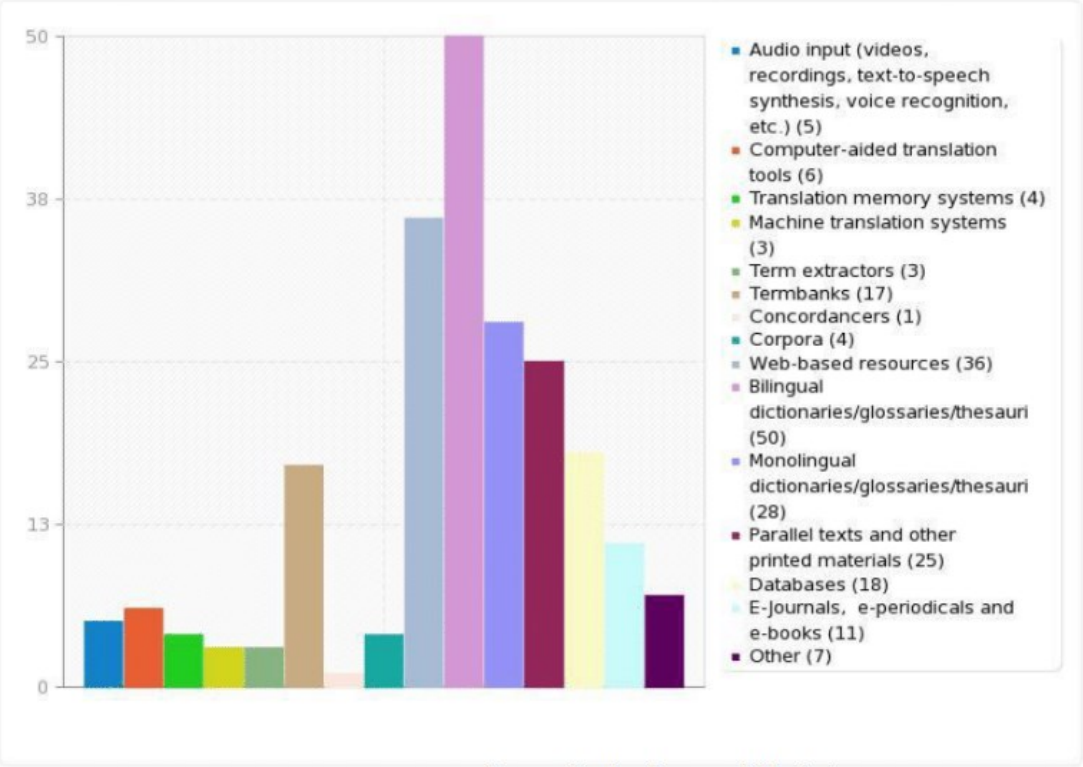
Zatímco se tedy někteří badatelé neshodnou, zda ASR patří či nepatří do CAI nástrojů (např. Defrancq & Fantinuoli, 2020, a Pisani & Fantinuoli, 2021, vnímají ASR odděleně), dle této definice lze ASR považovat za součást CAI nástrojů, což tak činíme i v této práci.

2.3.1 Obeznámenost tlumočníků s CAI nástroji a jejich využívání

Wang & Wang (2019) mluví o dvou možných dopadech CAI nástrojů: zlepšení kvality tlumočení, pokud tlumočník umí nástroje CAI efektivně využít (to podporuje také např. Defrancq & Fantinuoli, 2020, či Desmet et al., 2018), nebo zhoršení kvality tlumočení, pokud tlumočník nedokáže správně rozdělit pozornost. Zdá se tedy, že klíč k pozitivnímu dopadu CAI nástrojů tkví hlavně v obeznámenosti tlumočníků s danými technologiemi.

Kasík (2023) v dotazníkovém šetření zkoumal postoj tlumočníků na českém trhu vůči novým technologiím. Dle jeho zjištění se o ně 68 % zajímá pasivně, 22 % aktivně a 10 % vůbec. Jako

„novou technologii“ přitom tlumočníci vnímají i tlumočení na dálku, dále pak nejčastěji zmiňují AI (umělou inteligenci), DeepL (strojový překlad) a ASR (automatické rozpoznávání řeči). Co se ale týče využívání, 42 % zmiňuje, že tlumočí na dálku, 30 % žádnou technologii nepoužívá a 20 % používá Zoom a 5 % uvádí ASR. Zároveň drtivou většinu nenapadá, jaké jiné technologie by chtěli využívat, jeden respondent ovšem zmiňuje automatické titulky. Tyto poznatky podporuje také dotazníkové šetření, které provedly Pastor & Fern (2016). Ze 133 mezinárodních profesionálních tlumočníků pouze okolo 48 % používá technologie, nicméně přímo během výkonu tlumočení nejvíce z těchto zahrnuje online mono- a bilingvální slovníky a glosáře (pouze 5 jich používalo rozpoznávání řeči a tomu podobné nástroje).



Obrázek 5: Výsledky dotazníkového šetření: jaké technologie tlumočníci používají během tlumočení. (Pastor & Fern, 2016)

Kasík (2023) své šetření dále zaměřil na dedikované CAI nástroje. Vyplynulo z něj, že 48 % respondentů nezná žádné, 32 % zná automatické titulky v Zoomu, 18 % zná InterpretBank, 12 % Kudo Boothmate a 7 % Interpreter’s Help. Přitom 88 % z nich ale ani jednu z těchto technologií nevyužívá.

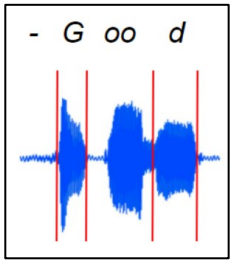
Důvody, proč dané technologie tlumočníci nevyužívají, Kasík rozvádí pomocí rozhovorů s profesionálními tlumočníky a výzkumníky. Uvádí, že jsou nástroje zatím ve vývoji a nerozšířily se na trh, tlumočníci nemají rádi změny a o technologie se příliš nezajímají, protože s nimi nemají zkušenosti. Hned několik se jich shodne, že hlavním důvodem je, že CAI technologie během tlumočení nepřinášejí přidanou hodnotu: nesnižují kognitivní zátěž, nezpříjemňují proces tlumočení, ani nezvyšují produktivitu. Pastorová a Fernová (2016) shrnují, že mezi tlumočníky panuje velká neznalost, obecná nedůvěra vůči technologiím a pocit, že je spíše vyrušují, než pomáhají.

Dle Kasíka jsou tlumočníci nejvíce obeznámení s CAI nástroje ve fázi přípravy (k vyhledávání termínů, ke zpracování glosáře apod.). Pokud by se s nimi měli naučit, hlavním kritériem pro ně bude snížení kognitivní zátěže a spolehlivost. Z dotazníkového šetření vyplynulo, že 65 % by mělo zájem o technologie, které by mohly zobrazovat či zvýrazňovat termíny, jména, čísla atd.

2.4 ASR

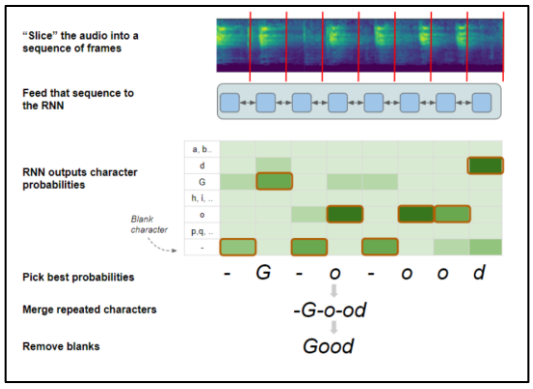
V této kapitole promluvíme o tom, co je to ASR, odkud vzniklo, jak funguje, jaké jsou jeho limitace a jaké metriky se používají k výpočtu jeho přesnosti. Této kapitole je věnována zvláštní pozornost právě proto, že se jedná o stěžejní téma empirické části.

ASR, neboli *Automatic Speech Recognition* (automatické rozpoznávání řeči), je technologie, která je schopna převést mluvenou řeč na psaný text. Lze toho docílit pomocí dvou modelů, které spolupracují: akustického modelu a jazykového modelu. Akustický model pracuje se zvukovým vstupem, který obdrží v podobě spektrogramu a následně ho rozdělí na jednotlivé hlásky. (Weingartová, 2020) Takto rozkouskovaná řeč je poté předána jazykovému modelu, který na základě pravděpodobnosti a gramatických pravidel určuje, o jaká se jedná slova. (Rev, 2022)



Obrázek 6: Zjednodušená vizualizace rozkouskování spektrogramu na hlásky. (Doshi, 2021)

Protože každý vyslovuje trochu jinak, s největší pravděpodobností bude zachycený zvuk podobný více než jedné hlásce. Akustický model proto srovná i okolní hlásky a na základě pravděpodobnosti vybere, o jakou se nejspíše jedná (např. v angličtině je pravděpodobnější, že po „*h*“ přijde „*e*“ než „*i*“). (Weingartová, 2020)



Obrázek 7: Opravdová vizualizace dekódování zvukového vstupu za pomoci CTC (Connectionist Temporal Classification, algoritmus pro trénování hlubokých neuronových sítí). (Doshi, 2021)

Řetězec hlásek je poté předán jazykovému modelu. Ten opět operuje na bázi pravděpodobnosti, tentokrát ale mezi jednotlivými slovy (např. v angličtině je pravděpodobnější fráze „*hello my name is"* než fráze „*hell oh my nay miss*"). (ibid.)

U většiny[1] systémů ASR je v současné době (červen 2024) pro každý jazyk potřeba jiný akustický model. Protože je tedy potřeba akustický model natrénovat na každý jazyk zvlášť, vede to ke zvýšeným nákladům. Navíc, akustický model bude fungovat tím lépe, čím více jazykových dat má k dispozici. U tzv. jazyků bohatých na zdroje (*resource-rich*, jedná se o jazyky, které jsou často zastoupené ve webovém prostředí, popř. pro ně lze najít digitální zdroje ve formě textových či zvukových korpusů, typicky např. angličtina, pro kterou lze dohledat až 30 000 hodin zpracovaných dat; Mirishkar, 2022) tak mívá ASR lepší výsledky než u jazyků, kde není k dispozici dostatek reprezentativních dat. (Heigold et al., 2013)

2.4.1 Historie

Technologie ASR se objevila v 50. letech minulého století, konkrétně 1952, kdy Bell Labs vytvořili technologii Audrey, která zpočátku fungovala jako rozpoznávač čísel, ale o deset let později dokázala přepsat základní slova jako „*hello*". (Foster, 2023) Od té doby ASR zaznamenalo velký vývoj, obzvláště v posledním desetiletí s příchodem umělé inteligence (*artificial intelligence, AI*) a hlubokého učení (*deep learning*). (ibid.)

V dnešní době je technologie ASR zabudovaná v chytrých telefonech, autech a dalších elektrozařízeních. Nejlepšími příklady této technologie jsou virtuální asistenti (Siri, Google Assistant, Alexa atd.), ale můžeme se s nimi setkat i na jiných platformách, například Google Meet nebo v této práci probíraný Zoom. (ibid.)

2.4.2 Přesnost a faktory ovlivňující přesnost

Již od samotného vzniku této technologie je hlavním tématem přesnost. Ta se totiž odvíjí od několika faktorů:

- vnějších (jako např. rychlost internetového připojení, kvalita mikrofonu, výslovnost mluvčího, přízvuk mluvčího apod.)
- vnitřních (např. data, na kterých byl nástroj trénován). (University of Melbourne, 2024)

[1] Ačkoli se již pracuje také na multilingválních ASR systémech, např. ve společnosti Meta, viz Tjandra et al. (2023).

Čím větší množství dat z dané oblasti, od daného mluvčího apod. má ASR k dispozici, tím přesněji dokáže zvukový vstup zaznamenat. Pokud tedy ASR trénuje na bílých mluvčích americké angličtiny mužského pohlaví, kteří mluví o politice, bude nejlépe zaznamenávat projevy z této oblasti od podobných mluvčích, v kontrastu s černošskou ženou, která bude mluvit newyorským dialektem o zahradničení.

V momentě, kdy text přepsaný ASR neodpovídá textu, který byl reálně pronesen, mluvíme o chybě. Chyby ASR mohou pocházet hned z několika zdrojů. Rozdělili jsme je do několika kategorií:

2.4.2.1 Chyby pocházející z nekvalitního zvukového vstupu

U této kategorie jsou chyby zapříčiněny nekvalitními či nedostatečnými externími podmínkami, tedy nepochází z ASR samotného, nýbrž ze zhoršené kvality zvuku. Pracovali jsme zde na základě studií Rodrigues et al. (2019), Graham & Roll (2024), Glasser et al. (2017) a také doporučení Zoomu. (Zoom, 2024a) Patří sem:

- **Ruch v pozadí**. (Zoom; Rodrigues et al., 2019)
- **Mluva několika řečníků najednou**. (Glasser et al., 2017)
- **Nekvalitní vstup**. Např. rychlost internetu, hlasitost a jasnost hlasu řečníka (Zoom), či vzdálenost od mikrofonu. (Rodrigues et al., 2019)
- **Emocionální stav řečníka**. Jedná se o křik, šepot nebo pláč. Všechno jsou to stavy, které negativně ovlivňují kvalitu zvukového vstupu. (Glasser et al., 2017)
- **Rychlost projevu**. U ASR není problém rychlost jako taková, jako spíš nepřesná výslovnost způsobená rychlejší mluvou. Řečníci totiž redukují jednotlivé fonémy, polykají koncovky nebo nenaznačují pauzy mezi slovy a větami. Pro ASR je ideální čtený projev (i se zvýšenou rychlostí), protože spontánní projev je neplynulý, vyplněný hezitacemi a dlouhými pauzami, což zhoršuje přesnost. (Graham & Roll, 2024)

2.4.2.2 Chyby vycházející ze zaujatých tréninkových dat

ASR technologie může být pouze tak kvalitní, jak kvalitní jsou data, na kterých se trénuje. Bohužel modely trénují na určitých typech projevů a řečníků více než na jiných (např. proto, že jsou snáze dostupní, nejčastěji zastoupeni apod.), a proto je poté lépe přepisuje. To následně vede

k tomu, že pokud je určitá skupina řečníků podreprezentována v trénovacích datech, ASR její mluvu přepisuje hůře. Do těchto zaujatostí patří například:

- **Zaujatost na základě pohlaví.** Ngueajio & Washington (2022) sestavily komplexní přehled studií, které se týkají zaujatosti ASR. Z výsledků jednotlivých studií vyplývá, že ASR si v přepisu mužských a ženských hlasů nevede stejně, ale neexistuje konsensus, zda si vede lépe u mužských či ženských hlasů. Podle všeho záleží na datech, na jakých ASR trénuje, ideálně by však mělo být trénované na obou stejně – a možná dokonce více na tradičně podreprezentovaných skupinách, neboť se například ukázalo, že systém, jehož tréninková data tvořily ze 70 % ženské hlasy, je stále přepisoval o něco hůře než mužské. (Garnerin, Rossato & Besacier, 2021)
- **Zaujatost na základě věku**. Věk je dalším faktorem, který hraje roli v tom, nakolik dobře dokáže ASR přepisovat řeč. Například děti a senioři jsou velmi odlišné skupiny, které nejsou v tréninkových datech dostatečně zastoupeny, navíc obě skupiny občas používají jiná slova, než jaká jsou v dané době rozšířená. Dále to pak souvisí také s celkově nekvalitním zvukovým vstupem (např. šišlání). (Feng et al., 2021)
- **Zaujatost na základě rasy a přízvuku**. Mluva černošských komunit, britská angličtina, skotská angličtina a další představují pro ASR problém, protože jsou tyto skupiny v trénovacím vzorku většinou podreprezentovány. (Graham & Roll, 2024, Koenecke et al., 2020) Můžeme do této kategorie ale také zařadit nerodilé mluvčí, protože jejich výslovnost bývá nestandardní – konkrétně se zjistilo, že čím bližší je výslovnost rodilého jazyka mluvčího cizímu jazyku (v tomto případě angličtině), tím lepších výsledků ASR dosahuje. (Graham & Roll, 2024) Zároveň by se sem dala zařadit zaujatost na základě zdravotního stavu, která je ale spojená hlavně se srozumitelností mluvčího – u řeči sluchově postižených lidí, případně lidí s dysartrií (motorická řečová porucha) vykazuje ASR podstatně vyšší chybovost. (Glasser et al., 2017; Tu et al., 2016)
- **Menší jazyky a cizí jazyky.** U menších jazyků bývá méně dostupných trénovacích dat, a proto u nich přesnost ASR bývá nižší. (Heigold et al., 2013) Akustické modely bývají navíc natrénované na specifické hlásky jednoho jazyka, a proto působí problémy, pokud se v jednom jazyce objevují hlásky z jazyka jiného. (Verbit Editoral & Opher, 2024)

2.4.2.3 Chyby vycházející z obsahu

U této kategorie se jedná o chyby pramenící z obsahu projevu, tzn. téma či používané lexikum.

- **Problémy s hranicí slov.** ASR pracuje na základě nejvyšší pravděpodobnosti, a to včetně hranic slov. Může tedy dojít ke špatné segmentaci hlásek do jiných slov, než jaká byla vyřčena. (Petkar, 2016)
- **Velikost slovníku**. ASR modely jsou trénované na specifických datech a pracují s pravděpodobností výskytu slova s ohledem na slova sousedící. To znamená, že méně častá slova nemusí být ve slovníku ASR vůbec zastoupena, a ASR je tak nemusí rozpoznat. Jedná se například o termíny (například z oblasti medicíny nebo historie) či zkratky. (Kuhn et al., 2023; Aksënova et al. 2021)
- **Nejednoznačnost slov**. Jedná se například o homofony, tedy slova, která znějí stejně, ale jinak se píšou. (Petkar, 2016) Čeština má takových slov poměrně málo, protože ve většině případů slovo vyslovujeme tak, jak ho píšeme (nejzastoupenější kategorií homofonů proto bývají slova lišící se v y/i, např. *byl* x *bil*). Angličtina oproti tomu operuje s širokou škálou homofonů, které jsou významově velmi odlišné (*flower* x *flour*). U španělštiny pak jde hlavně o hlásky b/v a tiché h (*vaca* x *baca*, *ola* x *hola*).

2.4.3 Metriky pro výpočet přesnosti automatických titulků

Nyní se budeme věnovat existujícím metrikám, pomocí kterých se měří přesnost automatických titulků. V naší práci rozdělujeme metriky od jednoduchých a automatizovatelných až po experimentální a sofistikované.

2.4.3.1 Model WER

Nejjednodušší a nejčastěji používanou metrikou je tzv. WER (*Word Error Rate*). (Fox, 2023) U výpočtu WER se v automatických titulcích každá odchylka od původního projevu považuje za chybu. To znamená, že pokud se nechá vyhotovit transkript, který 100% odpovídá projevu, lze tento transkript jednoduše porovnat s transkriptem automatických titulků vytvořených ASR.

Vzoreček k vypočítání WER je následující: $WER = \frac{S+I+D}{N} \text{x } 100$

S (*substitutions*) udává počet nahrazení slov, I (*insertions*) udává počet přidání slov, D (*deletions*) počet vynechání slov a N (*number of words spoken*) celkový počet slov v projevu.

Výsledek se pak dále vynásobí stem, aby se došlo k celkovému procentu. Každé kategorii (nahrazení, přidání, vynechání) je přiřazena stejná váha (1). Čím nižší je procento WER, tím přesnější ASR je. (Rev, 2023)

Za chybu se počítá všechno: v angličtině např. *I am* napsané jako *I'm*, interpunkce na špatném místě, nebo i špatně přepsané slovo (v češtině např. *objetí* místo *obětí*). Tím, že váha zůstává stejná u všech chyb, je nutno zvažovat relevanci výsledků – rozdílu mezi *new york* a *New York* nelze přisuzovat stejnou váhu jako vynechání slova. (Fox, 2023)

Výhodou tohoto modelu je jednoduchost a automatizovatelnost. Nevýhodou je ovšem to, že tento model počítá za vážné chyby i takové rozdíly oproti výchozímu projevu, které čtenáři nevnímají jako závažné, popřípadě jsou jimi dokonce vítané (např. vypouštění fatických výrazů).

2.4.3.2 Model NER

Nevýhody modelu WER reflektuje upravený model zvaný NER (*Number, Edition and Recognition error*), který navrhli Romero-Fresco & Martínez (2015). Tento model operuje s rozdílnou váhou chyb. Původně byl navržen pro hodnocení tzv. *live subtitles* („živé titulky") od respeakerů (viz kapitola 2.1 *Titulky*), ačkoli do budoucna ho autoři vidí jako užitečný i při hodnocení automatických titulků. Protože u respeakerů je stroj natrénovaný na hlas daného člověka, nemusí se potýkat s různými proměnnými, které mají jinak vliv na jeho výkon: přízvuk, rychlost, nekvalitní vstup a další. Respeaker zároveň může opravovat problémy jako čárky či interpunkce a dělá takové úpravy, aby zachoval původní smysl projevu. Pokud je projev velmi hutný, používá respeaker strategie velmi podobné těm tlumočnickým – vynechávání nepodstatných informací či fatických výrazů, kondenzování informace apod. (AI Media, 2023a) WER model ovšem tyto vynechávky a úpravy počítal za chyby, ačkoli z hlediska čtenářského pohodlí a smyslu byly vítané a správné. Proto byl NER model navržen, aby skutečně zkoumal, nakolik výstup respeakerů odpovídá původnímu projevu bez penalizace adekvátních úprav.

V rámci výpočtu NER se rozlišují tyto typy chyb či změn:

- *Edition errors* (nesprávná úprava, vl. překl.). Jedná se o chyby pramenící ze špatného rozčlenění původního projevu. U respeakerů tato kategorie zahrnovala chyby, kdy respeaker sám špatně pochopil myšlenku, nesprávně něco vynechal, nebo naopak přidal.

Pro automatické titulkování Romero-Fresco & Martínez zmiňují hlavně špatně zapsaná velká písmena, interpunkci či rozlišení mluvčího.

- *Recognition errors* (nesprávné rozpoznání, vl. překl.). Jedná se o chyby pramenící z nesprávného zapsání původního projevu. U respeakerů se jednalo hlavně o chyby technologie samotné, když špatně „slyšela“. V případě automatického titulkování je toto hlavní typ chyb.
- *Correct editions* (správná úprava, vl. překl.). Jedná se o speciální kategorii, která byla zavedena pro případy, kdy respeaker sice vynechal slova či informace, ale nemělo to žádný dopad na úplnost obsahu. U plně automatických titulků se neočekává, že by k těmto případům docházelo.

Zároveň autoři zdůrazňují poslední krok modelu, a sice slovní hodnocení – např. jaké strategie respeaker používal, zda byly úspěšné, zdá nestrávil příliš mnoho času opravami apod.

Vzoreček k výpočtu NER je následující: $NER = \frac{N-E-R}{N} \text{x } 100$

N (*number of words in the spoken text*) představuje celkový počet slov, E (*edition errors*) představuje celkovou váhu chyb pramenících z nesprávného roztřídění, R (*recognition errors*) představuje celkovou váhu chyb pramenících z nesprávného rozpoznání. Výsledek se vynásobí stem, aby se došlo k procentu. V rámci výpočtu celkového E a R se rozlišují tři úrovně chyb, kterým model přisuzuje odlišnou váhu:

1. *Minor error* (malá chyba, vl. překl.; váha 0,25)

   Tato chyba je lehce odhalitelná, nevede k nepochopení obsahu. Může se jednat např. o špatnou pádovou koncovku či jinou gramatickou chybu. Případně může jít o vynechání vedlejší věty / větných členů. Dojde sice k nepřesnému předání obsahu, ale nejedná se o příliš rušivou chybu, lze ji snadno odhalit. Podobně je to také s velkými a malými písmeny, v případě angličtiny pak např. rozdíl mezi *were* a *we’re*, *for* a *four*, nebo ryze gramatické chyby jako *what you do then* místo *what do you do then*.
2. *Standard error* (standardní chyba, vl. překl.; váha 0,5)

   Tato chyba se dá stále odhalit, ale je to o něco složitější. Může se jednat např. o špatnou interpunkci, když dojde k rozdělení jedné větu na dvě, které samostatně nedávají úplně

smysl. Navíc se mezi standardní chybu započítávají vynechání hlavních vět / větných členů, bez kterých vedlejší věty / větné členy postrádají smysl a relevanci. Z hlediska nesprávného rozpoznání se může jednat o těžce pochopitelné zapsání: v angličtině např. *paid in full by pizza* místo *paid in full by Visa*, *he's a buy you a bull asset* místo *he's a valuable asset*

3. *Serious error* (vážná chyba, vl. překl.; váha 1)
   Tato chyba mění obsah sdělení a nelze ji vyvodit, může proto vést ke špatnému pochopení. Může se jednat např. o vynechaní klíčového slova, které mění smysl. Důležité kritérium je, že chyba v kontextu může dávat smysl, a proto u čtenáře nevzbudí podezření. V angličtině například *he's having problems with the cheques* místo *he's having problems with the Czechs*; *he was born in 1986* místo *he was born in 1996*.

Průměrné přijatelné skóre při používání modelu NER na *live subtitles* (tzn. u zapojení respeakerů) je přesnost 98 % a výše. Automatické titulky by proto měly splňovat stejné požadavky, pokud mají nahradit práci respeakerů.

Nevýhodou NER je, že k výpočtu je potřeba lidské asistence (pokud by se tato metoda měla automatizovat, lze přikročit k použití jazykových modelů jako např. ChatGPT, to je ale zatím v počáteční fázi). (Hughes, 2022) Rozdělení chyb podle závažnosti navíc může být subjektivní.

2.4.3.3 Model Semantic-WER, ACE model a další

Mezi vysoce sofistikované metriky lze řadit takové, které počítají např. se sémantickou vzdáleností mezi původním slovem a chybně rozpoznaným slovem či zda byla chybně zapsaná slova klíčová pro porozumění. Mezi tyto metriky by patřily například Semantic-WER (Roy, 2021) nebo WKER (*Weighted Keyword Error Rate*; Kawahara & Nanjo, 2005) a další. Tyto metriky pracují s korpusy, jazykovými modely a pravděpodobnostními výpočty. Model ACE (*Automated Caption Evaluation*; Kafle & Huenerfauth, 2017) například analyzuje sémantickou vzdálenost mezi původním slovem a automatickými titulky, tedy nakolik se titulky blíží původnímu významu slova, a dále počítá pravděpodobnost výskytu slova, tedy jak často se dané slovo na daném místě vyskytuje a nakolik tedy člověk dokáže případnou chybu v titulcích pochopit a opravit. Z toho se pak pokouší spočítat, nakolik jsou titulky i přes svou chybovost

pochopitelné. Díky sémantické analýze tak bere v potaz více „pochopitelnost“ titulků než jejich konkrétní přesnost.

Nevýhoda těchto modelů je, že veškeré výpočty, které v sobě zahrnují práci s korpusem, se většinou dělají na angličtině, pro kterou existuje mnoho kvalitních korpusů s různými zaměřeními – to samé se bohužel nedá říci o jiných jazycích. Navíc jsou výpočty pro komerční využití příliš složité a vyžadují zahrnutí jazykových a IT odborníků.

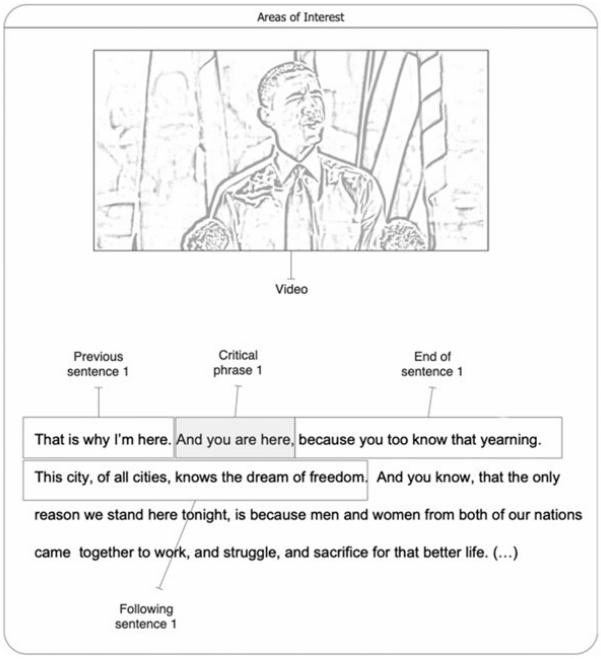
2.5 Tlumočení a ASR

V této kapitole propojujeme poznatky z oblasti tlumočení a ASR. Zaměřujeme se na to, jak automatické titulky ovlivňují výkon tlumočníků, a to jak pozitivně, tak negativně.

2.5.1 Kognitivní zátěž u tlumočení při využívání ASR

Při zapojení automatických titulků u tlumočení (např. na Zoomu) dochází k propojení dvou disciplín, u kterých je prokázána vyšší kognitivní zátěž (viz 2.2.3 *Kognitivní zátěž u tlumočení*): tlumočení videokonferencí a simultánního tlumočení s textem, který ale tlumočník nemá předem k dispozici.

Jak tlumočníci pracují s přepisem projevu, který vidí před sebou, zkoumali Seeber et al. (2020). Experiment se zaměřoval na fixaci pohledu tlumočníků (kde ulpívá tlumočník pohledem). Tlumočníci měli k dispozici přepsaný text projevu, který viděli na obrazovce před sebou spolu s videozáznamem řečníka. V experimentu byl psaný text rozdělen na tři části: předchozí větu, kritickou větu a následující větu. Kritická věta byla ta, kterou mluvčí právě vyslovoval.



Obrázek 8: Rozhraní experimentu s fixací pohledu tlumočníků. (Seeber et al., 2020)

Experiment prokázal, že tlumočníci se pohledem nejdéle fixovali na předchozí větu. Dle autora nejspíše proto, aby tím ulevili své krátkodobé paměti. Na následující větu se fixovali pouze v 10 % případů.

Jeho poznatky jsou částečně podpořeny také experimentem, který provedli Zou et al. (2022). V rámci výzkumu měli profesionální tlumočníci simultánně tlumočit projev, ke kterému dostali i text promítaný na obrazovce. Výzkumníci srovnali pohled tlumočníků a rychlost přetlumočení informací a vytvořili několik kategorií tlumočníků: ti, u kterých byl dominantní zvukový vstup (ED, *ear-dominant*), ti, u kterých byl dominantní vizuální vstup (ID, *eye-dominant*), a ti, u kterých byly vstupy vyvážené (EIB, *ear-eye-balanced*). ID (vizuální) tlumočníci produkují přesnější tlumočení, zatímco u ED (sluchoví) tlumočníků je tlumočení plynulejší. U EIB tlumočníků se ukázalo, že snaha vybalancovat obojí může vést ke snížení kvality tlumočení.

Dále se na plynulost versus přesnost při využívání technologií zaměřili Wang & Wang (2019). Pracovali sice s konsekutivním tlumočením (tedy takovým tlumočením, kdy tlumočník tlumočí až po vyslechnutí určitého celku ve výchozím jazyce), ale zapojili do něj právě ASR a MT (*Machine Translation*, strojový překlad). Výchozí projev byl nejprve zachycen technologií ASR (konkrétně *Dragon Anywhere*) a následně výzkumníkem manuálně vložen do nástroje Google Translate. Takto vyhotovený překlad byl poté zobrazen tlumočníkům na obrazovce a měli ho k dispozici během konsekutivního tlumočení. Rozdělili tlumočníky na dvě skupiny a připravili dva projevy: první skupina tlumočila první projev se strojovým překladem a druhý projev bez strojového překladu; druhá skupina naopak.

Table 6: Interpreting performance for E1

Interpreting with reference			Interpreting without reference		
Participants	**accuracy**	**fluency**	**Participants**	**accuracy**	**fluency**
CI01	3.0	3.5	CI02	3.7	3.0
CI03	3.7	3.8	CI04	1.3	2.8
CI05	3.7	3.8	CI06	4.7	4.8
CI07	4.9	3.5	CI08	2.7	3.5
CI09	2.7	3.0	CI10	3.0	3.0
Average	**3.6**	**3.52**	**Average**	**3.08**	**3.42**

Table 7: Interpreting performance for E2

Interpreting with reference			Interpreting without reference		
Participants	**accuracy**	**fluency**	**Participants**	**accuracy**	**fluency**
CI02	3.7	3.0	CI01	2.7	4.0
CI04	2.4	2.0	CI03	2.9	3.5
CI08	1.0	4.3	CI05	3.7	3.8
CI10	3.3	3.3	CI06	5.0	4.8
			CI07	3.3	4.0
			CI09	2.7	3.0
Average	**2.60**	**3.15**	**Average**	**3.38**	**3.85**

Obrázek 9: Výsledky dvou skupin (with reference = se strojovým překladem, without reference = bez strojového překladu) u dvou projevů (E1 a E2). (Wang & Wang, 2019)

Z jejich výsledků vyplynulo, že v prvním projevu byla skupina se strojovým překladem plynulejší i přesnější než skupina bez něj; ale při druhém projevu byla skupina se strojovým překladem naopak méně plynulá a přesná než skupina bez něj. Dle všeho byly vylučující se výsledky důsledkem špatného rozdělení skupin: v jedné skupině byli zkušenější tlumočníci, kteří lépe zvládali kognitivní zátěž u přidání strojového překladu, ale zároveň dokázali dobře tlumočit i bez něj. Ačkoli byl ale experiment proveden na konsekutivním tlumočení a se strojovým překladem, jsou jejich výsledky relevantní i pro naše účely, neboť dokazují, že při zapojení technologií může mít zkušenost tlumočníka dopad na jeho výkon.

Pokud je tedy ASR přesná, může tlumočníkův výkon vylepšit, a to hlavně na poli přesnosti. Na druhou stranu ale přidaný text zvyšuje kognitivní nároky na tlumočníka, což může vést ke zhoršení plynulosti a teoreticky k celkovému zhoršení kvality, pokud tlumočník neumí s dostupnými zdroji informací zacházet (Pisani & Fantinuoli, 2021). Všechny tyto studie ale opomíjí důležitý fakt: ASR se může i mýlit.

Na dopad nepřesností v přepisu textu na tlumočnický výkon se zaměřili Chmiel et al (2020). V jedné verzi veškerá jména, čísla a obsahová slova odpovídala řečenému, v druhé nikoli. Dle jejich výsledků u verze se správným přepisem nedošlo k signifikantnímu zlepšení celkové kvality tlumočení, došlo ale k nižší chybovosti u převodu čísel. U nesprávného přepisu pak pomocí délky fixace pohledu zjistili, že tlumočník musí vynakládat větší kognitivní úsilí na pochopení informace v případech, kdy se vizuální a zvuková stopa lišily, což může mít na kvalitu tlumočení negativní vliv.

O detailní prozkoumání tlumočnického výkonu při využívání konkrétně automatických titulků se pokusili Yuan & Wang (2023). Experiment proběhl na platformě Zoom. V rámci experimentu byl projev rozdělený na dvě poloviny: v jedné byly automatické titulky zapnuté, ve druhé ne. Zároveň každá polovina obsahovala 8 „míst zájmu“ (*points of interest*): 4 věty, které obsahovaly čísla a jména, a 4 věty, které je neobsahovaly. Obrazovka byla rozdělena do dvou oblastí: obličej mluvčího a automatické titulky. Výzkumníci se zajímali o to, jak moc tlumočníci přebíhají pohledem mezi těmito oblastmi, kde spočívají pohledem a jak dlouho, a to v závislosti na tom, zda se ve větě objevují čísla a jména, či nikoli.

Použitým materiálem byl článek z časopisu *The Economist* upravený rodilým anglickým mluvčím tak, aby působil jako mluvený projev. Všichni účastníci experimentu měli za sebou alespoň 6 měsíců studia simultánního tlumočení a jejich mateřským jazykem byla čínština, nejednalo se ale o profesionální tlumočníky.

Yuan & Wang se soustředili na 3 výzkumné otázky s následujícími hypotézami:

1. Jak tlumočníci zpracovávají vizuální informace poskytované automatickými titulky?
   Hypotéza 1a: Když se ve větě objeví jména a čísla, budou tlumočníci přebíhat pohledem mezi obličejem a automatickými titulky méně (oproti tomu, když ve větě čísla a jména nejsou).
   Hypotéza 1b: Když se ve větě objeví jména a čísla, tlumočníci setrvají pohledem déle na automatických titulcích.
2. Na jaký typ informací vynaloží tlumočníci více kognitivního úsilí?
   Hypotéza 2a: Tlumočníci se budou pohledem fixovat častěji na věty, které obsahují čísla a jména, než na ty, která je neobsahují.
   Hypotéza 2b: Tlumočníci se budou pohledem fixovat déle na věty, které obsahují čísla a jména, než na ty, která je neobsahují.
3. Mění se nějak zásadně přesnost tlumočení při tlumočení s automatickými titulky a bez nich?
   Hypotéza 3: Díky automatickým titulkům bude tlumočení čísel a jmen přesnější.

Hypotéza 1a byla vyvrácena: ať už se ve větě objevila čísla a jména, či nikoli, tlumočníci přebíhali pohledem mezi obličejem a automatickými titulky stejně často. Podle výzkumníků je to proto, že kromě jmen a čísel se v automatických titulcích objevují i další slova, která ne vždy odpovídají řečenému, což tlumočníky nutí pohled přesunout jinam, aby tomuto nesouladu nevěnovali příliš svojí mentální kapacity a nezvyšovali tak riziko jejího zahlcení.
Hypotéza 1b se potvrdila: u vět s čísly a jmény tlumočníci na automatických titulcích průměrně setrvají pohledem déle než na obličeji mluvčího.
To znamená, že tlumočníci celkově přebíhají pohledem mezi titulky a obličejem stejně často, nezávisle na tom, zda se objevila čísla a jména. Když už se ale čísla a jména objeví, setrvají pohledem na titulcích o něco déle než na obličeji mluvčího.

Hypotéza 2a byla potvrzena: z hlediska počtu fixací se tlumočníci pohledem fixovali častěji na čísla a jména, což vypovídá o vyšším kognitivním úsilí vydaném na zpracování tohoto typu informace. Také to vypovídá o tom, že tlumočníci tento typ informací aktivně vyhledávají. Hypotéza 2b se ale nepotvrdila: z hlediska času stráveného fixací pohledu se tlumočníci pohledem nefixovali na čísla a jména déle než na věty, které je neobsahovaly. To vyvrací předchozí zjištění (Korpal & Stachowiak-Szymczak, 2018). Autoři to vysvětlují tím, že v předchozích experimentech se využívalo statické prezentace ve formě slidů, zatímco Zoom vytváří titulky dynamicky a text tak neustále plyne, proto nemají tlumočníci čas se číslům a jménům věnovat déle. To dokazuje i Defrancq & Fantinuoli (2020): dynamický charakter zobrazování titulků je pro tlumočníky vyrušující. V jejich vlastním experimentu jim tlumočníci nevěnovali pozornost ani v případech, kdy titulky mohly využít.

Hypotéza 3 se potvrdila: přítomnost titulků snížila chybovost tlumočení o 30 %, což je v souladu i s dalšími výzkumy (Defrancq & Fantinuoli, 2020; Desmet et al., 2018).

2.6 V čem mohou automatické titulky tlumočníkům pomoci

Jak jsme popsali v kapitole o simultánním tlumočení, jedná se o činnost, která představuje pro tlumočníky vysokou kognitivní zátěž. Existují ovšem specifické součásti projevů, které jsou pro tlumočníky těžší než jiné – Gile (2009) mezi tyto spouštěče problémů (v orig. *problem triggers*, vl. překlad) řadí např. jména, čísla, akronymy a výčty.

Pokud si představíme virtuální konferenci, která využívá ASR, můžeme očekávat, že právě v těchto případech budou tlumočníci vyhledávat pomoc u kolegů či právě automatických titulků. Je tedy vysoce důležité, aby se v těchto kategoriích ASR nezmýlilo.

Fantinuoli (2017) přidává, že ASR může tlumočníkům pomoci také v případě terminologie, nicméně pouze v případě, že do ní má tlumočník šanci zasáhnout tak, jako je tomu v případě jeho vlastního CAI nástroje InterpretBank Digital Boothmate, který pracuje s tlumočníkovým předem vytvořeným glosářem. U Zoomu či dalších komerčních platforem ale nic takového (zatím) neexistuje.

Rozebereme si tedy kategorie, které jsou pro tlumočníky problematické, a ve kterých jim ASR může pomoci během samotného tlumočení. Jedná se o tyto:

- Čísla
- Jména, zkratky, akronymy
- Termíny
- Výčty
- Negace

2.6.1 Jména a názvy

Podle Gila (2009) jsou jména a názvy pro tlumočníky problematické, protože jsou málo redundantní, často krátké (což znamená, že sebemenší pokles pozornosti může vést ke ztrátě informace) a navíc se hlásky mohou velmi podobat a splývat (např. u jména *Jim Joseph*).

Meyer (2008) srovnával předání jmen a názvů mezi simultánním a konsekutivním tlumočení. Jednalo se o opravdová tlumočení pro neziskovou organizaci, jejíž brazilská mluvčí navštívila tři německá města (Berlín, Hamburg a Heidelberg), ve kterých přednášela na stejné téma a byla

tlumočena (2x konsekutivně, 1x simultánně). U konsekutivního tlumočení tlumočil vždy pouze jeden tlumočník (v každém městě jiný), oba s rodilou němčinou a portugalštinou jako jazykem B (tzn. jazyk, u kterého tlumočník může tlumočit oběma směry, z i do daného cizího jazyka). Jeden z nich zároveň později tlumočil řečnici i simultánně. U simultánního tlumočení byly najaty dva týmy profesionálních tlumočníků (celkem 4 tlumočníci s rodilou němčinou a portugalštinou jako jejich jazykem B, tři zvyklí na evropskou portugalštinu, jeden zvyklý na brazilskou portugalštinu), které tlumočily stejný projev ve stejný čas (celkem byla tedy zkoumána 2 konsekutivní tlumočení a 2 simultánní). Tlumočníkům byl poskytnut přepis projevu předem, ale řečnice od textu často odbíhala. Jména a názvy, které se vyskytly v tlumočení, byly následně porovnány se jmény a názvy v příslušném výchozím projevu.

Byly stanoveny 4 kategorie jmen: jména osob, názvy institucí, místní názvy a produktová jména. V projevu v Berlíně se objevilo 57 jmen, v Hamburgu 63 jmen, v Heidelbergu 81 jmen. Výzkumníci identifkovali 5 typů předání:

1. N = předání jména (občas s minimální fonetickou úpravou)
2. Ø = vynechání jména
3. Pro = nahrazení jména zájmenem
4. Gen = generalizace jména, kdy tlumočníci nahradili jméno jiným podstatným jménem (pův. *Soja Roundup-Ready -> „jejich produkt“)*
5. N+ = specifikace jména, kdy tlumočníci předali jméno, a navíc přidali dodatečné informace pro cílového posluchače (pův. *Národní Kongres -> „Brazilský Národní Kongres“).*

V Berlíně a Hamburgu se jednalo o konsekutivní tlumočení, v Heidelbergu byla dvě simultánní tlumočení. Z výsledků nevyplývá signifikantní rozdíl mezi simultánním a konsekutivním tlumočení.

	N	Ø	Pro	Gen	N+
Heidelberg 09 (n=81)	38 (47%)	10 (12%)	6 (7%)	10 (12%)	17 (21%)
Heidelberg 10 (n=81)	60 (74%)	2 (2%)	6 (7%)	5 (6%)	8 (10%)
Hamburg (n=63)	33 (52%)	5 (8%)	3 (5%)	5 (8%)	17 (27%)
Berlin (n=57)	34 (60%)	4 (7%)	1 (2%)	5 (9%)	13 (23%)

Obrázek 10: Procentuální výsledky u převodu jmen a názvů pro jednotlivé tlumočnické strategie napříč čtyřmi různými tlumočeními. Meyer (2008)

Vidíme, že tlumočníci volí různé strategie, přičemž k předání celého jména (ať už s nebo bez dodatečné informace) došlo průměrně v 81 % případů při KT a 76 % při ST. Průměrně v 7,5 % případů došlo k vynechání u KT a 7 % při ST. K nahrazení (ať už zájmenem nebo generickým podstatným jménem) došlo průměrně v 12 % případů při KT a 16 % při ST.

Pokud by ASR bylo schopné tlumočníkům jména, názvy a zkratky přepisovat, mohla by se snížit kognitivní zátěž, kterou musí tlumočníci vynaložit na jejich převod a úspěšnost převodu by se mohla ještě zvýšit. Konkrétně Fantinuoli (2017) v této oblasti vidí velký potenciál ASR.

Zatímco je pravda, že pokud se jedná o častá, domácí jména daného jazyka, zachytává je ASR dobře, její výkon se ale rapidně zhoršuje, pokud se jedná o jména neobvyklá či dokonce cizí. Gaido et al. (2021) ve své studii zkoumají přesnost automatického překladu, který vychází z ASR. V jejich studii se přesnost u převodu jmen pohybuje kolem 40 %. Protože zkoumaná jména není v mnoha případech potřeba překládat, lze nízkou úspěšnost při převodu přičítat právě špatnému výkonu ASR. Je to kvůli tomu, že buď daná jména nejsou součástí trénovacích dat, obsahují hlásky jiného jazyka, než na jaký byl akustický model ASR natrénován, případně je mluvčí vyslovuje jinak, než jak by vyslovena být měla (protože např. není rodilým mluvčím jazyka, z něhož jméno pochází). ASR se dopouští chyb jako špatné vyhláskování, nahrazení jiným jménem, převedení na jiná slova (podstatná jména, slovesa apod.), či kompletní vynechání. (Gaido et al., 2022)

2.6.2 Zkratky a akronymy

Co se týče zkratek a akronymů, pro tlumočníky mohou představovat podobný problém jako jména, protože jsou krátká a nebývají redundantní. (Gile, 2009; Mankauskienė, 2018) Z hlediska ASR se nám nepovedlo dohledat studie, které by se zaobíraly přesností ASR u zapisování zkratek. O problematice krácení u ASR pojednávají Katsumaru, Komatani, Ogata & Okuno (2008), ti se ale ve své studii zaměřují na zkrácená slova (studie probíhala na japonštině, která má jiná morfologická pravidla než angličtina), nikoli zkratky (tvořené z písmen) jako takové. V úvodu ale vysvětlují, že krácení slov jakéhokoli druhu bývá pro ASR problematické. V tradičních systémech musí vývojáři zkratky manuálně přidávat, čímž ale naroste okruh slovní zásoby, s níž ASR pracuje. To může vést k vyšší chybovosti, neboť ASR funguje na základě pravděpodobnosti: pokud má více slov, ze kterých vybírat, může se častěji splést.

2.6.3 Čísla

Pisani & Fantinuoli (2021) shrnují, že podle studií se míra pochybení u čísel mezi tlumočníky pohybuje mezi 21 a 70 %. Vysoké rozpětí přičítají různým faktorům: zkušenostem tlumočníků, typu textů, složitosti textů, vyhodnovacím metrikám (zaokrouhlení čísla je někdy počítáno jako chyba, jindy ne), ale také, zda výzkumy probíhají v experimentálním prostředí, nebo během reálného tlumočení pro klienta. Autoři zmiňuji, že klíčovou roli v převedení čísel hraje možnost číslo vidět vizuálně. Například, pokud si tlumočníci mohou dělat během tlumočení poznámky, chybovost se sníží o 10 % (Mazza, 2001). Pokud mají k dispozici dokumenty, chybovost se sníží o 50 % (Lamberger-Felber, 2001).

Desmet et al. (2018) zjistili, že pokud tlumočníci (v tomto případě studenti) vidí čísla, která se objevují na obrazovce a jsou synchronizovaná s projevem řečníka, dělají při předání čísel o dvě třetiny méně chyb, a to i v případě, že nevidí, k čemu číslo patří. Za největší přínos je považován fakt, že tlumočníci se dopouštějí mnohem méně často vynechávání čísel či přibližných odhadů. Je tedy jasné, že vidět číslo před sebou zlepšuje tlumočníkův výkon při jeho předávání.

Pisani & Fantinuoli (2021) rozebírají CAI nástroj InterpretBank Digital Boothmate, který má velký potenciál tlumočníkům pomáhat právě v oblasti čísel. Tento nástroj dokáže zachytit čísla a ukázat je pak společně s tím, co následuje přímo po nich (např. měna, podstatné jméno apod.). Navíc nástroj dokáže převádět slova na číslice (například *padesát dva* na *52*) a naopak – miliony

či miliardy se objeví slovně. Během této studie Fantinuoli svůj nástroj teprve testoval, konkrétně na magisterských studentech bez předchozích zkušeností s CAI nástroji. I tak zaznamenal snížení chybovosti při tlumočení čísel o 25 %.

Co se týče přesnosti ASR používané nástrojem InterpretBank při zachycování čísel, Defrancq & Fantinuoli (2020) zaznamenali 96% přesnost. Zdůrazňují, že ASR chybovala v případech, kdy se mluvčí opravoval (konkrétně když číslo začal vyslovovat, zaseknul se, a poté číslo zopakoval znova: *four thou– four thousand five hundred*), případně u homofonie (*2000 and 2* místo *2002*; *to* místo *2*). Překvapivě z experimentu nevyplynulo, že by se tlumočníci v případech komplexních čísel obraceli na ASR výrazně více (na rozdíl od Desmeta et al., 2018). Podle komentářů autorů to tak může být proto, že zkoumaní tlumočníci byli studenti a nebyli s daným CAI nástrojem seznámeni dopředu. Alternativně také možná přečtení informace na obrazovce přesahovalo jejich dostupnou kapacitu, kterou museli napnout na zachycení komplexního čísla.

Zároveň ale z tohoto experimentu vyplynuly dva další fakty, a sice že skupina, která mohla v prvním kole využívat ASR a ta jim byla následně v druhém kole odebrána, si vedla hůře než skupina, která ASR používat nemohla ani v jednom z kol. To naznačuje, že se tlumočníci velmi lehce mohou stát na ASR závislí. Zároveň ale výsledky ukazují, že skupina, která ASR k dispozici měla, si vedla lépe v předávání čísel, ať už pomoc ASR využila, nebo ne (autoři to zjišťovali pomocí sledování pohledu tlumočníků). Z tohoto zjištění nelze zatím dělat žádné závěry, snad kromě spekulace, jak říkají sami autoři, že přítomnost ASR dodává tlumočníkům sebevědomí, že se mají o co opřít, což vede ke zlepšení psychického stavu, a to následně zlepšuje jejich celkový výkon.

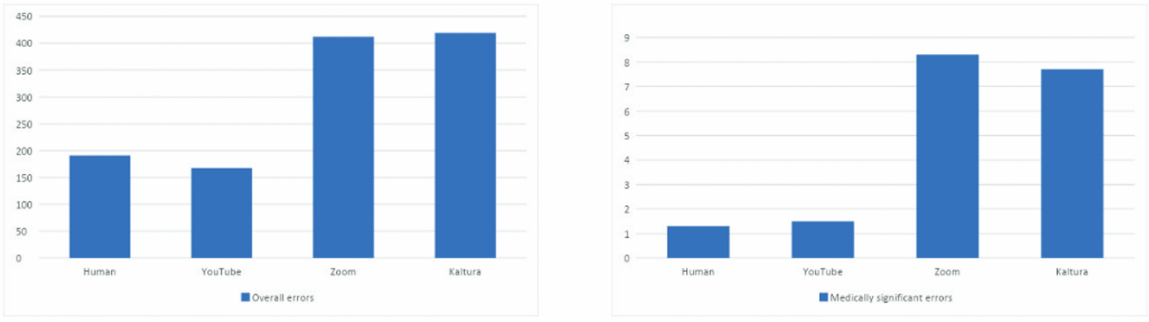
2.6.4 **Termíny**

Terminologie je mezi tlumočníky obecně považována za jeden z největších problémů. Mankauskienė (2016) ve své předběžné studii na téma spouštěčů problémů zkoumala 9 začínajících a 5 zkušených tlumočníků. Co se týče termínů, zjistila, že termíny jako *Tobin tax* (Tobinova daň), *urban renewal* (obnova měst) a *crowdfunding* (skupinové financování) nepřetlumočila více než polovina tlumočníků.

Problém s přetlumočením termínů nicméně nespočívá v tom, že by je tlumočníci neslyšeli: každý tlumočník se musí na projevy, které bude tlumočit, terminologicky připravit. (Pöchhacker, 2004) Pokud tedy termíny uvidí, ale nebude je znát, ASR mu nepomůže. Proto lze v této kategorii mluvit hlavně o snížení kognitivní zátěže tím, že tlumočníci vidí termín před sebou, mohou tedy ulevit své krátkodobé paměti (Seeber et al., 2020, prokázal, že pokud tlumočníci vidí text projevu, soustředí se na to, co právě tlumočí, aby ulevili paměti).

Nicméně, Fantinuoli (2017) vytvořil nástroj InterpretBank Digital Boothmate, který je schopen pomáhat i s termíny, přičemž v procesu částečně zapojuje právě ASR. Tento nástroj pracuje s tlumočníkovým vlastním, předpřipraveným glosářem. ASR z projevu vytváří text, který je pak dále „tříděn". Nástroj hledá shody mezi textem a glosářem, který tlumočník do nástroje předem nahrál. V momentě, kdy je objeví, navrhne tlumočníkovi pouze daný termín a překlad, který si tlumočník sám do glosáře vložil. Důležité je ale říct, že Fantinuoliho nástroj používá tzv. „*fuzzy match*" – tedy že termín přepsaný pomocí ASR a termín v glosáři nemusí být shodné, pouze dostatečně podobné. Díky tomu InterpretBank dokáže i některé chyby ASR „napravit".

Jak jsme ale vysvětlili v kapitole ASR, obecná ASR naopak většinou ve specifických termínech chybuje. Kustritz et al. (2023) srovnávají automaticky vytvořené titulky různými platformami (Kaltura, YouTube a Zoom) a titulky vytvořené člověkem. Jednalo se o veterinární kurz zaměřený na reprodukci zvířat, přičemž se zkoumala přesnost přepisu přednášek několika hostujících řečníků. Jakožto chyby se započítávala vynechaná slova, přidaná slova a špatně zapsaná slova. Autoři rozlišovali váhu chyby mezi obecnou chybou a medicínsky signifikantní chybou (kvůli které se studenti naučí slovo nesprávně). Výsledky z 25minutové lekce jsou následující:



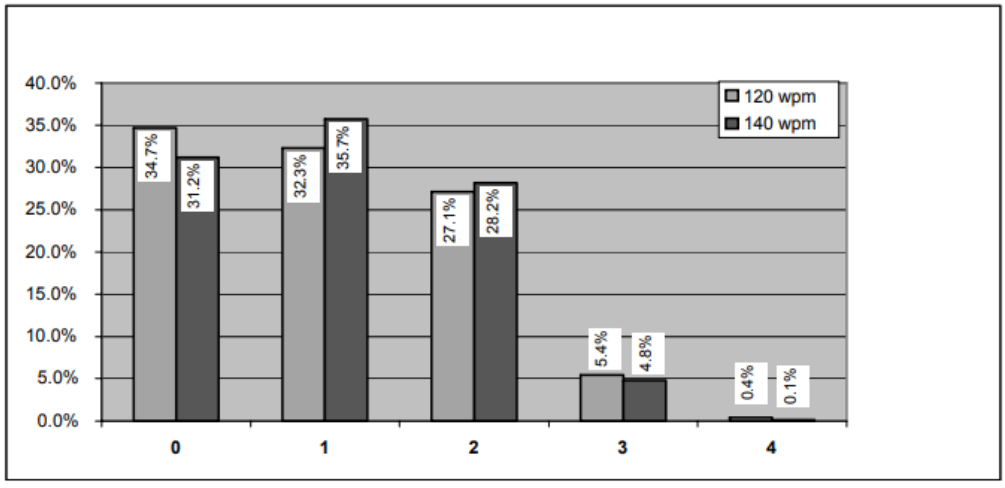
Obrázek 11: Počet obecných chyb a medicínsky signifikantních chyb mezi různými platformami a člověkem. Kustritz et al. (2023)

Jak vidíme, ASR na Zoomu a Kaltuře se dopustilo mnohem více obecných i medicínsky signifikantních chyb.

2.6.5 Výčty

Další oblastí, ve které může ASR potencionálně pomoci tlumočníkům, jsou výčty a negace. Gile (2009) zmiňuje právě výčty jako spouštěče problémů u tlumočníků. ASR v této oblasti může ulevit tlumočníkově krátkodobé paměti.

Schlesinger (2003) zkoumala kapacitu tlumočníkovy paměti právě na výčtech, konkrétně na podstatných jménech, která jsou rozvedena čtyřmi přídavnými jmény. Její výzkum probíhal mezi dvěma jazyky: angličtinou (ve které přídavná jména předcházejí podstatnému jménu, které rozvíjí) a hebrejštinou (ve které stojí podstatné jméno jako první následované rozvíjejícími přídavnými jmény). Zkoumaní profesionální tlumočníci (všichni s hebrejštinou jako mateřským jazykem) si tedy museli výčet anglických přídavných jmen zapamatovat, zatímco čekali na podstatné jméno, kterým v hebrejštině museli začít. Schlesinger zkoumala rozdíly ve výkonu tlumočníků podle rychlosti projevů (120 a 140 slov za minutu), nicméně rychlost projevu se neukázala být natolik signifikantní, jak se domnívala. Co je ale pro naše účely důležité, jsou celkové výsledky: z celkových 1920 případů (60 podstatných jmen s 4 přídavnými jmény × 16 tlumočníků × 2 různé rychlosti) byla všechna čtyři přídavná jména zachována méně než v 1 % případů. Tlumočníci zachovávali převážně žádné či jedno přídavné jméno.



Obrázek 12: Procentuální rozpětí podle toho, kolik bylo zachováno přídavných jmen (0-4), v závislosti na rychlosti přednesu (120 a 140 slov za minutu). Schlesinger (2003)

V této oblasti tedy můžou být titulky tlumočníkům nápomocné, protože mohou výčty v těchto případech dohledat a vidět před sebou.

2.6.6 Negace

Anekdoticky jsou negace také častým problémem u tlumočníků, ačkoli je žádná studie nezmiňuje konkrétně jako spouštěč problémů. Částečně lze však vysvětlení hledat u Gila (2009), který zmiňuje právě krátká a neredundantní slova jako potenciálně problematická. Z tohoto hlediska lze mezi taková slova počítat i zápor, protože jakýkoli pokles v pozornosti může vést k jeho kompletnímu přeslechnutí, což povede k velmi závažné chybě v přetlumočení. U negací proto ASR může fungovat jako kontrolní mechanismus, aby se tlumočník ujistil, že tlumočí správně.

3 EMPIRICKÁ ČÁST

U experimentu s ASR Zoomu jsme se rozhodli zkoumat celkovou přesnost ASR za dobrých podmínek (tzn. bez umělého šumu, s dobrým připojením, kvalitním mikrofonem apod.), ale zároveň se zaměřit na určité zdroje obsahových chyb (jakými jsou např. cizí jména a slova).

Toho docílíme vytvořením projevu, který bude obsahovat prvky, ve kterých ASR vyniká, ale také prvky, o kterých víme, že s nimi mívá ASR problém. Přidanou hodnotou experimentu je také prozkoumání toho, zda v může ASR v aktuálním stavu skutečně pomoci tlumočníkům v tom, co zmiňujeme v kapitole 2.6 *V čem mohou automatické titulky tlumočníkům pomoci*.

3.1 Metodologie

Experiment proběhl ve třech jazycích (čeština, angličtina, španělština). Pro každý jazyk byl vytvořen projev, který je napříč všemi jazykovými verzemi srovnatelný. Zároveň tyto projevy obsahují prvky spadající do kategorií, které chceme zkoumat z hlediska pomoci tlumočníkům: jména, zkratky, akronymy, čísla, termíny, výčty a negace. Následně byly projevy předneseny rodilými mluvčími (jeden za každý jazyk) v rámci živé Zoom schůzky, při které byla zapnuta funkce automatických titulků, kamera (kvůli realističnosti, neboť vizuální přenos může mít dopad na rychlost internetu, která může mít dále dopad na přesnost titulků) a natáčení (k vytvoření automatického přepisu a ke případné kontrole, co mluvčí skutečně vyslovil). V závěru byla přesnost titulků vyhodnocena pomocí metriky NER.

3.1.1 Příprava experimentu a účastníci

Zoom na svých stránkách píše, že přesnost automatických titulků se mj. odvíjí od ruchu v pozadí, hlasitosti a jasnosti hlasu mluvčího, použité slovní zásobě a specifickým dialektům. (Zoom, 2024a) Proto jsme se těmito pokyny rozhodli řídit a zajistit dobré vybavení, prostředí bez okolního ruchu a mluvčí, kteří mluví většinovým dialektem daných jazyků. Zároveň se ale jedná o realistické podmínky, kterých by poučený uživatel Zoomu mohl dosáhnout, nebo se k nim alespoň přiblížit.

K omezení okolního ruchu jsme využili kabin, které se nachází na hlavní budově Filozofické fakulty Univerzity Karlovy v místnosti 310. Kabiny jsou moderní, zvukotěsné, navíc mají každá uvnitř vlastní počítač s monitorem a vestavěnou kamerou. Připojení k internetu je kabelové. Ke

snímaní řeči jsme použili mikrofon Jabra Speak 510. ASR podává horší výkony u spontánních projevů plných hezitací, pauz a falešných začátků, proto budou mít naši řečníci k dispozici předpřipravený text, který budou oralizovat.

Protože neznáme specifika trénovacích dat u Zoomu, abychom předešli zaujatostem, považujeme za nejlepší, aby jednotliví řečníci byli mezi sebou srovnatelní. Jedná se tedy ve všech případech o muže středního věku, přičemž každý z nich je rodilým mluvčím jazyka, ve kterém projev přednáší. Pro angličtinu se jedná o amerického mluvčího, pro španělštinu o mluvčího kastilštiny a u češtiny o mluvčího z Čech. Všichni obdrželi informovaný souhlas o účasti v experimentu (viz přílohy č. 1, 2 a 3).

3.1.2 Tvorba projevu

Cílem projevu bylo, aby zněl přirozeně a realisticky. Zároveň ale všechny tři projevy musely být jazykově srovnatelné a obsahovat stejný počet zkoumaných prvků, aby se daly porovnávat mezi sebou. Nejsnazším postupem proto bylo připravit výchozí projev v češtině a ten poté přeložit do ostatních zkoumaných jazyků tak, aby byly jednotlivé zkoumané kategorie zastoupeny stejnou měrou a okolní text byl zachován s tím, že bude přeložen co nejpřirozeněji. Všechny jazykové verze naleznete v přílohách č. 4, 5 a 6.

Výchozí projev byl napsán v češtině na základě dvou videí z YouTube: přednášky pro mediky (Kieslich, 2023) a přednášky pro veřejnost (Rektorová, 2020). U obou případů se jedná o opravdové projevy a sloužily především jako východisko pro terminologii a zachycení mluvené řeči v jinak odborném a semiodborném prostředí. Vzniklý projev je koncipován jako přednáška o medicínském tématu, konkrétně centrální nervové soustavě, mozku a jeho nemocech. Hypotetický řečník se snaží publiku vysvětlit, jak mozek funguje, předat jim základní znalosti jako např. latinská označení jednotlivých částí a uvést je do tématu Alzheimerovy nemoci, která mozek postihuje.

U češtiny jsme se rozhodli odstranit nespisovné české koncovky, ačkoli jsou v mluvě velmi časté. Důvodem k tomu bylo, že se projev snaží simulovat reálného mluvčího na tlumočené akci v odborném prostředí, kde se očekává spisovnější mluva. Dalším důvodem bylo, že angličtina ani španělština nenabízí obdobný jazykový ekvivalent pro nespisovnost koncovek jako např. *v šedý hmotě* x *v šedé hmotě*.

Kompromisem tedy bylo, aby používaná čeština byla spisovná, ale stále působila mluveně, nikoli psaně. Časté jsou tedy anafory, vedlejší věty, vysvětlování, celkově kratší věty (nebo naopak delší souvětí, které je rozvinuto pomocí vedlejších vět přívlastkových). Také v projevu najdeme řečnické otázky.

> ***Minule*** *jsme mluvili o gastrointestinálním systému, tam máme trubici,* ***která*** *rozmělňuje jídlo a přetavuje ho na molekuly,* ***které*** *vstřebává. (...)* ***Že jsme vůbec pochopili,*** *jak mozek funguje, nám umožnila taková nenápadná nudná obyčejná věc (...). Když přeskočím tak 200 let, tak jsme se dostali až k optogenetice,* ***slyšeli jsme o tom někdy?***

Poté jsme do projevu přidali věty a slova, které se zaměřují hlavně na zkoumané kategorie. Při přípravě projevu jsme zvážili několik kategorií, na které se kromě celkové přesnosti soustředit. Tyto kategorie vychází z provedené rešerše založené na tom, co činí problémy ASR a co činí problémy tlumočníkům. Jsou jimi:

- Jména a názvy (domácí a cizí)
- Zkratky (domácí a cizí) a akronymy
- Čísla
- Termíny (domácí a cizí)
- Výčty
- Negace

Konkrétní zástupce každé z kategorií lze nalézt v přílohách č. 10, 11, 12 a 13.

Z faktického hlediska jsou tyto přidané informace někdy skutečné, jindy ne. Celkově, faktická stránka projevu bylo jedno z kritérií, kde jsme se rozhodli udělat kompromis v prospěch zkoumaných kategorií. Všechny přidané věty jsou variacemi na opravdové věty – nejedná se pouze o náhodný řetězec slov, který by neměl co do činění s tématem, nebo se nemohl v reálném projevu vyskytnout. Je zde potřeba zdůraznit, že cíl experimentu je zkoumat rozpoznávání řeči, nikoli porozumění textu. Dokud jsou tedy věty realistické a nejedná se o úplně náhodné řetězce slov, nemělo by to dle dohledatelných informací představovat problém. Aby projev netrval více než 10-15 minut na přečtení, bylo navíc potřeba dané informace zkondenzovat. Některé informace se proto v projevu objevují „zničehonic“ bez předchozího zmínění:

O kom se mluví méně, ale je neméně důležitý, je Oscar Fischer. Totiž, ona by se ta nemoc měla jmenovat Alzheimer-Fischerova nemoc, protože on už tou dobou, v roce 1897, popsal kolem 15 pacientů s touto nemocí. ***Bohužel, Marie Stejskalová, Bohumila Lískalová ani Karel Bouček nebyli jedni z nich, protože nebyli z Prahy ani z Brna, pocházeli z malé vesnice Plešovice, u které se poblíž nachází Dalešická přehrada.***

Vytvořený projev v češtině jsme poté přeložili do angličtiny a španělštiny. Poté jsme oba texty předali zkušeným proofreaderům (rodilým mluvčím), kteří text dále upravili (v anglické verzi se dokonce jednalo o mluvčího, který text přednášel). Takto upravený text byl ještě dále s předstihem zaslán jednotlivým mluvčím s tím, že mimo vyznačené kategorie mohou text upravit sobě na míru, jak by ho sami chtěli číst.

Z hlediska překladu bylo důležité, aby text zněl v jazyce autenticky, což vedlo k odlišné segmentaci vět. Jak anglický, tak španělský překlad jsou navíc o cca 200 slov delší. Závěrečný transkript, se kterým srovnáváme text automatických titulků, byl nakonec ještě manuálně upraven podle toho, co mluvčí reálně vyslovili během experimentu (vč. zakoktání se, oprav, opakování apod.).

Nyní se budeme věnovat rozboru jednotlivých zkoumaných kategoriích.

3.1.2.1 Jména a názvy

Do této kategorie se řadí jména osob a názvy míst, projektů či aplikací. Obojí je roztříděné do subkategorie domácí (tedy vlastní danému jazyku s výslovností vlastní danému jazyku, proto například *Madrid* figuruje jako domácí název ve všech třech jazycích, protože každý jazyk má zaužívanou vlastní výslovnost) a cizí (tedy cizí danému jazyku s teoreticky cizí výslovností, např. město *Östersund*). Protože jsou našimi mluvčími rodilí mluvčí daných jazyků, chceme prozkoumat, jakou přesnost ASR bude mít u domácích jmen či názvů oproti cizím jménům či názvům. Mluvčí byli instruováni, aby cizí jména a názvy vyslovovali tak, jak si myslí, že by je rodilý mluvčí vyslovil, bez ohledu na jejich vlastní specifické znalosti. Aby nedošlo k tomu, že mluvčí bude znát všechny jazyky, ze kterých jména či názvy pochází, bylo vybráno hned několik různých jazyků (mj. čínština, dánština, němčina, svahilština, švédština), kterými mluvčí nemluví.

Jednotlivá domácí i cizí jména či názvy byly v různých jazykových verzích substituovány tak, aby byla udržena domáckost či naopak cizost. Zachovali jsme také jejich počet, aby bylo možné jazyky srovnávat.

3.1.2.2 Zkratky a vyhláskování

Zkratky jsou rozděleny do podkategorií domácí, cizí a akronymy. Zároveň je k domácím zkratkám přiřazeno také vyhláskovávání slov (*Kiefer Schneider, to je německy a píše se to k-e-i-f-e-r…).* Je to proto, že ve všech jazycích jsou domácí zkratky vyhláskovávány jednotlivými písmeny (*NDR = /en dé er/*).

Co se týče cizích zkratek, v české verzi používáme zkratky z angličtiny (*FBI = /ef bí áj/*): je to proto, že jsou zaužívané s anglickou výslovností. Český mluvčí byl poučen, aby domácí zkratky vyhláskoval česky a cizí (anglické) vyhláskoval anglicky.

V angličtině a španělštině byla kategorie cizích zkratek vynechána, protože tyto jazyky ve většině případů nepřebírají zároveň i cizí výslovnost. Cizí zkratky by tedy mluvčí stejně vyslovovali vyhláskováním jednotlivých písmen v daném jazyce, a proto by se z hlediska výslovnosti v podstatě nelišily od domácích zkratek. Navíc pro jednotlivé mluvčí působily cizí zkratky rušivě a komentovaly, že by to takto rodilý mluvčí nikdy neřekl (např. *VŠE* by Angličan ani Španěl nezkoušel vyslovit).

Akronymy byly vybrány tak, aby byly společné všem jazykům (*Interpol, NASA, UNICEF, UNESCO).* Jediný akronym, který se lišil, byl *AIDS*, který španělština převádí jako *SIDA*.

3.1.2.3 Čísla

Rozhodli jsme se čísla kategorizovat podle jejich řádu. Rozdělujeme tedy čísla na jednotky, desítky, stovky, tisíce, miliony a miliardy. Každá kategorie v každém z jazyků obsahuje pět čísel. Kromě toho jsme zavedli také dvě další kategorie:

1. Roky, protože mají v jednotlivých jazycích různou výslovnost a v zájmu udržení poměru jednotlivých kategorií je jednodušší s nimi pracovat jako se separátní kategorii. V češtině je rok 1897 vysloven jako /*osmnáct set devadesát sedm*/ místo /*tisíc osm set devadesát sedm*/; v angličtině je cokoli před 2000 a po 2009 rozloženo do dvou desítkových čísel,

1897 je tedy /*eighteen ninety-seven*/; španělština jako jediná zachovává stejnou výslovnost jako u tisíců, 1897 je tedy /*mil ochocientos noventa y siete/.*

2. Speciální, do této kategorie spadají zlomky (2/3, ½), mocniny (10^{11}, 10^{-9}) a telefonní čísla, která se v každém z jazyků vyslovují a zapisují jinak. Zároveň, ačkoli mluvčí byli instruováni, aby telefonní čísla vyslovovali po jednotkách, každý z nich byl navyklý na jiný způsob a vyslovil číslo jako kombinaci jednotek a desítek, což znemožnilo zařazení do jiných kategorií.

Také jsme zvažovali, zda jako samostatnou kategorii zařadíme falešné začátky, tedy případy, kdy mluvčí začne vyslovovat číslo, zasekne se a vrátí se na začátek, případně se zmýlí, vrátí se na začátek a při opětovném vyslovení chybu opraví. Nakonec jsme se ale rozhodli tuto kategorii kompletně vynechat, protože ačkoli ji Fantinuoli et al. (2020) uvádí jako zdroj častých chyb u ASR, působila by v projevu už příliš uměle.

3.1.2.4 Termíny

Kvůli medicínské tématice se v projevu objevuje mnoho termínů. K tomu, co je termín a co tedy za termín považujeme i my, uvádí Nový encyklopedický slovník češtiny následující:

> *K vlastnostem [termínů] (a zároveň k požadavkům na ně kladeným) patří: (1)* ***ustálenost*** *(vlastnost zabezpečující bezporuchovost odborné komunikace), (2)* ***systémovost*** *(míra, do jaké daný t. zapadá do systému terminologie daného oboru a do jaké odráží systémové sepětí termínů daného oboru), (3)* ***přesnost*** *(t. má pojmový význam, který je definován) a jednoznačnost (t. má být jednoznačný, alespoň v rámci terminologie příslušného oboru, popř. oborů souvisejících), (4)* ***nosnost*** *(schopnost t. být východiskem při tvoření dalších t.), (5)* ***ústrojnost*** *(struktura t. má být v souladu se zákonitostmi a pravidly jaz. systému).* (Martincová & Bozděchová, 2017)

Napříč všemi jazyky jsme provedli důkladnou rešerši, aby si v jednotlivých jazycích termíny skutečně odpovídaly.

Jako cizí termíny jsou uváděna jména částí centrální nervové soustavy v latině a řečtině. Uvádíme je jako cizí, protože se jedná přímo o latinské či řecké názvy, nikoli, že z latiny či řečtiny pochází. Nejlépe je tento rozdíl vidět na španělštině, která se z latiny vyvinula:

prodloužená mícha je *médula espinal*, latinsky *medulla spinalis*. Na španělské verzi jsou vidět jednotlivé etymologické změny, kterými termín prošel, a proto je *médula espinal* vedena jako termín domácí, zatímco *medulla spinalis* jako termín cizí.

3.1.2.5 Výčty

U výčtů jsme podle vzoru Shlesinger (2003), viz 2.6.5. *Výčty*, volili tři a více položek (ve většině případů to byly výčty o čtyřech členech).

3.1.2.6 Negace

Negace se v textu vyskytovaly přirozeně. Bylo ovšem potřeba zvážit hledisko srovnatelnosti. Protože jeden z našich hlavních cílů bylo zachovat autenticitu a přirozenost v každém z jazyků, ačkoli se ve dvou z nich jednalo o překlad, netrvali jsme na stoprocentním zachování negace a dali jsme proofreaderům volnost v jejich případné opravě.

Např. věta *Za dob SSSR jsme o tom tolik nevěděli (...)* byla ve španělštině převedena jako *Durante la existencia de la URSS, sabíamos poco de él (...)* (doslova *věděli jsme málo*). Obdobně, věta *není to žádná lehká hydrolýza esterů* byla v angličtině přeložena jako *it makes ester hydrolysis look pretty easy in comparison* (doslova *hydrolýza esterů vypadá v porovnání poměrně snadně*).

Následně jsme tedy hodnotili a srovnávali pouze ty negace, které byly zachované ve všech jazykových verzích.

3.1.3 Vyhodnocení

Po ukončení schůzky Zoom vygeneruje několik souborů: mj. textový soubor s časovou stopou a textem titulků, audiosoubor s výchozím projevem a videosoubor, kde lze mluvčího vidět, slyšet a případně i zapnout titulky.

Všechny tyto soubory jsme si stáhli a obsah textového souboru zkopírovali do separátního dokumentu. Odstranili jsme časovou stopu a projev rozčlenili tak, aby odpovídal výchozímu projevu. Každá časová stopa má totiž vlastní řádek, což mnohdy neodpovídá našemu „standardnímu“ členění do odstavců. Upravovali jsme ale pouze vizuální stránku, obsahově zůstalo všechno stejné. Text automatických titulků ve všech jazycích naleznete v přílohách č. 7, 8 a 9.

V této podobě poté automatické titulky srovnáváme oproti výchozímu projevu. Jeho obsah jsme také upravili na základě toho, co mluvčí skutečně přečetli a řekli, a to včetně oprav a zakoktání se (viz přílohy č. 4, 5 a 6).

K vyhodnocení používáme metriku NER. Zvolili jsme tuto metriku kvůli jejím výhodám oproti WER a sofistikovanějším metrikám, které zmiňujeme v kapitole 2.4.3. *Metriky pro výpočet přesnosti automatických titulků.*

Tato metrika byla původně navržena pro ohodnocení výkonu respeakerů, ale do budoucna (Romero-Fresco & Martínez ji navrhli v roce 2015) autoři počítali s jejím využitím i u automatických titulků.

Rozděluje tři typy chyb:

- *Edition errors* (nesprávná úprava, vl. překl.)
- *Recognition errors* (nesprávné rozpoznání, vl. překl.)
- *Correct editions* (správná úprava, vl. překl.).

Správné úpravy byly původně zavedeny pouze pro případy, kdy respeakeři správně upravili výchozí projev – např. kondenzace informace, vypuštění fatických výrazů apod. U automatických titulků se s tímto typem úprav nepočítá, a proto ho nerozlišujeme ani my.

Dále metrika rozlišuje chyby podle stupně závažnosti a přisuzuje jim různou váhu:

1. *Minor error* (malá chyba, vl. překl.; váha 0,25)
2. *Standard error* (standardní chyba, vl. překl.; váha 0,5)
3. *Serious error* (vážná chyba, vl. překl.; váha 1)

Vzoreček k výpočtu je $NER = \frac{N-E-R}{N} \text{x } 100$. N je celkový počet slov, E je celková váha chyb v nesprávné úpravě, R je celková váha chyb v nesprávném rozpoznání.

V naší práci provádíme dvojí vyhodnocení: celkovou přesnost z hlediska čtenáře a přesnost v rámci jednotlivých zkoumaných kategoriích z hlediska tlumočníka. K tomuto dvojímu hodnocení jsme se uchýlili proto, že z rešerše (zejm. Yuan & Wang, 2023) vyplynulo, že tlumočníci nečtou titulky jako celek, nýbrž vyhledávají konkrétní informace.

Hledisko čtenáře pojímáme z pozice člověka, který neslyší originál a automatické titulky jsou jeho jediným zdrojem informací. Nemusí u toho tlumočit, ani se soustředit na cokoli jiného, co by ovlivňovalo jeho kognitivní kapacitu. Důležitým aspektem je kromě přesnosti také smysl.

Oproti tomu hledisko tlumočníka pojímáme z pozice člověka, který slyší výchozí projev, ale zároveň u toho musí tlumočit. Automatické titulky tedy nečte jako celek, ale vyhledává v nich pomoc u spouštěčů problémů, jakými jsou například jména, čísla apod. Desmet et al. (2018) shrnuje, že aby ASR tlumočníkům pomáhalo konkrétně v předání čísel, je potřeba, aby: 1. bylo rychlé a přesné, zvládalo různé mluvčí a přízvuky, 2. dokázalo ukázat samotná čísla bez okolních slov, 3. tlumočníkovi čísla ergonomicky ukázalo. Domníváme se, že tato kritéria lze kromě druhého uplatnit na ASR celkově. Kromě přesnosti je tedy důležitá snadná vyhledatelnost v textu.

Pro větší názornost vždy uvádíme všechna zvážená kritéria a hlediska u jednotlivých kategorií.

3.2 Výsledky experimentu

Nejdříve se budeme věnovat vyhodnocení celkové přesnosti a poté postupně probereme výsledky v rámci jednotlivých kategorií. Veškerá zvážená kritéria, na které jsme při hodnocení brali zřetel, jsou zmíněna u každé kategorie zvlášť, stejně tak hodnocení závažnosti (malá / standardní / vážná chyba).

3.2.1 Celková přesnost

V této části se věnujeme vyhodnocení celkové přesnosti automatických titulků v jednotlivých jazycích. Na chyby nahlížíme z hlediska čtenáře, tedy někoho, kdo (s největší pravděpodobností) neslyší originál. Hlavním kritériem je pro nás to, aby automatické titulky odpovídaly originálu a aby dávaly smysl. Text automatických titulků v jednotlivých jazycích s vyznačenými chybami je k dispozici v přílohách č. 7, 8 a 9.

Váhu chyby jsme posuzovali následovně:

Malá chyba:

- Nesprávně použitá velká či malá písmena, nesprávná interpunkce (tzn. tam, kde je jasně dané, jaká interpunkce je správná, příp. tam nemá být žádná; např. *probírat dýchací systém. Močové a pohlavní ústrojí. Srdce…).* (Nesprávná úprava)
- Chyba, kterou lze lehce odhalit a opravit (např. gramatické chyby). (Nesprávné rozpoznání)

Standardní chyba:

- Interpunkce, která mění význam či mate čtenáře (např. *přetavuje ho na molekuly, které vstřebává mozek je druhý (…)* – zde měla být tečka mezi *vstřebává* a *mozek)*. (Nesprávná úprava)
- Chyba, kterou lze odhalit, ale je složitá na zpětnou rekonstrukci (např. *těla neuronů -> děla neuronů*). (Nesprávné rozpoznání)
- Vynechaná slova, u kterých lze z kontextu dovodit, že došlo k nějaké vynechávce (*on Douyin -> on* __). (Nesprávné rozpoznání)

Vážná chyba

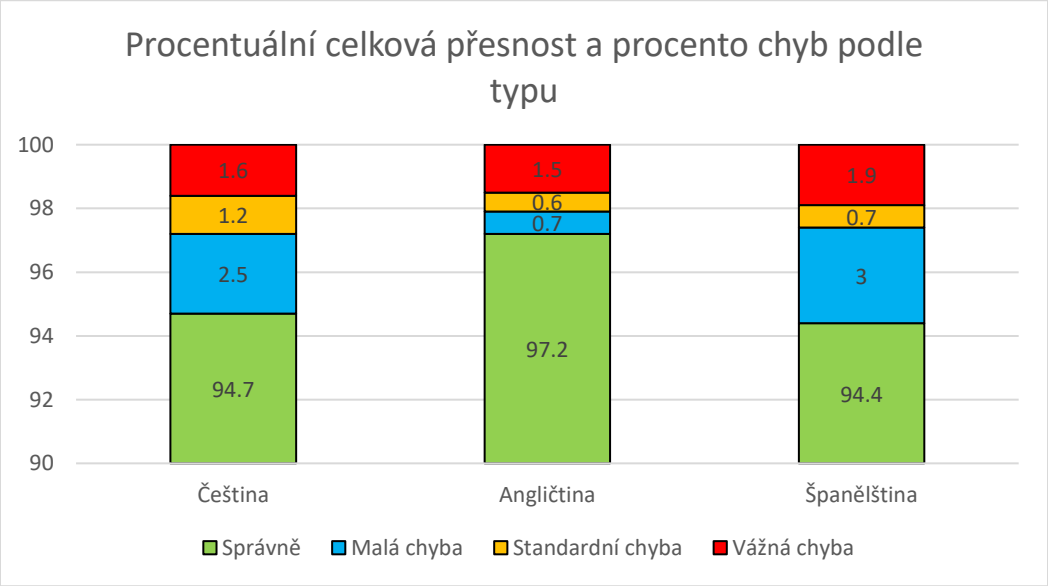
- Chyba, kterou nelze odhalit (např. nesprávná jména či čísla). (Nesprávné rozpoznání)
- Vynechaná slova, u kterých nelze z kontextu dovodit, že byla vynechána (*máme pons, mesencephalon, diencephalon a telencephalon. Chybí pak… -> máme Ponce. Chybí pak…).* (Nesprávné rozpoznání)

Jako chyby jsme nepočítali:

- Případy, kdy šlo na daném místě udělat tečku či čárku bez změny významu (např. *Dobré ráno*(./,) *Já jsem…)*.
- Nestandardní zápis čísel (např. v češtině slovy místo číslic).
- V angličtině jsme nepočítali za chyby zkrácení slov, kde je to gramaticky možné (např. *we had -> we'd*).

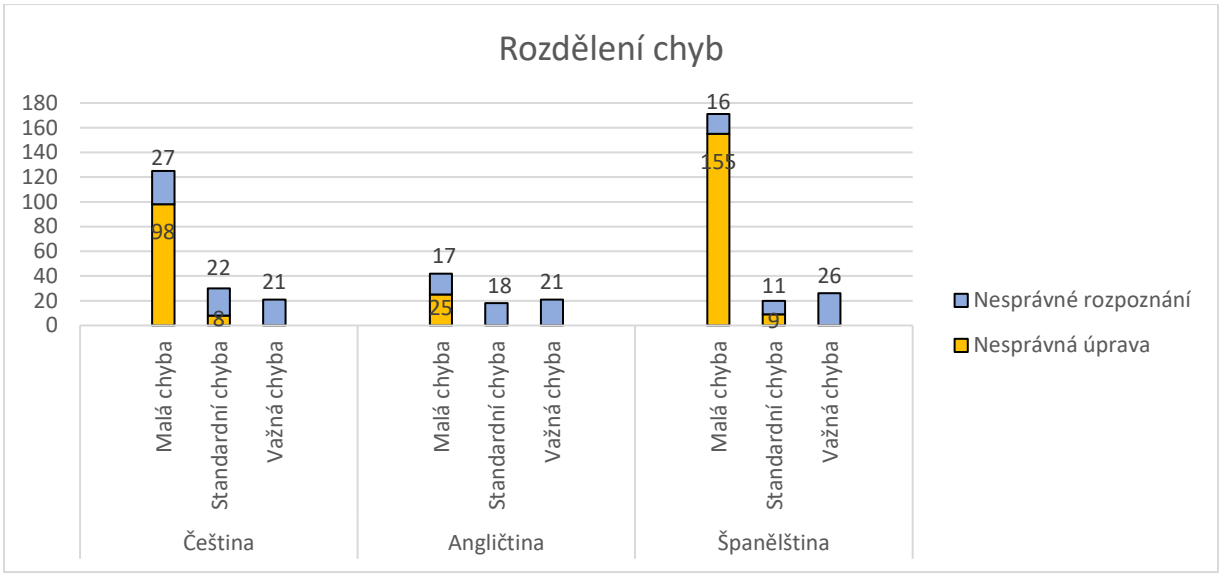
Pokud se nakupilo několik chyb v jednom slově, počítali jsme pouze tu nejzávažnější jako jednu chybu. Celkově obsahovaly české titulky 1276 slov, anglické 1436 slov a španělské 1400 slov.

Po vyhodnocení všech chyb jsme došli k následující celkové přesnosti v procentech:



Graf 1: Celková procentuální přesnost ve všech jazycích, vč. procenta chyb dle závažnosti.

V grafu vidíme, že angličtina dosahuje nejvyšší přesnosti (97,2 %), což je v souladu s našimi očekáváními (angličtina je tzv. jazyk bohatý na zdroje, viz 2.4 *ASR*). Čeština i španělština dosahují velmi podobného výsledku, 94,7 % a 94,4 %. Navíc, poměrně vysoké procento chyb (2,5 % a 3 %) jsou pouze malé chyby. Detailnější rozdělení typu chyb lze vidět v tomto grafu:



Graf 2: Rozdělení chyb podle závažnosti a typu ve všech jazycích.

Napříč všemi jazyky vidíme, že malá chyba v nesprávné úpravě (tzn. hlavně interpunkce a velká písmena) je nejčastějším typem chyby. Týká se to hlavně španělských titulků, které kompletně postrádají tečky, čárky a velká písmena, jinak dosahuje poměrně vysoké přesnosti. České titulky mají obdobný problém, ačkoli v menším měřítku.

Chyby u nesprávného rozpoznání jsou převážně gramatického rázu: v českých titulcích např. *mě/mně, bude/budu*; v angličtině např. *here you guys ever heard about it?*; ve španělštině např. *enseño -> enseñó*.

Nesprávná úprava u standardních chyb se týká pouze českých a španělských titulků, protože se jednalo o chybnou interpunkci (často chybějící, ale ne vždy), která mate čtenáře (*někde mezi 1545 a 20 000 spoji -> někde mezi tisíci pěti sty, čtyřiceti pěti a dvaceti tisíci spoji* – zde se vyskytla čárka v čísle, která působí, jako kdyby se jednalo o tři čísla místo dvou).

Nesprávné rozpoznání bylo vyhodnoceno jako standardní chyba v případech, kdy ASR špatně zapsalo slovo (*šedé hmotě -> že hmotě*).

Nesprávná úprava nebyla nikdy vyhodnocena jako vážná chyba.

K vážné chybě v nesprávném rozpoznání docházelo hlavně v případech cizích slov: cizí termíny a cizí jména či názvy. U španělštiny se také ale jednalo o chyby v zápisu milionů. Je nutno také podotknout, že ačkoli má angličtina v ostatních typech chyb výrazně lepší výsledky, v kategorii vážných chyb je srovnatelná s češtinou i španělštinou.

3.2.2 Výsledky v rámci jednotlivých kategoriích

V této části se věnujeme vyhodnocení jednotlivých kategoriích s použitím NER metriky, tentokrát ale bereme v potaz hledisko tlumočníka: tedy někoho, kdo slyší originál a automatické titulky nevnímá jako souvislý text, spíše v nich vyhledává informace potřebné k přetlumočení informace. Hlavní kritérii je tedy přesnost a snadná vyhledatelnost. Zároveň nerozlišujeme mezi nesprávnou úpravou a nesprávným rozpoznáním. Nesprávná úprava vede ke špatné vyhledatelnosti a nesprávné rozpoznání ke špatné přesnosti, proto je obojí stejně závažné a nevidíme tedy jako důležité mezi nimi rozlišovat. Zástupce kategorií spolu s tím, jak byly zapsány automatickými titulky a jak byly hodnoceny, je dostupné v přílohách č. 10, 11, 12 a 13.

3.2.2.1 Jména

U této kategorie jsme jako hlavní kritéria, podle kterých jsme určovali závažnost chyby, viděli tato:

- velké písmeno na začátku, aby bylo jméno snadno vizuálně vyhledatelné,
- pokud tlumočník přečte, co automatické titulky nabídly, bude to do přijatelné míry odpovídat původní výslovnosti (toto posuzování je vysoce subjektivní a ideálně by se na něm mělo shodnout několik tlumočníků, nicméně zapojení dalších tlumočníků do experimentu přesahovalo rozsah této práce),
- příjmení je důležitější než jméno,
- pokud se nakupí několik chyb (např. špatně přepsané příjmení a malé písmeno), počítáme pouze tu nejzávažnější.

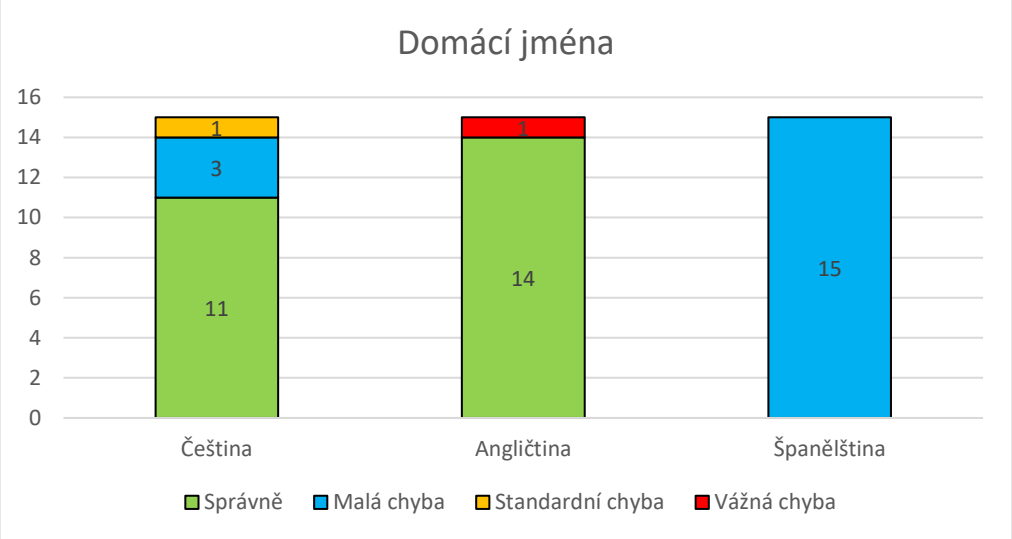
Chyby jsme hodnotili na této stupnici:

- Malá chyba: původní výslovnost se velmi blížila a tlumočníka by nezmátla, popř. chybělo velké písmeno.

- Standardní chyba: přepis by tlumočníka mohl rozhodit, ale pokud by jej foneticky přečetl (s tím, že jméno předtím slyšel), reprodukoval by jméno přijatelně úspěšně.
- Vážná chyba: přepis by tlumočníkovi vůbec nepomohl, nebo by ho naopak mohl mystifikovat.

3.2.2.1.1 Domácí jména

Celkově bylo v každém z jazyků 15 domácích jmen a názvů.



Graf 3: Počet správně a špatně zapsaných domácích jmen u všech jazyků, vč. závažnosti chyb.

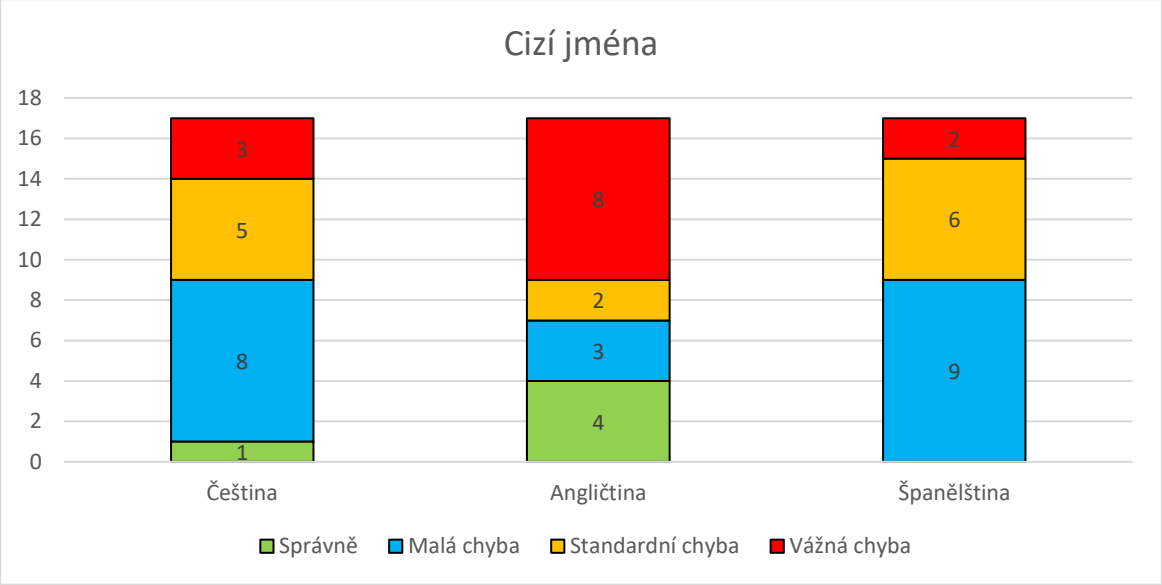
U českých titulků byla přesnost při převodu domácích jmen 91,7 %. Došlo k vynechání první slabiky příjmení (*Lískalová -> skalová*) a třem malým písmenům na začátku jmen.

U anglických titulků byla přesnost při převodu domácích jmen 93,3 %. Došlo k záměně celého jména *Lavenham* na *Lebanon*, jinak byla jména přepsána správně i s velkými písmeny.

U španělských titulků byla přesnost při převodu domácích jmen 75 %. Důvodem, proč španělské titulky obdržely tak nízké procento, jsou všechna malá písmena na začátku jmen. Tím splývají s okolním textem, což může ztěžovat jejich vyhledání během tlumočení. Jinak španělština měla celkově nejlepší výsledky, jediné menší posuny byly *Muñez -> muñiz, Lopez -> lópez* a *Ruíz -> ruiz*. Pokud bychom pouze u španělštiny pominuli hledisko velkých písmen a započítali pouze tři malé chyby zmíněné výše, dosahovala by úspěšnosti 95 %. Její přesnost je tedy poměrně vysoká, ale bohužel jména nelze snadno vyhledat.

3.2.2.1.2 Cizí jména

V kategorii cizích jmen bylo velmi důležité kritérium, zda by tlumočník dokázal jméno na základě přepisu reprodukovat přijatelně správně, za předpokladu, že jméno slyšel. Celkem bylo v každém z jazyků 17 cizích jmen a názvů.



Graf 4: Počet správně a špatně zapsaných cizích jmen u všech jazyků, vč. závažnosti chyb

U češtiny byla přesnost při převodu cizích jmen 55,9 %. Vážnými chybami byly *HjerneProjekt -> Jeden projekt, Östersund -> Ssteresson* a *Dieter -> Petr*. Jako příklad standardní chyby lze uvést např. *Juan Luis Jorge -> Juan Weiss Horgez,* či *Ujana Salama -> u Jana saláma*.

U angličtiny byla přesnost při převodu cizích jmen 42,6 %. Angličtina měla největší počet vážných chyb, jako příklady lze uvést *Hjärnfonden -> Hyung Fonda,* či *Octavia Velazquéz -> Octavian Los Angeles*, příp. kompletní vynechání názvu aplikace *Douyin*. K vynechání jména došlo také u *Atiquzzama Khan -> - Khan*, zde jsme to ale hodnotili jako malou chybu, neboť bylo zachováno příjmení a bylo přepsáno správně.

U španělštiny byla přesnost při převodu cizích jmen 57,4 %. Podobně jako v kategorii domácích jmen měla poměrně vysokou úspešnost, až na dvě vážné chyby *Dieter -> dicha* a *Brno -> horno*, ale opět kompletně vynechávala velká písmena. Pokud bychom opět pouze u španělštiny nepočítali nepřítomnost velkých písmen jako chybu, vystoupala by její úspešnost na 61,8 %.

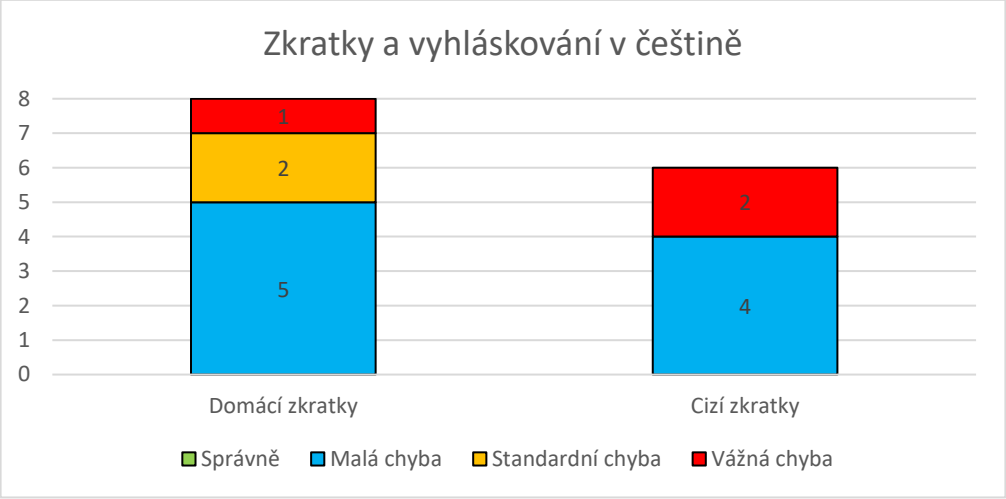
3.2.2.2 Zkratky a vyhláskování

Hlavní kritéria u zkratek a vyhláskování byla tato:

- velká písmena v celé zkratce (pokud tam mají být),
- správná písmena,
- pokud se nakupí několik chyb, počítáme pouze tu nejzávažnější.

3.2.2.2.1 Čeština

V rámci zkratek jsme pro češtinu vyčlenili vlastní kategorii, protože na rozdíl od angličtiny a španělštiny existují v češtině zkratky, které lze přečíst v jazyce, ze kterého přišly (konkrétně z angličtiny, např. *FBI* či *BBC*). Domácích zkratek bylo šest, navíc se dvěma vyhláskováními (tzn. osm celkem). Cizích zkratek bylo šest.



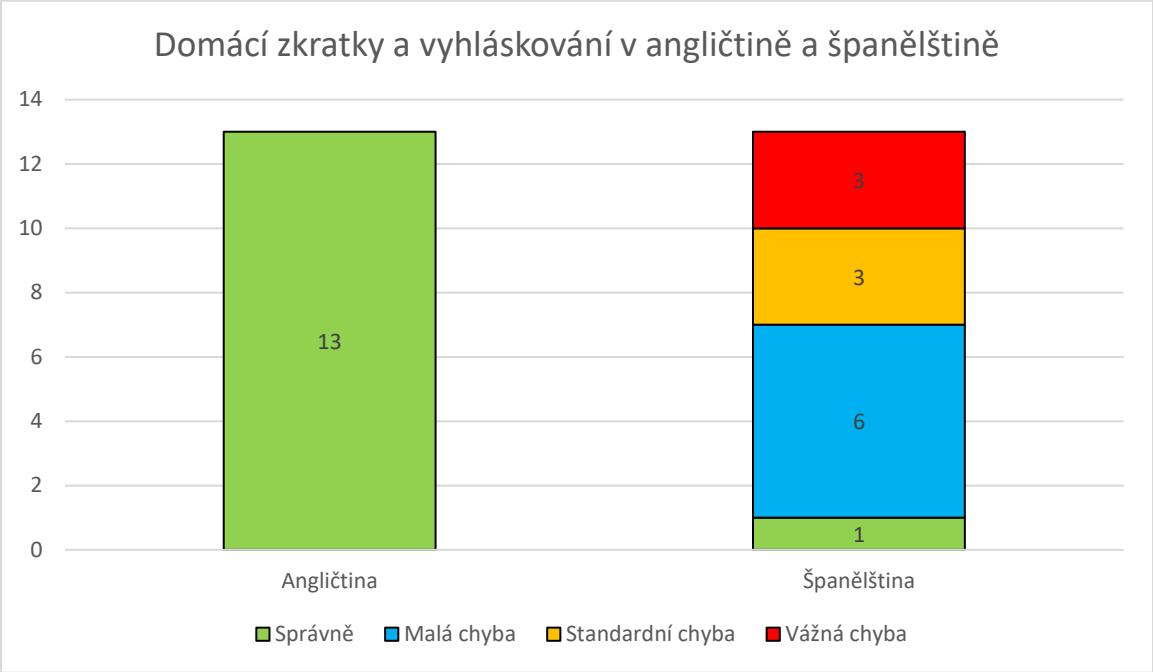
Graf 5: Počet správně a špatně zapsaných domácích a cizích zkratek a vyhláskování u českých titulků, vč. závažnosti chyb.

Přesnost při převodu českých domácích zkratek byla 59,4 %. Ani jedna zkratka či vyhláskování nebylo kompletně správně. Malými chybami byla velká písmena pouze na začátku, nikoli v celé zkratce (např. *EEG -> Eeg*). Standardní chyby v domácích zkratkách byla pomýlená písmena, konkrétně *SSSR -> ssr* a *SZ -> Cz*. Vážnou chybou bylo vyhláskování jména *k-i-e-f-e-r s-ch-n-e-i-d-e-r -> kádí Efers. Chm. Mhm.*

Přesnost při převodu cizích zkratek v českých titulcích byla 50 %. Všechny malé chyby spočívaly v tom, že velké písmeno bylo pouze na začátku zkratky. Vážnými chybami bylo *UCL -> Usil* a *AI -> ej.*

3.2.2.2.2 Angličtina a španělština

Celkově bylo v této kategorii 13 domácích zkratek v každém z jazyků.



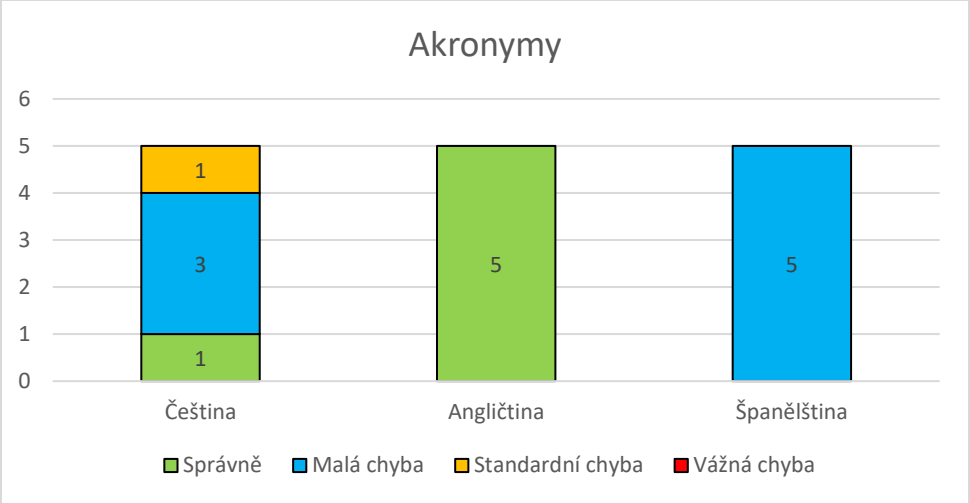
Graf 6: Počet správně a špatně zapsaných domácích zkratek u anglických a španělských titulků, vč. závažnosti chyb.

V anglických titulcích je přesnost při převodu zkratek 100 %. Všechny zkratky i vyhláskování byly správně, a to včetně všech velkých písmen. U vyhláskování dokonce skupiny písmen odpovídaly tomu, jak je po sobě vyslovoval mluvčí (např. *t-r-a-c-t-u-s c-o-r-t-i-c-o-s-p-i-n-a-l-i-s -> TRA CT US. CORT ICOS PINA LIS.*).

Ve španělských titulcích je přesnost při převodu zkratek 54 %. Většina zkratek postrádala jakékoli velké písmeno (kromě *FBI -> FBI* a *PhD -> pHD*). Standardní chybou bylo vynechání písmene (např. *NMDA -> nma*). Jako vážné chyby jsme viděli např. *EEG -> eje* či *IA -> ya.*

3.2.2.3 Akronymy

Tato kategorie byla napříč všemi jazyky stejná, proto ji lze opět srovnávat dohromady. Každý z jazyků obsahoval pět akronymů.



Graf 7: Počet správně a špatně zapsaných akronymů u všech jazyků, vč. závažnosti chyb.

U českých titulků byla přesnost při převodu akronymů 75 %. U akronymů, které měly všechna velká písmena (např. *NASA*) bylo zachováno pouze první velké (*Nasa*). Jako standardní chybu jsme hodnotili záměnu písmena *UNICEF -> Unicet*.

U anglických titulků byla přesnost při převodu akronymů 100 % včetně velkých písmen.

U španělských titulků byla přesnost při převodu akronymů 75 %, ale pouze proto, že nebyla použita žádná velká písmena, jinak by španělština také dosáhla 100 %.

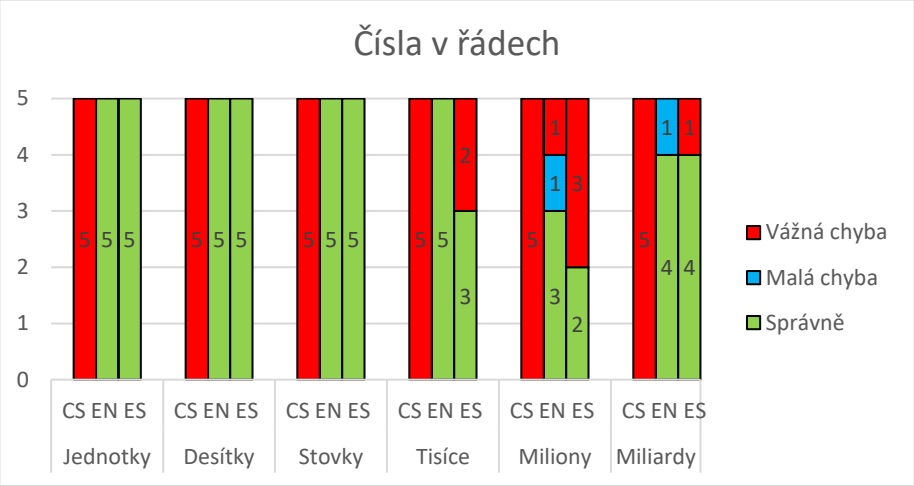
3.2.2.4 Čísla

U čísel došlo k velmi zajímavému zjištění, a sice že zatímco u anglických a španělských titulků jsou čísla převáděna na číslice, u češtiny jsou přepisována slovně. Z tlumočnického hlediska se tedy dá říci, že je slovní přepis čísel v českých titulcích téměř nepoužitelný, protože vyhledání čísla mezi ostatními slovy, přečtení čísla, jeho zpracování a následné přetlumočení může značně převýšit dostupnou kognitivní kapacitu tlumočníka. Ačkoli jsou tedy čísla v českých titulcích přepsána většinově správně, tlumočník nemůže automatické titulky využít. Vzhledem k tomu, že dané kategorie hodnotíme právě z tlumočnického hlediska, rozhodli jsme se proto veškerá čísla v českých titulcích označit za vážnou chybu. Jinak jsme za vážnou chybu v anglických a španělských titulcích považovali čísla, která se neshodovala s těmi ve výchozím projevu.

Standardní chybu jsme v této kategorii nerozlišovali, protože u čísel není možná: pokud je číslo zapsáno špatně, nelze se domýšlet, jaké číslo tam mělo být správně.

Za malou chybu jsme označovali nestandardní zápis: např. mezera mezi číslicemi, kde být neměla *1,75 -> 1,7 5*). Domníváme se totiž, že v jazycích, kde jsou čísla celkově zapisovaná číslicemi, může nestandardní zápis vést k lehce zvýšenému úsilí ze strany tlumočníka.

Pokud se nakupí několik chyb, počítáme pouze tu nejzávažnější. Celkem bylo v každém z jazyků pět stejných čísel v každé podkategorii.



Graf 8: Počet správně a špatně zapsaných čísel u všech jazyků, vč. závažnosti chyb.

Jak vidíme, u jednotek, desítek a stovek je anglické a španělské ASR velmi přesné.

Problém tisíců se u španělských titulků zdá být v tom, že označení pro tisíce *mil* je velmi krátké a nepříliš jasně artikulované. Tak tomu bylo v případě *20 000 -> 20*. V druhém případě byl vynechán celý konec čísla: *54 520 -> 54,5*.

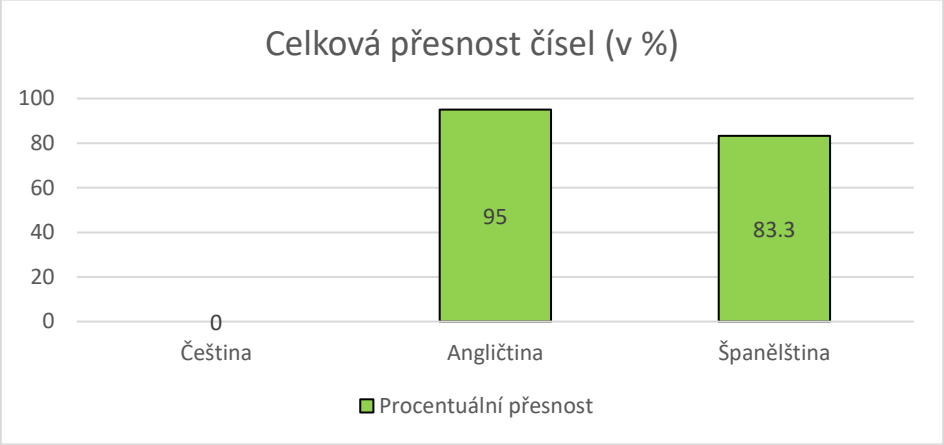
U milionů se ve španělských titulcích ukázalo, že španělské slovo *millones* je automaticky převáděno na šest nul. To se ukázalo být problematické tam, kde bylo číslo zkrácené: *28,5 millones -> 25,8 000 000; 1,75 millones -> 1 coma 75 000 000.* Třetí vážnou chybou byl případ, kdy po sobě následovala dvě různá čísla: *113 millones, 2/3 de… -> 113 000 002 tercios.*

Stejná chyba u milionů se vyskytla také u anglických titulků, kde se mluvčí nejdříve dopustil chyby a následné opravy: *130- and by 2050 about 113 million, 2/3 of (…) -> 130 and by 2,050 about a hundred 13,000,002 thirds of* (…). Vidíme zde, že ASR správně přepsalo i počáteční opravu, ale potom číslo přepsalo špatně. Možným důvodem mohlo být to, že mluvčí po opravě začal mluvit mnohem rychleji a překotněji, což vedlo k nejasné artikulaci.

Co se týče miliard, u španělských titulků došlo ke stejnému problému jako v milionech, protože španělština nepoužívá samostatné slovo pro miliardy, nýbrž mluví o tisících milionů: *9.700 millones -> 9,7 000 000*.

V závěru považujeme za důležité zmínit také několik jednotlivostí pro konkrétní jazyky:

- Ačkoli mají české titulky kompletně slovní přepis, i v tomto se vyskytly určité chyby. Mezi slovy byla dvakrát čárka (*1 545 -> tisíci pěti sty, čtyřiceti pěti; 200 -> dvě, stě*), bylo vynecháno jedno procento (*2-3 % pacientů -> dvou až tří pacientů*) a v miliardách došlo k chybě *8 miliard -> osoba*.
- Anglické titulky variují v zápisu jedničky: někdy ji zapíšou jako *one* (*1-5 years -> one to 5 years*), jindy jako *1* (*there's one other person -> there's 1 other person*).
- Zdá se, že španělské titulky mají celkový problém se znaménky: nezachycují totiž %, a to ani slovně, což je pro tlumočníky velmi důležitá informace (*no el 20% -> no el 20*). Zároveň vyslovenou čárku zapisují slovně (*1,75 millones -> 1 coma 75 000 000*), ale tam, kde není, ji přidávají (*54 520 -> 54,5*).

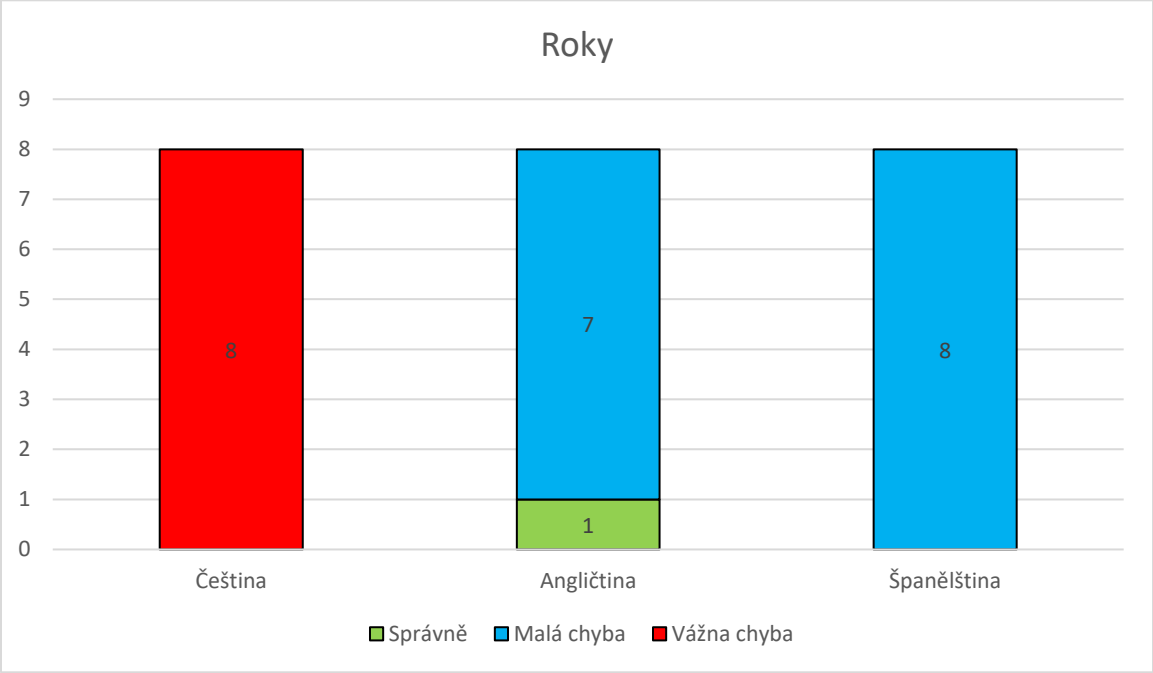


Graf 9: Celková procentuální přesnost zapisování čísel u všech jazyků.

Jak vidíme, anglické titulky dosahují značně vyšší přesnosti u čísel než ty španělské. České titulky obdržely nula procent, protože v zápisu vůbec nepoužívají číslice, nýbrž slova. Pokud bychom se soustředili čistě na správnost a slovní zápis opomenuli, dosahovaly by české titulky 95 %.

3.2.2.4.1 Roky

U zápisu roků jsme jako malou chybu hodnotili nestandardní zápis. Záměrně jsme vybrali několik různých letopočtů, abychom otestovali, jak různá výslovnost zápis ovlivní. Celkem se v každém z jazyků objevilo osm stejných roků. Standardní chybu jsme zde opět nerozlišovali a jako vážnou chybu jsme stejně jako u předchozí kategorie hodnotili slovní přepis.



Graf 10: Počet správně a špatně zapsaných roků u všech jazyků, vč. závažnosti chyb.

České titulky obdržely 0 %, protože mají opět kompletně slovní zápis. Věcně správně byly ale všechny roky (tedy 100% přesnost). Zápis je dvojího typu:

- roky po 2000 jako tisíce: *2030 -> dva tisíce třicet*
- roky před 2000 včetně slova *set*: *1910 -> devatenáct set deset*

Anglické titulky z hlediska nestandardního zápisu dosáhly 78 %, celkově byly jinak věcně správně všechny roky. U angličtiny jsme pozorovali tři různé typy zápisu:

- roky po 2000 jako tisíce s čárkou: *2030 -> 2,030*
- roky před 2000, které se vyslovují po dvou částech, bez mezery: *1897-> 1897*
- roky, které obsahují 0 vyslovovanou /*oh*/, s mezerou: *1907 -> 19 0 7*

Tento vzorec ale porušuje rok *1910-> 1,910*, ačkoli ho mluvčí vyslovil po dvou částech, /*nineteen ten*/, což mělo být zapsané jako *1910*.

Obdobně, španělské titulky měly 75% přesnost z hlediska zápisu, věcně správně byly všechny roky. U španělštiny vidíme ve všech případech zápis stejný jako u tisíců, neboť je španělská výslovnost stálá u všech roků: *2010 -> 2 010, 1897 -> 1 897.*

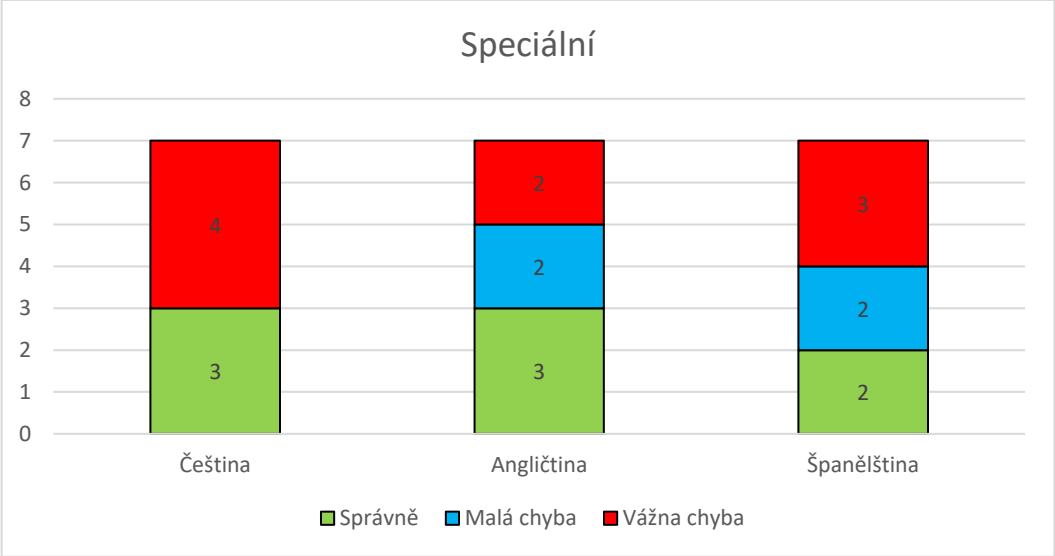
3.2.2.4.2 Speciální

Kategorie speciální obsahuje hned několik různých druhů čísel (zlomky, mocniny a telefonní čísla, která jsou specifická svým odlišným zápisem v každém z jazyků). Nelze zde ale jednotně penalizovat slovní zápis, neboť např. ½ se běžně zapisuje také jako *polovina*. Zlomky (1/10, 2/3 a ½) jsme se proto rozhodli považovat za správné i v případě, že byly zapsané slovně.

Jako malou chybu jsme hodnotili případy, kdy sice číslo nebylo zapsané tak, jak by mělo (10^{11}), ale obsahovalo alespoň nějaké číslice (*10 to the 11th power*).

Jako vážnou chybu jsme hodnotili kompletně slovní přepis mocnin, telefonních čísel a také nesprávně zapsaná čísla.

Celkově se objevily v každém z jazyků tři zlomky, dvě mocniny a dvě telefonní čísla.



Graf 11: Počet správně a špatně zapsaných mocnin, zlomků a telefonních čísel u všech jazyků, vč. závažnosti chyb.

Celkově měly české titulky přesnost 43 %, anglické 64 % a španělské 50 %.

V anglických a španělských titulcích jsme jako malou chybu hodnotili zápis mocniny pomocí kombinace slov a číslic (*10^{11} -> 10 to the 11th power, 10 elevado a 11*).

U českých titulků jsme jako vážnou chybu hodnotili kompletně slovní zápis mocnin a telefonních čísel (*10^{11} -> deset na jedenáctou; + 420 602 875 483 -> plus čtyři, dva, nula, šest, nula, dva, osm, sedm, pět, čtyři, osm, tři*).

V anglických a španělských titulcích pak vážná chyba nastala ve stejných případech: telefonní čísla a číslo následované zlomkem.

- U zlomku se konkrétně jednalo o část „*113 milionů, 2/3 z nich ve vyspělých zemích*“. Tato juxtapozice dvou odlišných čísel vedla ke špatnému přiřazení zlomku k předcházejícímu číslu. Vzniklo tak *hundred 13,000,002 thirds* v angličtině a *113 000 002 tercios* ve španělštině.

Telefonní čísla se také ukázala být problematická, a to hlavně proto, že každý jazyk má odlišné způsoby, jakými je vyslovuje, konkrétně u telefonní předvolby.

- U angličtiny se vyslovuje /country code/ a poté jednotlivé číslice. U českého telefonního čísla byl zápis přesný: *country code 420 602 875 483 -> country code 4 2 0 6 0 2 8 7 5 4 8 3.*

  Ale u španělského čísla došlo k změně na americký zápis: *country code 34 662 00 57 36 -> country code **(346)** 62**0**-5736.* Nejenom, že tedy byla předvolba špatně, ale aby číslo vyhovovalo formátu, byla odstraněna jedna nula uprostřed čísla.
- U španělštiny se na začátku vysloví dvě nuly a poté jednotlivé číslice. Zdá se ale, že nuly představují pro ASR problém, protože jich několik vymazalo: *00 420 602 875 483 -> 0 0 4 **2 6** 0 2 8 7 5 4 8 3* u českého čísla; *00 34 662 00 57 36 -> **0 3**4 6 6 2 **0 5** 7 36* u španělského čísla.

3.2.2.5 Termíny

U termínů jsme brali v potaz to, že tlumočník termín slyšel a automatické titulky používá tedy čistě k tomu, aby ulevil krátkodobé paměti. Celkově se v projevu objevilo 55 unikátních termínů, přičemž některé se opakovaly, a tak se celkový počet výskytů rovnal 72. K hodnocení termínů jsme používali stejná kritéria jako u hodnocení celkové přesnosti.

Graf 12: Počet správně a špatně zapsaných termínů u všech jazyků, vč. závažnosti chyb.

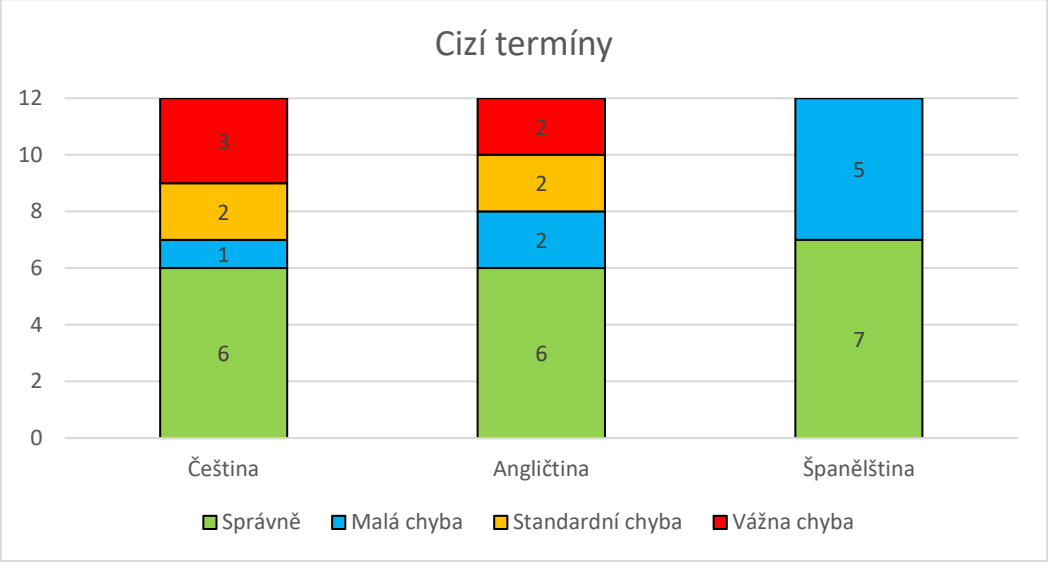
České tituly měly v termínech přesnost při převodu termínů 92 %. Malé chyby spočívaly v nesprávném umístění mezer či vyhláskování (*optogenetice -> obtogenetice, synapsí -> syna psí, somato-senzitivní -> soma to senzivitní*). Do kategorie standardních chyb byly zařazeny případy, kdy došlo k tvoření neexistujících slov: *myelin -> myjeli, sagitální pohled -> zgitální pohled, beta-amyloid -> Beta. A melot*. Vážná chyba pramenila z kompletního vynechání slova (*kognitiva*).

Anglické titulky dosahují přesnosti při převodu termínů 94 %. Jedna vážná chyba pramenila nejspíše z toho, že se mluvčí zakoktal a několikrát se opakoval, což dokazuje, že ASR mívá problém se spontánní řečí (viz 2.4.2.1 *Chyby pocházející z nekvalitního zvukového vstupu*), která je plná hezitací a opakování se (*Acelo-acetylcholinestera- acetylcholinesterasis inhibitors -> Acetal, That's that I. Settle collinear stories inhibitors*). Druhá vážná chyba bylo přepsání *senile plaques* jako *C*. Také je zde ale vidět, že výkon není stálý a záleží na jazykovém kontextu. Termíny *axons, dementia* a *Alzheimer's disease* byly několikrát zapsány správně, ale v jednom z výskytů se u každého termínu vyskytla chyba: *accents, dimension* a *Allzheimer's disease.*

Ve španělských titulcích je přesnost při převodu termínů 96,6 %. Hned dvě standardní chyby přitom pramenily ze špatného zaznamenání zkratek (*EEG -> eje, NMDA -> nma*). Jinak se španělština dopouštěla hlavně menších hláskovácích chyb (*placa alar -> placa lar, bulbo raquídeo -> bulbo raquídio*).

3.2.2.5.1 Cizí termíny

V případě cizích termínů se jednalo o 11 unikátních latinských a řeckých termínů, přičemž se jeden opakoval (*cerebrum*), celkových počet byl tedy 12. Z tlumočnického hlediska jsme nehodnotili, zda je termín zapsán správně (na rozdíl od celkové přesnosti), ale zda ho může tlumočník pouze přečíst a bude správně.



Graf 13: Počet správně a špatně zapsaných cizích termínů u všech jazyků, vč. závažnosti chyb.

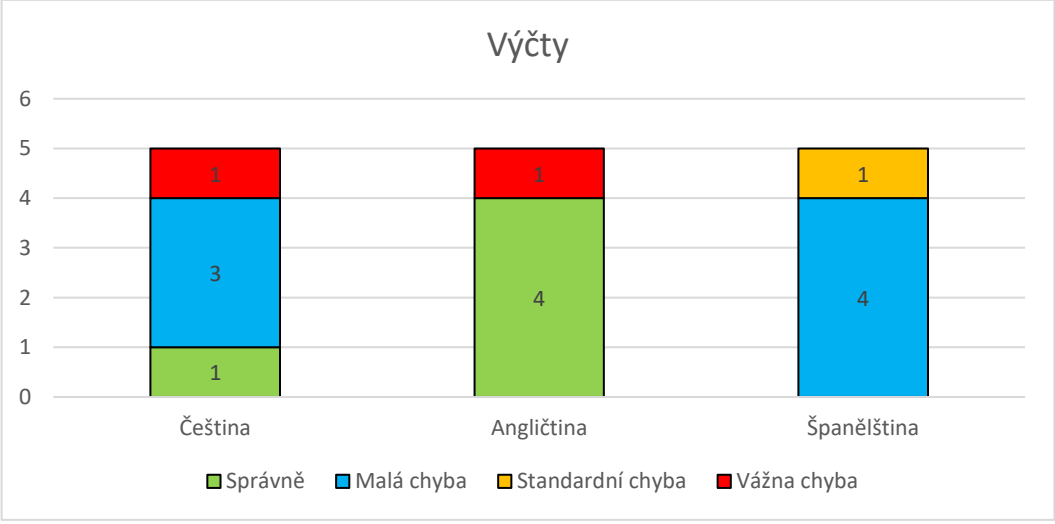
České titulky dosáhly u cizích termínů přesnosti 64,6 %. Některé termíny sice nebyly přepsány úplně správně (*medulla oblongata -> medula oblongáta*), ale tlumočník by je po přečtení stále vyslovil správně. Jakožto standardní chyby jsme vyhodnotili následující: *pons -> ponce* a *cerebellum -> ce rebelům*. Došlo ale ke třem vynechávkám ve výčtu za sebou, což jsme hodnotili jako vážné chyby.

Anglické titulky dosáhly u cizích termínů přesnosti 70,8 %. Jakožto standardní chyby jsme vyhodnotili případy, kdy bylo *cerebrum* zapsáno jako *cerebral* (ačkoli první výskyt tohoto slova byl zapsán správně), a spojení dvou slov *while encephalon -> wellencephalon*. Jakožto vážné chyby jsme podobně jako u českých titulků hodnotili jednu vynechávku a jedno zkrácení (*diencephalon -> D*) za sebou.

Španělské titulky dosáhly u cizích termínů přesnosti 89,5 %. Lze se domýšlet, že takto vysoké skóre pramení z podobnosti mezi latinou a španělštinou, nicméně i řecké termíny byly zaznamenány dostatečně správně (*diencephalon -> dien cefalón*). Občas ale právě kvůli podobnosti ASR měnilo latinské termíny na španělské (*medulla spinalis -> médula espinalis*).

3.2.2.6 Výčty

U výčtů jsme brali v úvahu, zda jsou úplné a zda jsou jednotlivé části jasně oddělené, a tím pádem snadno rozeznatelné. V projevu bylo celkem pět výčtů (v jednom případě obsahoval tři členy, ve čtyřech případech obsahovaly čtyři členy)



Graf 14: Počet správně a špatně zapsaných výčtů u všech jazyků, vč. závažnosti chyb.

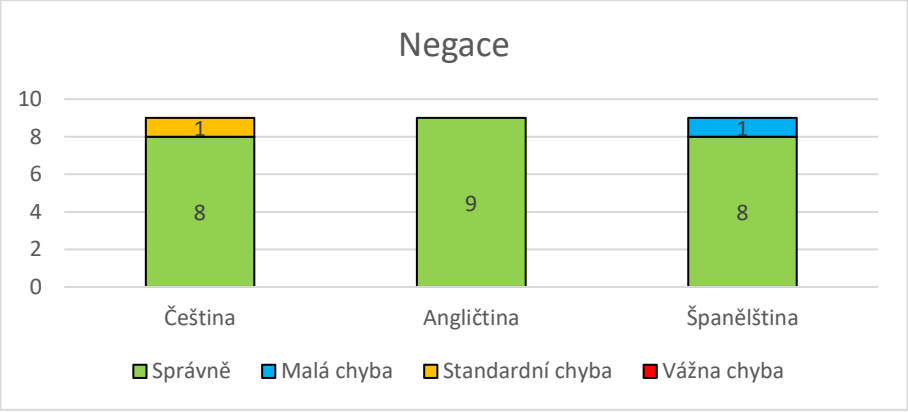
Malé chyby jsme počítali tam, kde byla špatná/chybějící interpunkce, ale lze stále snadno rozeznat jednotlivé části (*dýchací systém. Močové a pohlavní ústrojí. Srdce a lymfatický systém; el pons el mesencefalón el dien cefalón y el telencefalón*).

Jako standardní chybu jsme u španělských titulků viděli výčet jmen. V českých titulcích došlo sice k podobné chybě, ale u češtiny na rozdíl od španělštiny lze jasně určit, co je jméno a co příjmení a kde končí (jména osob se liší od podstatných jmen, ženská příjmení mají koncovku *-ová* apod.). U španělských jmen může být *Francisco* jak jméno, tak příjmení; ne vždy se liší od podstatných jmen, např. *dolores* znamená *bolesti*, ale existuje i jako jméno; a zároveň mohou mít Španělé více než tři jména. Proto jsou interpunkce a správná velká písmena mnohem důležitější. Zatímco jsme to tedy u češtiny vnímali jako malou chybu, u španělských titulků to dle našeho uvážení byla standardní chyba (srov. *Marie Stejskalová bohumila skalová ani Karel bouček* X *juan santiago muñiz francisco javier ruiz y lucía dolores bustamante*).

Za vážnou chybu jsme počítali vynechání některých členů výčtu. K oběma vážným chybám v anglických a českých titulcích došlo ve stejném výčtu, který obsahoval cizí termíny. Domníváme se, že šlo nejspíše o kombinaci cizích termínů a výčtu, protože se cizí termíny objevily těsně za sebou. Jinak totiž v cizích termínech i výčtech dosahuje ASR poměrně dobrých výsledků.

3.2.2.7 Negace

U negací jsme se soustředili hlavně na to, zda byl jasně zaznamenán zápor a zda je přítomné sloveso, ke kterému patří. Celkem se všechny jazykové verze shodovaly v devíti negacích.



Graf 15: Počet správně a špatně zapsaných negací u všech jazyků, vč. závažnosti chyb.

Ve většině případů ASR zápor zaznamenalo a zapsalo správně. U češtiny nicméně došlo k přidání kladného slovesa: *ztráta paměti zcela jistě nepatří přirozeně ke stáří -> ztráta paměti **je** zcela jistě nepatří přirozeně ke stáří*. Nelze přesně říci, proč k tomu došlo, snad kromě pochybení pravděpodobnostního vyhodnocování, protože mluvčí v dané části nezpomalil, nezaváhal ani neřekl nic, co by mohlo sloveso *je* připomínat. U španělštiny došlo k vynechání slovesa, ke kterému zápor patří: *que no por ello es menos importante -> que no por **ellos** menos importante*. Jak vidíme, místo toho bylo *es* přiřazeno k *ello*. Jedná se ale o chybu, kterou si lze lehce domyslet.

3.3 Shrnutí výsledků a závěrečná diskuze

Pro vyhodnocení výsledků jsme použili metriku NER. ASR Zoomu napříč všemi jazyky dosahuje průměrně 95% celkové přesnosti (94,7 % u českých titulků, 97,2 % u anglických titulků, 94,4 % u španělských titulků). Romero-Fresco & Martínez (2015) zmiňují, že respeakeři dosahují hranice 98% přesnosti, což poté uvádí jako požadovanou minimální hodnotu přesnosti pro automatické titulky. Můžeme tedy vidět, že ani angličtina s nejlepšími výsledky této hranice nedosahuje, ačkoli se jí velmi blíží. Můžeme se domnívat, že kdybychom v experimentálním projevu nezahrnovali tolik různých problematických kategorií, celková přesnosti by byla vyšší.

Celkově ASR ve všech jazycích přepisuje dobře domácí jména, termíny, čísla do řádu tisíců a negace. Výčty jsou většinou také zaznamenány správně, jedinou chybou u nich bývá nesprávná interpunkce. Co se týče negativ, roky ani u jednoho z jazyků nemají sjednocený zápis, obtíže činí také telefonní čísla (domníváme se, že se jedná o opakování stejných číslic za sebou). U vyšších čísel (tisíce a výše) nastávají problémy v nesprávném přepisu (obzvláště u španělských titulků), problémová jsou také dvě rozdílná čísla vedle sebe. Pokud se mluvčí zakoktá při výslovnosti čísla, může to vést ke špatnému přepisu. Zkratky a akronymy zvládají jedině anglické titulky, u ostatních jazyků jsou nesprávně zapsány, či jsou vynechána některá písmena. Cizí jména a cizí termíny činí problémy napříč automatickými titulky všech jazyků.

Jednotlivé jazyky dále mají svá vlastní specifika:

- V českých titulcích dochází k chybám v interpunkci, velkých písmenech (a to obzvláště ve zkratkách a akronymech) a gramatických jevech (koncovky, *i/y, mně/mě* apod.). Hlavním negativem pro tlumočníky je slovní zápis čísel, ačkoli díky tomu mají české titulky vyšší správnost. U termínů občas zapisují neexistující slova, místy fonetický přepis ale naopak pomáhá u přepisu cizích jmen a cizích termínů.
- Anglické titulky dosahují velmi dobrých výsledků (u zkratek a akronymů mají dokonce 100% úspešnost), ale největšími problémy se zdají být cizí slova (cizí termíny, ve kterých krátí či naopak spojují slova; a cizí jména, v této kategorii měly více než dvojnásobek vážných chyb oproti ostatním jazykům). Zápis čísel bývá v pořádku, ke špatnému přepisu ale dochází v momentě, kdy jsou dvě čísla vedle sebe, nebo se mluvčí zasekne v jejich výslovnosti.

- Španělské titulky prakticky nepoužívají interpunkci a velká písmena, až na několik málo náhodných výjimek. V kategorii čísel jsou problémem miliony automaticky přepisované jako šest nul. Navíc nezvládají přepis čísel, která obsahují čárku (*25,8 milionu* apod.), a vynechávají znaménka (např. procenta). Mají ale nejvyšší úspešnost v přepisu cizích slov (termínů i jmen).

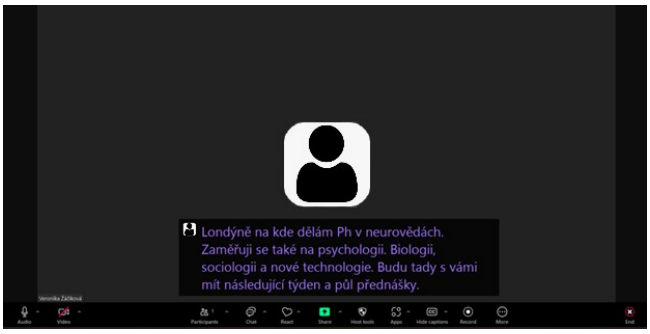
Největším rozdílem v hledisku mezi čtenářem a tlumočníkem se ukázaly být čísla. Zatímco čtenářovi by nejspíše nevadil tolik slovní zápis či mezery mezi čísly, u tlumočníků to může působit zmatení. Velký rozdíl byl také u jmen – zatímco z tlumočnického hlediska jsme hodnotili, nakolik zápis odpovídal foneticky (protože tlumočníci jméno předtím slyšeli a je důležité ho zopakovat správně), u čtenářského hlediska jsme hodnotili správnost, např. pokud by se čtenář rozhodl jméno dohledávat, zda by jméno odpovídalo.

V názvu naší práce jsme si kladli otázku, zda se mohou tlumočníci na automatické titulky spolehnout a v čem jim mohou pomoci. Myslíme si, že hlavní pomoc ASR poskytuje hlavně u domácích jmen, výčtů, negací, v anglických a španělských titulcích také u nižších řádů čísel (pod miliony) a pouze v angličtině u zkratek a akronymů, protože tam zaznamenáváme velmi vysokou přesnost bez přílišných výkyvů. U termínů automatické titulky napříč jazyky také dosahují vysoké přesnosti, ale u této kategorie tlumočníci stále musí spoléhat hlavně na vlastní znalosti (na rozdíl např. u InterpretBank Digital Boothmate). Domníváme se, že u ostatních kategorií (hl. vyšší čísla, cizí jména a cizí termíny) se na automatické titulky zatím nedá zcela spoléhat, protože jejich výkon kolísá.

Defrancq & Fantinuoli (2020) v závěru své studie komentují, že tlumočníci, kteří se mohou spolehnout na vlastní schopnosti, v jejich studii konkrétně při převádění čísel, pomoc ASR ani nevyhledávají. Z toho plyne otázka, zda má ASR u profesionálních tlumočníků vůbec budoucnost. Pokud ale vezmeme v potaz možné pozitivní psychologické dopady v podobě zvýšeného sebevědomí, lepší obeznámenost tlumočníků s tlumočením videokonferencí a CAI nástroji v budoucnu a rostoucí přesnost ASR, domníváme se, že může dojít k vytvoření symbiotického vztahu mezi tlumočníky a ASR.

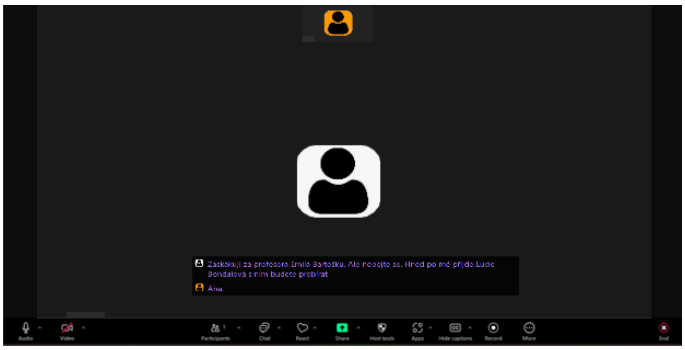
Zda tlumočníci budou automatické titulky využívat nebo ne, ale nezáleží pouze na jejich přesnosti. Velký důraz je také kladen na vizuální prezentaci a rozhraní, v čemž spatřujeme zásadní problém automatických titulků na Zoomu.

V uživatelském rozhraní Zoomu je u automatických titulků možné si upravit velikost písmen a barvu písmen, navíc lze box s titulky přesouvat a roztahovat či zužovat (ale pouze do určité šířky). Výška je pevně dána na 4 řádky (viz obrázek č. 13 níže). Nový text se generuje na spodku a starý text nahoře postupně mizí. Jak poznamenávají Yuan & Wang (2023), tento dynamický charakter automatických titulků často znamená, že tlumočníci nemají mnoho času se zaměřit na jednotlivá slova, a pokud tedy to, co hledají (jména, výčty, čísla apod.) není jasně označeno (velkým písmenem, interpunkcí, číslicí), může dojít snadno k zahlcení jejich kognitivní kapacity.



Obrázek 13: Zobrazení titulků na Zoomu. (vlastní fotka)

Problém také nastává, pokud je přítomen další mluvčí. Jakýkoliv zvuk od druhého mluvčího totiž „usekne“ horní polovinu textu (viz obrázek č. 14 níže). Text tam navíc zůstane poměrně dlouho, tzn. že pokud druhý mluvčí řekne pouhé „*aha*“, toto slovo tam zůstane několik sekund a řeč toho, kdo mluví, se bude zobrazovat pouze na dvou řádcích.



Obrázek 14: Zobrazení titulků na Zoomu v případě dvou mluvčích, menší font. (vlastní fotka)

3.3.1 Srovnání s předchozími výzkumy, limitace našeho výzkumu a budoucí výzkum

Považujeme za relevantní srovnat námi dosažené výsledky s výsledky ostatních výzkumníků. Veškerý dosavadní výzkum, který nám je k dispozici, probíhal na angličtině, proto srovnání omezíme pouze na automatické titulky v angličtině.

Co se týče celkové přesnosti ASR Zoomu, máme k dispozici dva výsledky ke srovnání, konkrétně výzkum Univerzity v Denveru, který zmiňuje 90% přesnost, a Univerzity Santa Barbara, který mluví o 80% přesnosti. Bohužel nám ale nejsou známé konkrétní výpočty ani metodologie, které univerzity k výpočtu použily. Pomocí metriky NER jsme v anglických titulcích zaznamenali celkovou přesnost 97,2 %, která převyšuje oba zmiňované výzkumy. Pokud bychom měli odhadovat, že byla ve výzkumech použita častější a snadnější metrika WER (tedy, že jakýkoli rozdíl oproti řečenému je chyba a všechny chyby mají stejnou váhu), i tak by po přepočtu dosahovalo anglické ASR přesnosti 94,4 %.

V jednotlivých kategoriích máme výsledky ke srovnání u čísel, jmen a termínů. Defrancq & Fantinuoli (2021) u čísel dospěli k přesnosti ASR 96 %, jednalo se o jejich nástroj InterpretBank Digital Boothmate a prokazatelně používali metriku WER. Za použití NER metriky jsme u anglického ASR Zoomu zaznamenali přesnost 95 %, přičemž se jednalo o jednu vážnou chybu a dvě malé. Pokud bychom tento výsledek měli přepočítat pomocí WER metriky, tedy všechny tři chyby považovat za vážné, dospěli bychom k výsledku 90 %. Ačkoli se tedy pohybujeme stále nad 90 %, Zoom v porovnání dosahuje o něco horších výsledků.

U jmen zaznamenali Gaido et al. (2021) přesnost v anglickém ASR 40 %. Ve své studii nezmiňují, o jaká jména či názvy se přesně jednalo, proto nelze určit, zda se jednalo o domácí či cizí jména. Nicméně, na Zoomu jsme v kategorii anglických domácích jmen zaznamenali přesnost 93,3 % a v kategorii anglických cizích jmen přesnost 42,6 %. Pokud bychom opět převedli výsledky z NER metriky na WER metriku, jednalo by se o přesnost 93,3 % v domácích jménech a 23,5 % u cizích jmen. Zatímco tedy přesnost anglických titulků u domácích jmen tyto výsledky značně přesahuje, u cizích jmen je její výsledek naopak horší.

Co se týče termínů, Kustritz et al. (2023) konkrétně u Zoomu zaznamenali kolem 8 medicínsky signifikantních chyb (tedy chyby, při kterých by se studenti medicíny naučili něco špatně). V našem experimentu u anglických titulků došlo k 8 chybám v domácích termínech a 6 chybám

v cizích termínech (u obojího se jednalo o chyby různého stupně závažnosti). Kustritz et al. zkoumali výkon ASR Zoomu na 25minutové lekci, u nás se jednalo o 10minutový projev. Na druhou stranu ale náš projev je záměrně kondenzovaný, aby byl sice kratší, ale obsahoval velké množství termínů. Teoreticky bychom ale mohli uvažovat, že pokud by byl projev ještě delší, došlo by k vícero chybám. Z tohoto hlediska jsme tedy dospěli k horšímu výsledku.

Dále bychom se rádi vyjádřili k limitacím našeho výzkumu. Jasným nedostatkem je subjektivita metriky NER. V ideálním případě bychom měli tým čtenářů a tlumočníků, kteří by se na rozdělení jednotlivých chyb shodli. Zároveň jsme také měli pouze jednoho mluvčího za každý jazyk, tento vzorek by bylo lepší zvětšit.

Z hlediska dalšího výzkumu by bylo zajímavé zjistit, jak tlumočníci s automatickými titulky pracují, co jim dělá potíže a co jim naopak pomáhá. Důležitým faktorem takového výzkumu by ale bylo, aby se jednalo o tlumočníky zvyklé na práci s automatickými titulky. Budoucí výzkum založený na naší práci by také mohl zkoumat, jak různé druhy chyb (malé / standardní / vážné v úpravě / rozpoznání) ovlivňují tlumočníkův výkon a důvěru v ASR. V neposlední řadě je pak pro tlumočníky důležitý také aspekt rychlosti generování automatických titulků (Desmet et al., 2018). ASR se totiž musí trefit do tzv. *ear-voice span (EVS*). Jedná se o zpoždění mezi poslechem výchozího projevu a začátkem tlumočení (jinak také nazývané dekaláž). U tlumočníků se EVS pohybuje okolo 2,5 až 3 sekund (Olerón & Nanpon, 1965/2002), automatické titulky se tedy musí zobrazovat rychleji. Jedná se o hledisko, které jsme v naší práci nerozebírali, proto se možný budoucí výzkum může ubírat i tímto směrem.

4 ZÁVĚR

Tato diplomová práce si kladla za cíl prozkoumat přesnost automatických titulků na platformě Zoom.

V teoretické části jsme se soustředili na vysvětlení přínosu automatických titulků na poli inkluze, kognitivní zátěže v tlumočení, jak ASR zapadá do CAI nástrojů a jak funguje. V neposlední řadě jsme také pohovořili o tom, jak mohou automatické titulky pomoci tlumočníkům.

V empirické části jsme vytvořili experimentální projevy, které zkoumaly celkovou přesnost a přesnost v předem určených kategoriích (jména, zkratky, akronymy, čísla, termíny, výčty a negace) ve třech zkoumaných jazycích (čeština, angličtina, španělština). K vyhodnocení jsme použili metriku NER, díky které můžeme chyby rozdělovat do třech různých úrovní závažnosti a dvou typů podle toho, v čem se chyba vyskytla. Výsledky jsme posuzovali z hlediska čtenáře, ale také z hlediska tlumočníka. Pro oboje hlediska je důležitá přesnost, ale zatímco hledisko čtenáře bere v potaz celkový smysl, hledisko tlumočníka se zaobírá také snadnou vyhledatelností a nenahlíží na text jako celek.

Z hlediska čtenáře jsme dospěli k celkové přesnosti 94,7 % u českých titulků, 97,2 % u anglických titulků a 94,4 % u španělských titulků. Z hlediska tlumočníka považujeme za klíčové vysoké výsledky v přesnosti u domácích jmen (91,7 % u českých titulků, 93,3 % u anglických a 75 % u španělských, zde konkrétně kvůli malým písmenům) a termínů (92 % u českých, 94 % u anglických a 96,6 % u španělských). U čísel lze hovořit o pozitivních výsledcích u anglických titulků (přesnost 95 %); u španělských (přesnost 83,3 %) a českých (0 % kvůli slovnímu zápisu) přesnost nedosahuje dostatečných hodnot.

V závěru jsme identifikovali slabé a silné stránky ASR (jako například interpunkce, znaménka či velká písmena v českých a španělských titulcích) v každém z jazyků a vyjádřili se k tomu, zda a v čem se na něj mohou tlumočníci spolehnout.

5 SEZNAM POUŽITÉ LITERATURY

Odborné prameny

AKHULKOVA, Yulia, Sarah HICKEY a Rosemary HYNES. *Language Technology Atlas* [online]. Nimdzi, 2022. Dostupné z: https://www.nimdzi.com/wp-content/uploads/2022/09/NimdziAtlas22v3.pdf

AKSËNOVA, Alëna, Daan van ESCH, James Flynn a Pavel Golik, 2021. How Might We Create Better Benchmarks for Speech Recognition? *Proceedings of the 1st Workshop on Benchmarking: Past, Present and Future. Association for Computational Linguistics* [online], s. 22–34. Dostupné z: 2021.bppf-1.4.pdf (aclanthology.org)

BAČÁKOVÁ, Jitka, 2023. Služby online tlumočení znakového jazyka a vliv pandemie covid-19 na tuto oblast tlumočení v České republice. Bakalářská práce, vedoucí Andrejsková, Helena. Praha: Univerzita Karlova, Filozofická fakulta, Ústav jazyků a komunikace neslyšících.

BRAUN, Sabine, 2015. Remote Interpreting. *Routledge Handbook of interpreting*. Dostupné z: https://www.researchgate.net/publication/280305163_Remote_Interpreting

COMMISSION OF THE EUROPEAN COMMUNITIES, 2003. Joint report on social inclusion. *Results from the National Action Plans for Social Inclusion (2003-2005).* Dostupné z: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2003:0773:FIN:EN:PDF

CORPAS PASTOR, Gloria a Lily May FERN, 2016. *A survey of interpreters' needs and practices related to language technology* [online]. Dostupné z: https://www.researchgate.net/publication/303685153_A_survey_of_interpreters'_needs_and_practices_related_to_language_technology

CORPAS PASTOR, Gloria, 2018. Tools for interpreters: The challenges that lie ahead. *Current Trends in Translation Teaching and Learning* E, 5, s. 157–182.

ČEŇKOVÁ, Ivana, 2008. Úvod do teorie tlumočení. Česká komora tlumočníků znakového jazyka, o. s., 99 s. ISBN 978-80-87153-74-1.

DEFRANCQ, Bart a Claudio FANTINUOLI, 2020. Automatic Speech Recognition in the Booth. Target. *International Journal of Translation Studies* 33, no. 1, s.73–102. DOI: https://doi.org/10.1075/target.19166.def

DESMET, Bart, Mieke VANDIERENDONCK a Bart DEFRANCQ, 2018. Simultaneous interpretation of numbers and the impact of technological support. In: FANTINUOLI, Claudio (Ed.), *Interpreting and Technology,* s. 13–27. Berlin: Language Science Press. DOI: 10.5281/zenodo.1493281

FANTINUOLI, Claudio, 2017. Speech recognition in the interpreter workstation. *Proceedings of the Translating and the Computer 39*. London. Dostupné z: https://www.researchgate.net/publication/321137853_Speech_Recognition_in_the_Interpreter_Workstation

FANTINUOLI, Claudio, 2018. Computer-assisted Interpreting: Challenges and Future Perspectives. In: CORPAS PASTOR, Gloria a Isabel DURÁN MUÑOZ. *Trends in e-tools and resources for translators and interpreters* [online]. Leiden: Brill, s. 153-174. ISBN 9789004351790.

FENG, Siyuan, Olya KUDINA, Bence Mark HALPERN a Odette SCHARENBORG, 2021. *Quantifying Bias in Automatic Speech Recognition* [online]. Dostupné z: https://arxiv.org/pdf/2103.15122

GAIDO, Marco, Matteo NEGRI a Marco TURCHI, 2022. Who Are We Talking About? Handling Person Names in Speech Translation. *Proceedings of the 19th International Conference on Spoken Language Translation (IWSLT 2022)*, s. 62–73. Association for Computational Linguistics. Dostupné z: https://arxiv.org/pdf/2205.06755

GAIDO, Marco, Susana RODRÍGUEZ, Matteo NEGRI, Luisa BENTIVOGLI a Marco TURCHI, 2021. Is "moby dick" a Whale or a Bird? Named Entities and Terminology in Speech Translation. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, s. 1707–1716. Association for Computational Linguistics. Dostupné z: https://arxiv.org/pdf/2109.07439

GARNERIN, Mahault, Solange ROSSATO a Laurent BESACIER, 2021. Investigating the Impact of Gender Representation in ASR Training Data: a Case Study on Librispeech. *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing*, s. 86-92. Association for Computational Linguistics. Dostupné z: https://aclanthology.org/2021.gebnlp-1.10.pdf

GILE, Daniel, 2009. Basic Concepts and Models for Interpreter and Translator Training. *Benjamins Translation Library*. Amsterdam: John Benjamins Publishing Company. ISBN 9789027224330. DOI: https://doi.org/10.1075/btl.8

GLASSER, Abraham, Kesavan KUSHALNAGAR a Raja KUSHALNAGAR, 2017. Deaf, Hard of Hearing, and Hearing Perspectives on Using Automatic Speech Recognition in Conversation [online]. *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, NY, USA: ACM, 2017, s. 427–432. Dostupné z: https://doi.org/10.1145/3132525.3134781

GRAHAM, Calbert a Nathan ROLL, 2024. Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits [online]. *JASA Express Letters*. 2024, roč. 4, č. 2. ISSN 2691-1191. Dostupné z: https://doi.org/10.1121/10.0024876

GUO, Meng, Lili HAN a Marta Teixeira ANACLETO, 2022. Computer-Assisted Interpreting Tools: Status Quo and Future Trends [online]. *Theory and Practice in Language Studies*, roč. 13, č. 1, s. 89-99. ISSN 2053-0692. DOI: https://doi.org/10.17507/tpls.1301.11

HEIGOLD, Georg, Vincent VANHOUCKE, Andrew SENIOR, Patrick NGUYEN, Marc'aurelio RANZATO et al. Multilingual acoustic models using distributed deep neural networks. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing,* s. 8619-8623. IEEE. ISBN 978-1-4799-0356-6. DOI: https://doi.org/10.1109/ICASSP.2013.6639348

CHMIEL, Agnieszka a Agnieszka LIJEWSKA, 2019. Syntactic processing in sight translation by professional and trainee interpreters: Professionals are more time-efficient while trainees view the source text less. *Target* 31.3, s. 378-397

CHMIEL, Agnieszka, Przemysław JANIKOWSKI a Agnieszka LIJEWSKA, 2020. Multimodal processing in simultaneous interpreting with text: Interpreters focus more on the visual than the auditory modality. Target. *International Journal of Translation Studies*, roč. 32, č. 1, s. 37–58. DOI: https://doi.org/10.1075/target.18157.chm

KAFLE, Sushant a Matt HUENERFAUTH, 2017. Evaluating the Usability of Automatically Generated Captions for People who are Deaf or Hard of Hearing. *ASSETS '17: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, s.165–174. Dostupné z: https://arxiv.org/ftp/arxiv/papers/1712/1712.02033.pdf

KASÍK, Jan, 2023. Nové technologie v tlumočení a budoucí vývoj tlumočnické profese. Diplomová práce, vedoucí Ešnerová, Kateřina. Praha: Univerzita Karlova, Filozofická fakulta, Ústav translatologie.

KATSUMARU, Masaki, Kazunori KOMATANI, Tetsuya OGATA, Hiroshi OKUNO, 2008. Expanding Vocabulary for Recognizing User's Abbreviations of Proper Nouns without Increasing ASR Error Rates in Spoken Dialogue Systems. *INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association*, Brisbane, Australia, 22-26. září, 2008, s. 187–190. Dostupné z: https://www.researchgate.net/publication/221489568_Expanding_Vocabulary_for_Recognizing_User's_Abbreviations_of_Proper_Nouns_without_Increasing_ASR_Error_Rates_in_Spoken_Dialogue_Systems

KOENECKE, Allison, Andrew NAM, Emily LAKE, Joe NUDELL, Minnie QUARTEY et al. Racial disparities in automated speech recognition [online]. *Proceedings of the National Academy of Sciences*. 2020, roč. 117, č. 14, s. 7684-7689. ISSN 0027-8424. Dostupné z: https://doi.org/10.1073/pnas.1915768117

KORPAL, Paweł a Katarzyna STACHOWIAK-SZYMCZAK, 2018. The whole picture: Processing of numbers and their context in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics*, roč. 54, č. 3, s. 335–354. DOI: https://doi.org/10.1515/psicl-2018-0013

KUHN, Korbinian, Verena KERSKEN a Gottfried ZIMMERMANN. Accuracy of AI-generated Captions With Collaborative Manual Corrections in Real-Time. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2023, s. 1–7. ISBN 9781450394222. Dostupné z: https://doi.org/10.1145/3544549.3585724.

LAMBERGER-FELBER, Heike, 2001. Text-oriented research into interpreting: Examples from a case-study [online]. *Hermes – Journal of Language and Communication in Business,* roč. 14(26), s. 39–64. DOI: https://doi.org/10.7146/hjlcb.v14i26.25638

MANKAUSKIENĖ, Dalia, 2016. Problem Trigger Classification and its Applications for Empirical Research. *Procedia - Social and Behavioral Sciences*, 231, s. 143-148. ISSN 1877-0428. Dostupné z: https://www.sciencedirect.com/science/article/pii/S1877042816311946

MANKAUSKIENĖ, Dalia, 2018. Problem triggers in simultaneous interpreting from English into Lithuanian. Summary of Doctoral Dissertation. Vilnius University. Dostupné z: https://www.researchgate.net/publication/325319733_Problem_triggers_in_simultaneous_interpreting_from_English_into_Lithuanian_Summary_of_my_PhD_Thesis

MAZZA, Cristina, 2001. Numbers in simultaneous interpretation. *The Interpreters' Newsletter,* roč. 11, s. 87–104. Dostupné z: https://www.openstarts.units.it/server/api/core/bitstreams/4b10bda2-eacb-460c-85b8-be34c780b4ba/content

MEYER, Bernd, 2008. Interpreting proper names: Different interventions in simultaneous and consecutive interpreting? *Trans-kom*, 1(1), s. 105–122. Dostupné z: https://www.trans-kom.eu/bd01nr01/trans-kom_01_01_08_Meyer_Interpreting_Proper_Names.20080707.pdf

MIRISHKAR, Sai Ganesh, 2022. Building ASR systems for Resource-Rich and Resource-Poor languages. *Interspeech.* Dostupné z: http://www.isca-students.org/sacweb/images/files/resources/8thDC/5_8DC_paper_4534.pdf

MOSER-MERCER, Barbara, 2003. *Remote interpreting: Assessment of human factors and performance parameters.*

MOSER-MERCER, Barbara, 2005. Remote Interpreting: Issues of Multi-Sensory Integration in a Multilingual Task. *Meta*, roč. 50, č. 2, s. 727-738. ISSN 1492-1421. DOI: https://doi.org/10.7202/011014ar

NANJO, Hiroaki a Tatsuya KAWAHARA, 2005. A New ASR Evaluation Measure and Minimum Bayes-Risk Decoding for Open-domain Speech Understanding. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing 2005*. IEEE. Dostupné z: http://sap.ist.i.kyoto-u.ac.jp/EN/bib/intl/NAN-ICASSP05.pdf

NGUEAJIO Mikel a Gloria WASHINGTON, 2022. *Hey ASR System! Why Aren't You More Inclusive? Automatic Speech Recognition Systems' Bias and Proposed Bias Mitigation Techniques. A Literature Review* [online]. Dostupné z: https://arxiv.org/ftp/arxiv/papers/2211/2211.09511.pdf

OLÉRON, Pierre a Hubert NANPON, 1965/2002. Research into Simultaneous Translation. In PÖCHHACKER, Franz a Miriam SHLESINGER (Ed.), *The Interpreting Studies Reader*, s. 42–51. London: Routledge.

PETKAR, Harshalata, 2016. A Review of Challenges in Automatic Speech Recognition [online]. *International Journal of Computer Applications,* roč. 151, s. 23-26. Dostupné z: https://typeset.io/pdf/a-review-of-challenges-in-automatic-speech-recognition-3j7as0m8ku.pdf

PISANI, Elisabetta a Claudio FANTINUOLI, 2021. Measuring the Impact of Automatic Speech Recognition on Number Rendition in Simultaneous Interpreting. *Empirical Studies of Translation and Interpreting.* Routledge. DOI: http://dx.doi.org/10.4324/9781003017400-14

PÖCHHACKER, Franz, 2004. *Introducing Interpreting Studies*. First Edition. Londýn a New York: Routledge. ISBN 0-203-50480-1

PRANDI, Bianca, 2020. The use of CAI tools in interpreter training: Where are we now and where do we go from here? [online]. *inTRAlinea. Special Issue: Technology in Interpreter Education and Practice*. Dostupné z: https://www.intralinea.org/specials/article/2512

RAYAA, Bachir Mahyub a Anne MARTIN, 2022. Remote Simultaneous Interpreting: perceptions, practices and developments. *The Interpreters' Newsletter, 27*.

RODRIGUES, Ana, Rita SANTOS, Jorge FERRAZ ABREU, Pedro BEÇA, Pedro ALMEIDA a Sílvia FERNANDES, 2019. Analyzing the performance of ASR systems: The effects of noise, distance to the device, age and gender. *XX International Conference of Human Computer Interaction (Interacción'19)*, 25-28. červen, 2019, Donostia, Gipuzkoa, Spain. Dostupné z: https://www.researchgate.net/publication/333915170_Analyzing_the_performance_of_ASR_systems_The_effects_of_noise_distance_to_the_device_age_and_gender

ROMERO-FRESCO, Pablo a Juan Martínez PERÉZ, 2015. Accuracy Rate in Live Subtitling: The NER Model. *Audiovisual Translation in a Global Context*, s. 28–50. Palgrave Studies in Translating and Interpreting. Palgrave Macmillan, London. DOI: https://doi.org/10.1057/9781137552891_3

ROOT KUSTRITZ, Margaret, Ryan RUPPRECHT a Perle ZHITNITSKIY, 2023. Comparison of accuracy of machine-generated or human-generated captions of Zoom live lectures in a comparative theriogenology course. *Clinical Theriogenology*, roč. 15, s. 52–56. DOI: https://doi.org/10.58292/ct.v15.9596

ROY, Somnath, 2021. *Semantic-WER: A Unified Metric for the Evaluation of ASR Transcript for End Usability* [online]. Dostupné z: https://arxiv.org/pdf/2106.02016

ROZINER, Ilan a Miriam SHLESINGER, 2010. Much ado about something remote: Stress and performance in remote interpreting. *Interpreting, 12*(2), s. 214-247. DOI: https://doi.org/10.1075/intp.12.2.05roz

SEEBER, Kilian G., 2017. Multimodal Processing in Simultaneous Interpreting. *The Handbook of Translation and Cognition* [online], s. 461-475. Wiley. ISBN 9781119241430. DOI: https://doi.org/10.1002/9781119241485

SEEBER, Kilian G., Laura KELLER a Alexis HERVAIS-ADELMAN, 2020. When the ear leads the eye – the use of text during simultaneous interpretation. *Language, Cognition and*

Neuroscience, roč. 35, č. 10, s. 1480–1494, DOI: https://doi.org/10.1080/23273798.2020.1799045

SHLESINGER, Miriam, 2003. Effects of presentation rate on working memory in simultaneous interpreting. *The Interpreters' Newsletter*, 12, s. 37–50. Dostupné z: https://www.openstarts.units.it/server/api/core/bitstreams/8e4470ef-47f2-4a59-bb95-3dbd60962f7d/content

SPINA, Carli. 2021. Video Accessibility Tools: chapter 2, captions. *Library Technology Reports,* vol. 57, no. 3. Dostupné z: https://journals.ala.org/index.php/ltr/article/view/7526/10401

TJANDRA, Andros, Nayan SINGHAL, David ZHANG, Ozlem KALINLI, Abdelrahman MOHAMED et al., 2023. *Massively multilingual asr on 70 languages: tokenization, architecture, and generalization capabilities* [online]. Meta AI, USA. Dostupné z: https://research.facebook.com/file/1198767717493667/Draft_MultiASR.pdf

TU, Ming, Alan WISLER, Visar BERISHA, a Julie M. LISS. The relationship between perceptual disturbances in dysarthric speech and automatic speech recognition performance [online]. *The Journal of the Acoustical Society of America*. 2016, roč. 140, č. 5, s. EL416-EL422. ISSN 0001-4966. Dostupné z: https://doi.org/10.1121/1.4967208

WANG, Xinyu a Caiwen WANG, 2019. Can Computer-assisted Interpreting Tools Assist Interpreting? Transletters. *International Journal of Translation and Interpreting* [online], 2(3), s. 109–139. ISSN 2605-2954.

YUAN, Lu a Binhua WANG, 2023. Cognitive processing of the extra visual layer of live captioning in simultaneous interpreting. Triangulation of eye-tracked process and performance data [online]. *Ampersand*. 2023, roč. 11. ISSN 22150390. Dostupné z: https://doi.org/10.1016/j.amper.2023.100131

ZOU, Longhui, Michael CARL a Jia FENG, 2022. Patterns of Attention and Quality in English-Chinese Simultaneous Interpreting with Text. *International Journal of Chinese and English Translation & Interpreting*. ISSN 2753-6149. DOI: https://doi.org/10.56395/ijceti.v2i2.50

Internetové zdroje

AI MEDIA, 2023a. What is Respeaking? The Conclusive Guide. *AI MEDIA* [online]. Dostupné z: https://www.ai-media.tv/knowledge-hub/insights/guide-to-respeaking/ [citováno 2024-02-15]

AI MEDIA, 2023b. 5 Great Moments in the History of Closed Captioning. *AI MEDIA* [online]. Dostupné z: https://www.ai-media.tv/knowledge-hub/insights/great-moments-history-captioning/ [citováno 2024-07-25]

BALL, David, 2022. Bridge the Language Gap and Connect People With 8810 Translated Captions. *Zoom* [online]. Dostupné z: https://news.zoom.us/bridge-the-language-gap-and-connect-people-with-zoom-translated-captions/ [citováno 2024-06-11]

DOSHI, Ketan, 2021. Audio Deep Learning Made Simple - Automatic Speech Recognition (ASR), How it Works. *Ketan Doshi Blog* [online]. Dostupné z: https://ketanhdoshi.github.io/Audio-ASR/

EŠNEROVÁ, Kateřina a Barbara TITE, 2020. Jak jsme tlumočily přes Zoom. *Jinými slovy* [online]. Dostupné z: https://www.jinymislovy.cz/jak-jsme-tlumocily-pres-zoom/ [citováno 2024-07-25]

FANTINUOLI, Claudio, 2024. What is Machine Interpreting. *Medium* [online]. Dostupné z: https://medium.com/@info_48425/defining-machine-interpreting-3bb06220f211 [citováno 2024-07-25]

FOSTER, Kelsey, 2023. What is Automatic Speech Recognition? A Comprehensive Overview of ASR Technology. *AssemblyAI* [online]. Dostupné z https://www.assemblyai.com/blog/what-is-asr/ [cit. 2024-2-25]

FOX, Dylan, 2023. Is Word Error Rate Useful? *AssemblyAI* [online]. Dostupné z https://www.assemblyai.com/blog/word-error-rate/ [citováno 2023-12-16]

GOOGLE CHROME, 2024a. Manage captions and translations in Chrome. *Google Chrome* [online]. Dostupné z: https://support.google.com/chrome/answer/10538231?hl=en [citováno 2024-06-11]

GOOGLE CHROME, 2024b. Gestionar subtítulos y traducciones en Chrome. *Google Chrome* [online]. Dostupné z: https://support.google.com/chrome/answer/10538231?hl=es&sjid=4226063105191602010-EU [citováno 2024-06-11]

HUGHES, John, 2022. The Future of Word Error Rate (WER). *Speechmatics* [online]. Dostupné z: https://www.speechmatics.com/company/articles-and-news/the-future-of-word-error-rate [citováno 2024-07-25]

KIESLICH, Karel, 2023. Centrální nervová soustava, úvod (přednáška, 3. LF, 2023). *YouTube* [online]. Dostupné z: https://www.youtube.com/watch?v=DMZlJqGMflM&ab_channel=KarelKieslich [citováno 2024-07-25]

MARTINCOVÁ, Olga a Ivana BOZDĚCHOVÁ, 2017. TERMÍN. In: KARLÍK, Petr, Marek NEKULA a Jana PLESKALOVÁ (eds.), *CzechEncy - Nový encyklopedický slovník češtiny* [online]. Dostupné z: https://www.czechency.org/slovnik/TERM%C3%8DN [citováno 2024-06-05]

MOLLA, Rani, 2020. Microsoft, Google, and Zoom are trying to keep up with demand for their now free work-from-home software. *Vox* [online]. Dostupné z: https://www.vox.com/recode/2020/3/11/21173449/microsoft-google-zoom-slack-increased-demand-free-work-from-home-software [citováno 2024-06-05]

MULLIN, Joe, 2012. Netflix agrees to subtitle all films by 2014. *CNN Business* [online]. Dostupné z: https://edition.cnn.com/2012/10/11/tech/web/netflix-subtitles-settlement/index.html [citováno 2024-07-25]

POŠTA, Miroslav, 2023. Jak se liší titulky pro neslyšící od „normálních“ titulků? *Titulkujeme: O profesionálním titulkování* [online]. Dostupné z: https://titulkujeme.cz/titulky-pro-neslysici/ [citováno 2024-07-25]

REKTOROVÁ Irena, 2020. Mozek a ztracené neurony: Alzheimerova a Parkinsonova nemoc. *YouTube* [online]. Dostupné z:

https://www.youtube.com/watch?v=VobQbAv4BiU&ab_channel=Hv%C4%9Bzd%C3%A1rnaaplanet%C3%A1riumBrno [citováno 2024-07-25]

REV, 2022. What Role Does an Acoustic Model Play in Speech Recognition? *Rev* [online]. Dostupné z: https://www.rev.com/blog/resources/what-is-an-acoustic-model-in-speech-recognition [citováno 2024-07-02]

REV, 2023. What is WER? What Does Word Error Rate Mean? *Rev* [online]. Dostupné z: https://www.rev.com/blog/resources/what-is-wer-what-does-word-error-rate-mean [citováno 2024-06-05]

TIMELINE MEDIA, 2022. Steno Captioning and Voice Writing for Live Television Explained. *YouTube* [online]. Dostupné z: https://www.youtube.com/watch?v=K7G7dUldrzs&ab_channel=TimeLineMedia [citováno 2024-07-25]

UNIVERSITY OF CALIFORNIA, SANTA BARBARA, 2021. Zoom: Closed-Captioning and Transcription of Synchronous and Asynchronous Lectures for Classroom Instruction. *UC Santa Barbara* [online]. Dostupné z: https://help.lsit.ucsb.edu/hc/en-us/articles/360042185611-Zoom-Closed-Captioning-and-Transcription-of-Synchronous-and-Asynchronous-Lectures-for-Classroom-Instruction [citováno 2024-04-04]

UNIVERSITY OF DENVER, 2021. ZOOM Closed captioning and live transcription. *University of Denver* [online]. Dostupné z: https://otl.du.edu/knowledgebase/zoom-closed-captioning-and-live-transcription/ [citováno 2024-04-03]

UNIVERSITY OF MELBROUNE, 2024. Accuracy. *The University of Melbourne* [online]. Dostupné z: https://www.unimelb.edu.au/accessibility/automatic-speech-recognition/accuracy [citováno 2024-07-25]

VERBIT EDITORIAL a Irit OPHER, 2024. Our Research Expert's Take: How Speech Recognition Works in Multilingual Environments. *Verbit* [online]. Dostupné z: https://verbit.ai/general/the-complex-world-of-speech-recognition-in-multilingual-environments/ [citováno 2024-07-25]

WEINGARTOVÁ, Lenka, 2020. Automatic Speech Recognition: Chapter 1. *YouTube* [online]. Dostupné z: https://www.youtube.com/watch?v=vlr_3kNQZAk&ab_channel=NEWTONTechnologies

WORLD HEALTH ORGANIZATION, 2019. Deafness and hearing loss. *World Health Organization* [online]. Dostupné z: https://www.who.int/health-topics/hearing-loss [citováno 2024-07-26]

YOUTUBE, 2023. How (and why) to add captions to your YouTube videos. *YouTube Official Blog* [online]. Dostupné z: https://blog.youtube/news-and-events/caption-my-youtube-videos/ [citováno 2024-07-25]

ZOOM, 2023a. Administración de subtítulos automatizados. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/es/article?id=zm_kb&sysparm_article=KB0062817 [citováno 2024-06-11]

ZOOM, 2023b. K-12 School Time Limit Removal. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/cs/article?id=zm_kb&sysparm_article=KB0062185 [citováno 2024-06-05]

ZOOM, 2023c. Using Language Interpretation in your meeting or webinar. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0064768 [citováno 2024-06-11]

ZOOM, 2024a. Managing automated captions. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0062813 [citováno 2024-06-11]

ZOOM, 2024b. Enabling automated captions. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/cs/article?id=zm_kb&sysparm_article=KB0058810 [citováno 2024-06-11]

ZOOM, 2024c. Enabling and configuring translated captions. *Zoom* [online]. Dostupné z: https://support.zoom.com/hc/cs/article?id=zm_kb&sysparm_article=KB0059081 [citováno 2024-06-11]

6 SEZNAM OBRÁZKŮ A GRAFŮ

Obrázky

Obrázek 1: B. M. Rayaa & A. Martin (2022), dotazníkové šetření remote tlumočníků na téma které platformy používají nejčastěji, od 0 = nikdy jsem na této platformě nepracoval/a až po 5 = na této platformě pracuji nejčastěji. 15
Obrázek 2: Zobrazení automatických titulků v aplikaci Zoom. (Ball, 2022) 16
Obrázek 3: Diagram, v čem se musí jednotlivé funkce nástroje protínat, aby byl považován za CAI nástroj. (Guo et al., 2022) 22
Obrázek 4: Kategorizace CAI nástrojů. (Guo et al., 2022; vl. překlad) 23
Obrázek 5: Výsledky dotazníkového šetření na téma jaké technologie tlumočníci používají během tlumočení. (Pastor & Fern, 2016)..... 24
Obrázek 6: Zjednodušená vizualizace rozkouskování spektogramu na hlásky. 26
Obrázek 7: Opravdová vizualizace dekódování zvukového vstupu za pomoci CTC (Connectionist Temporal Classification, algoritmus pro trénování hlubokých neuronových sítí). 26
Obrázek 8: Rozhraní experimentu s fixací pohledu tlumočníků. (Seeber et al., 2020) 35
Obrázek 9: Výsledky dvou skupin (with reference = se strojovým překladem, without reference = bez strojového překladu) u dvou projevů (E1 a E2). (Wang & Wang, 2019) 36
Obrázek 10: Procentuální výsledky u převodu jmen a názvů pro jednotlivé tlumočnické strategie napříč čtyřmi různými tlumočeními. Meyer (2008) 42
Obrázek 11: Počet obecných chyb a medicínsky signifikantních chyb mezi různými platformami a člověkem. Kustritz et al. (2023)..... 45
Obrázek 12: Procentuální rozpětí podle toho, kolik bylo zachováno přídavných jmen (0-4), v závislosti na rychlosti přednesu (120 a 140 slov za minutu). Schlesinger (2003)..... 46
Obrázek 13: Zobrazení titulků na Zoomu. (vlastní fotka) 78
Obrázek 14: Zobrazení titulků na Zoomu v případě dvou mluvčích, menší font. (vlastní fotka) 78

Grafy

Graf 1: Celková procentuální přesnost ve všech jazycích, vč. procenta chyb dle závažnosti…...58
Graf 2: Rozdělení chyb podle závažnosti a typu ve všech jazycích. 59
Graf 3: Počet správně a špatně zapsaných domácích jmen u všech jazyků, vč. závažnosti chyb. 61

Graf 4: Počet správně a špatně zapsaných cizích jmen u všech jazyků, vč. závažnosti chyb 62
Graf 5: Počet správně a špatně zapsaných domácích a cizích zkratek a vyhláskování u českých titulků, vč. závažnosti chyb. 63
Graf 6: Počet správně a špatně zapsaných domácích zkratek u anglických a španělských titulků, vč. závažnosti chyb. 64
Graf 7: Počet správně a špatně zapsaných akronymů u všech jazyků, vč. závažnosti chyb. 65
Graf 8: Počet správně a špatně zapsaných čísel u všech jazyků, vč. závažnosti chyb. 66
Graf 9: Celková procentuální přesnost zapisování čísel u všech jazyků. 67
Graf 10: Počet správně a špatně zapsaných roků u všech jazyků, vč. závažnosti chyb. 68
Graf 11: Počet správně a špatně zapsaných mocnin, zlomků a telefonních čísel u všech jazyků, vč. závažnosti chyb. 69
Graf 12: Počet správně a špatně zapsaných termínů u všech jazyků, vč. závažnosti chyb. 71
Graf 13: Počet správně a špatně zapsaných cizích termínů u všech jazyků, vč. závažnosti chyb. 72
Graf 14: Počet správně a špatně zapsaných výčtů u všech jazyků, vč. závažnosti chyb. 73
Graf 15: Počet správně a špatně zapsaných negací u všech jazyků, vč. závažnosti chyb. 74

7 PŘÍLOHY

Příloha č. 1: Informovaný souhlas s účastí ve výzkumu 97
Příloha č. 2: Informed consent about research participation 98
Příloha č. 3: Consentimiento informado sobre la participación en la investigación 99
Příloha č. 4: Text originálu – čeština 100
Příloha č. 5: Text originálu – angličtina 102
Příloha č. 6: Text originálu – španělština 105
Příloha č. 7: Text automatických titulků – čeština 109
Příloha č. 8: Text automatických titulků – angličtina 112
Příloha č. 9: Text automatických titulků – španělština 115
Příloha č. 10: Kategorie: čeština – jména, názvy a zkratky 119
Příloha č. 11: Kategorie: angličtina – jména, názvy a zkratky 121
Příloha č. 12: Kategorie: španělština – jména, názvy a zkratky 123
Příloha č. 13: Kategorie: všechny jazyky – akronymy, čísla, termíny a cizí termíny, výčty a negace 125

Příloha č. 1: Informovaný souhlas s účastí ve výzkumu

Informace o výzkumu:

Byl jste požádán o přečtení předem připraveného textu pro účely diplomové práce s názvem „Přesnost automatických titulků generovaných platformou Zoom – je na ně spoleh?“. Tato práce si klade za cíl prozkoumat, jakou přesnost mají automatické titulky na platformě Zoom.

Projev budete přednášet během opravdové schůzky na Zoomu, kde budete jediným účastníkem. Z vašeho projevu bude pořízena audio i videonahrávka, které budou sloužit výhradně pro zpracování této diplomové práce a nebudou dostupné nikomu jinému než řešitelce diplomové práce. V diplomové práci bude k dispozici pouze přepis vyhotovený řešitelkou pořízený na základě nahrávek.

V případě dotazů se můžete obrátit na řešitelku práce na adrese: veronika.zacikova@seznam.cz

Prohlášení

Já níže podepsaný potvrzuji, že

a) jsem se seznámil s informacemi o cílech a průběhu výše popsaného výzkumu;
b) dobrovolně souhlasím s účastí své osoby v tomto výzkumu;
c) rozumím tomu, že se mohu kdykoli rozhodnout ve své účasti na výzkumu nepokračovat;
d) jsem srozuměn s tím, že jakékoliv užití a zveřejnění dat a výstupů vzešlých z výzkumu nezakládá můj nárok na jakoukoliv odměnu či náhradu, tzn. že veškerá oprávnění k užití a zveřejnění dat a výstupů vzešlých z výzkumu poskytuji bezúplatně.

Zároveň souhlasím s (nehodící se škrtněte):

a) předneseným a nahráním projevu ANO/NE

b) použitím přepisu projevu v textu diplomové práce ANO/NE

c) zveřejněním svého jména v textu diplomové práce ANO/NE

Potvrzuji, že jsem převzal podepsaný stejnopis tohoto informovaného souhlasu.

Dne:

Jméno účastníka výzkumu:

Podpis:

Příloha č. 2: Informed consent about research participation

About the research:

You have been asked to read a speech prepared beforehand for the purpose of the master's thesis entitled "The accuracy of automated captions generated by Zoom – are they reliable?". The thesis aims to test the accuracy of Zoom's automatic captions.

You will deliver the speech during a real Zoom meeting of which you will be the sole participant. An audio and videorecording will be made during the speech, which will be used exclusively for the purpose of this thesis and will not be made available to anyone else except for the author of this thesis. The only part available in the thesis will be a written transcript of the speech created by the author based on the recordings.

If you have any questions, you can contact the author of this thesis at: veronika.zacikova@seznam.cz

Statement

I, the undersigned, certify that

a) I have read and understood the information about the aims and conduct of the research described above;
b) I voluntarily agree with my participation in this research;
c) I understand that I may choose to opt out from the research at any time;
d) I understand that any use and publication of the data and findings from the research does not entitle me to any remuneration or compensation, i.e. I grant all rights to use and publish the data and findings resulting from the research free of charge.

I also agree with (delete as appropriate):

a) delivering and recording the speech YES/NO

b) use of the speech transcript in the text of the thesis YES/NO

c) publishing my name in the text of the thesis YES/NO

I confirm that I have received a signed copy of this informed consent.

Date:

Name of the participant:

Signature:

Příloha č. 3: Consentimiento informado sobre la participación en la investigación

Sobre la investigación:

Se le ha pedido que lea un discurso preparado de antemano para la tesis de máster titulada "La precisión de los subtítulos automáticos generados por Zoom – ¿son fiables?". El objetivo de la tesis es comprobar la precisión de los subtítulos automáticos de Zoom.

Pronunciará el discurso durante una reunión real de Zoom en la que usted será el único participante. Se realizará una grabación de audio y vídeo durante el discurso, que se utilizará exclusivamente para los fines de esta tesis y no estará disponible a nadie más que la autora de la tesis. La única parte disponible estará una transcripción escrita del discurso creada por la autora basándose en las grabaciones.

Si tiene alguna pregunta, puede ponerse en contacto con la autora de la tesis en: veronika.zacikova@seznam.cz

Declaración

Yo, el abajo firmante, declaro que

a) he leído y comprendido la información sobre los objetivos y la realización de la investigación descrita anteriormente;
b) acepto voluntariamente mi participación en esta investigación;
c) entiendo que en cualquier momento puedo optar por no participar en la investigación;
d) entiendo que cualquier uso y publicación de los datos y resultados de la investigación no me da derecho a ninguna remuneración o compensación, es decir, concedo de forma gratuita todos los derechos para utilizar y publicar los datos y resultados de la investigación.

También estoy de acuerdo con (tache según proceda):

a) pronunciar y grabar el discurso SÍ/NO

b) el uso de la transcripción del discurso en el texto de la tesis SÍ/NO

c) la publicación de mi nombre en el texto de la tesis SÍ/NO

Confirmo que he recibido una copia firmada de este consentimiento informado.

Fecha:

Nombre del participante:

Firma:

Příloha č. 4: Text originálu – čeština

Dobré ráno, já jsem Kiefer Schneider, to je německy a píše se to „k-i-e-f-e-r s-ch-n-e-i-d-e-r“. Kdysi jsem tu taky studoval, to ještě existovala NDR, a teď tu učím anatomii. Aktuálně pobývám v Londýně na UCL, kde dělám PhD v neurovědách, zaměřuji se také na psychologii, biologii, sociologii a nové technologie. Budu tady s vámi mít následující týden a půl přednášky na centrální nervový systém. Zaskakuji za profesora Emila Bartošku, ale nebojte se, hned po mně přijde Lucie Bohdalová. S ní budete probírat dýchací systém, močové a pohlavní ústrojí, srdce a lymfatický systém. Také společně proberete AIDS.

Když se snažíme pochopit systémy v těle, je dobré si vytvořit metaforu. Minule jsme mluvili o gastrointestinálním systému, tam máme trubici, která rozmělňuje jídlo a přetavuje ho na molekuly, které vstřebává. Mozek je trochu jiný, není to žádná lehká hydrolýza esterů. Za dob SSSR jsme o tom tolik nevěděli, ale teď v době AI si k ní mozek přirovnáváme, myslíme, že sbírá spoustu dat a na jejich základě predikuje. S tímto modelem konkrétně přišel Jaromír Doležal. Podle všeho používáme jen asi tak 1/10 kapacity našeho mozku.

Že jsme vůbec pochopili, jak mozek funguje, nám umožnila taková nenápadná nudná obyčejná věc: histologické barvení, díky kterému jsme byli schopní nabarvit mozek a vidět neurony. Těch je určitě více než 1 676, jak jsme se kdysi domnívali. A díky tomuto speciálnímu barvení, které vymysleli Anthony Goldie a Lukáš Kachal, jsme viděli, že to nejsou jen tak ledajaké buňky, ale že mezi sebou mají nějaké spoje, někde mezi 1 545 a 20 000 spoji. Dále tuto metodu pak rozvinul Juan Luis Jorge z univerzity v Madridu ve spolupráci s čínským vědcem jménem Yu Xin, který působí na univerzitě v Hangzhou. Když přeskočím tak 200 let, tak jsme se dostali až k optogenetice, slyšeli jsme o tom někdy?

Pracujeme s geneticky upravenými myšmi, kterým dáme nějaké geny, takže jejich proteiny, jejich transmitery reagují na světlo. Když na ně pak posvítíme světlem o určité vlnové délce, kolem 550 nanometrů, tak se tím ten neuron aktivuje a my jsme tím schopní ovládat populace neuronů, které exprimují ten daný gen. Mimochodem, 1 nanometr je 10^{-9} metrů. Můžeme to použít i u lidí, dáme jim elektrický EEG senzor, který snímá jejich mozkovou aktivitu. Dělá se to v rámci projektu HjerneProjekt, který vede Honza Šedivý, ředitel institutu Hjärnfonden ve městě Östersund. Lidé tak mohou ovládat třeba robotickou paži, můžete se podívat na videa na BBC, případně na americké televizi MTV. Tato technologie se dá uplatnit taky třeba u FBI, v Interpolu nebo v NASA. UNICEF má vlastní projekt Ujana Salama, na který bylo vyhrazeno kolem 5,656 miliard euro. Z toho 25,8 milionů čerpáme my v rámci tuzemského projektu a další 1,75 milionu je věnováno na podporu výzkumu. A dále čekáme také na grant z UNESCO.

Z čeho se ale mozek skládá? V šedé hmotě jsou těla neuronů, a protože jsou u sebe těsně, tak to vypadá tmavší. Mozek má asi 10^{11} neuronů a kolem 19 857 synapsí. 78,5 % tvoří voda, ne 20 %, jak se kdysi říkalo. V bílé hmotě jsou axony a ty axony, když mají kolem sebe myelin, tak to vypadá světleji. Mozek se řekne cerebrum, mícha se řekne medulla spinalis. Mozek pak dělíme na další části, důležitá je medulla oblongata, říká se jí také prodloužená mícha. Pak tam máme pons, mesencephalon, diencephalon a telencephalon. Chybí pak jen cerebellum, jen v něm je 69,3 miliard neuronů. Všimněte si, že máme dvě slova, cerebrum, což je latinsky a encephalon, to je řecky. Tady na obrázku vidíme sagitální pohled, díky tomu můžeme vidět část neokortexu a uvnitř nuclei. Bílou hmotu tvoří dráhy, které vedou přes axony až do dalšího neuronu. Podle jejich směru je také pojmenováváme, například dráha z kortexu do míchy by byla tractus corticospinalis, to se píše jako „t-r-a-c-t-u-s mezera c-o-r-t-i-c-o-s-p-i-n-a-l-i-s“.

A z čeho vůbec vzniká CNS? Z vnějšího ektodermu vznikne neurální brázda a ta se zanoří, až vytvoří trubici. Část z toho se oddělí a vznikne z toho neurální lišta. Další důležité struktury jsou alární ploténka, která přijímá signál, tedy je somato-senzitivní, a bazální ploténka, která je motorická.

Pojďme se teď podívat na různá onemocnění mozku. Budeme postupně v rámci hodin mluvit o Alzheimerově nemoci, Parkinsonově nemoci, epilepsii a roztroušené skleróze. Alzheimer- pardon, Alzheimerova nemoc je spojená s demencí, víme že v roce 2010 bylo na světě 35 milionů lidí trpících demencí. Za dalších 20 let se ten počet zdvojnásobí, tzn. že v roce 2030 mluvíme asi o 65 milionech lidí s demencí, 2050 už asi 113 milionů, 2/3 z nich ve vyspělých zemích. Jen dnes na světě žije přesně 8 103 926 441 lidí, ale čísla rostou a odhaduje se, že tou dobu bude na planetě žít asi 9,7 miliard lidí, takže daný počet se nezdá jako příliš veliký, ale stále se jedná o vysoké procento. Navíc, poplatky za léčbu také neustále rostou, v roce 2019 zaplatily pojišťovny průměrně za pacienta 54 520 korun, celkem za celý rok se tak částka vyšplhala na 1 480 116 000 korun. Nejčastější příčinou demence je právě Alzheimerova nemoc, konkrétně v 75 % všech případů demence.

Co to ale je, Alzheimerova nemoc? Je to mimo jiné porucha krátkodobé paměti – ztráta paměti zcela jistě nepatří přirozeně ke stáří. Zároveň touto nemocí netrpí pouze staří lidé, může ji mít i třicátník. Mnozí se také domnívají, že nemoc není smrtelná. Onemocnění ničí mozkové buňky, mozek degeneruje, mění se hlavně nervové přenašeče v mozku pro ukládání paměťové stopy, konkrétně acetylcholinu a glutamátu. Na jejich základě jsme vytvořili přípravky, tzv. kognitiva, která slouží k léčbě pacientů, jednak inhibitory acetylcholinesterázy, jednak antagonisté NMDA receptorů. Dokáží pozastavit průběh nemoci na zhruba 1 až 5 let. Pomocí biomarkerů pozorujeme i další změny, např. beta-amyloid, který se ukládá v senilních placích.

Nemoc dostala jméno podle Aloise Alzheimera, který ji poprvé popsal u pacientky Auguset Dieter v roce 1901. 1907 pak popsal neuropatologický obraz, ale teprve 1910 Atiquzzama Khan v Mnichově zmiňuje Alzheimerovu nemoc. Dieter přitom trpěla vzácnou variantou, obvykle se nemoc projevuje po 60, ale ona ji měla geneticky mnohem dříve, to je u 2-3 % pacientů. Čestnost výs- četnost výskytu také roste s věkem, kolem 65 let je postiženo 1-5 % lidí, ve věku 90 let až 1/2 populace. U nás žije 724 lidí, kterým je nad 100 let, to znamená, že minimálně 362 z nich trpí touto nemocí.

O kom se mluví méně, ale je neméně důležitý, je Oscar Fischer. Totiž, ona by se ta nemoc měla jmenovat Alzheimer-Fischerova nemoc, protože on už tou dobou, v roce 1897, popsal kolem 15 pacientů s touto nemocí. Bohužel, Marie Stejskalová, Bohumila Lískalová ani Karel Bouček nebyli jedni z nich, protože nebyli z Prahy ani z Brna, pocházeli z malé vesnice Plešovice, u které se poblíž nachází Dalešická přehrada.

Pokud vás zajímá více, ozvěte se mi na mail, pošlete mi SZ v aplikaci Douyin, nebo na číslo +420 602 875 483, případně na číslo mého španělského kolegy z Córdoby +34 662 00 57 36.

Příloha č. 5: Text originálu – angličtina

Hello everyone, my name is Kiefer Schneider, and you spell that k-i-e-f-e-r s-c-h-n-e-i-d-e-r. I'm German, but I studied here when the NDR still existed and now I'm back and teaching anatomy. I live in London, and I am a PhD student at UCL in neuroscience. But I also dabble in psychology, biology, sociology and new technology. I'll be teaching for a week and a half about the central nervous system. I am subbing in for Professor John Fowler. When my session is over, you'll have Lucy Morgan and with her, you'll be looking at the respiratory, urinary and reproductive systems, as well as the heart and lym- /inadible/ I apologise, as well as the heart and lymphatic system. You'll also be discussing AIDS.

Now, it's good to use metaphors to try to understand the human body. It looks like you've been talking about the gastrointestinal system – with that, basically there's a tube that takes food and makes it into molecules that it then absorbs. The brain works a bit differently; it makes ester hydrolysis look pet- pretty easy in comparison to be honest. Now, back when the USSR still existed, we didn't know much about the brain. Now, in the era of AI, we can draw parallels between it and the brain. Essentially, it collects tons of data and based on that, it makes predictions. Henry Fletcher is the one who came up with that specific model. Oh, and it seems we only use about 1/10 of our brain capacity.

Now, the fact that we even understand how the brain works was made possible thanks to a single boring, mundane thing: histological staining. Because of this, we can mark the brain and actually see neurons. As opposed to what we had- opposed to- as opposed to what we had previously believed, there's certainly more than 1,676 of them. And thanks to this special staining, which was created by Anthony Goldie and Lukasz Kachal, we saw that they're- that they are no ordinary cells, but that they're connected, through somewhere between 1,545 and 20,000 connections. This method was then further developed by Juan Luis Jorge from the University of Madrid in cooperation with the Chinese scientist Yu Xin, who works out of a university in Hangzhou. Now let me skip about 200 years, so we can get to optogenetics – have you guys ever heard about it?

Well, we work with genetically modified mice: we give them genes, so that their proteins, their transmitters react- react to light. Then, if we shine a light of a certain wavelength (around 550 nanometers) on them, it'll activate the neuron and thus, we are able to control the neuronal population that expresses that specific gene. Oh, by the way, 1 nanometer is equal to 10^{-9} of a meter. We can also do this with people, we give them an electric EEG sensor that records their brain activity. This is what is being done at the HjerneProjekt, led by Judith Fairfield, the director of the institute Hjärnfonden in the city of Östersund. Using this technique, people can control a robotic arm – you can watch some videos on BBC or MTV. The technology may also be used by the FBI, Interpol, and NASA. UNICEF has a funding of it- UNICEF is funding its own project called Ujana Salama, with 5.656 billion euro. Out of that, we drew 24- we drew 25.8 million for our local project and another 1.75 million has been ear-marked for research support. We're also awaiting a grant from UNESCO.

But what exactly constitutes the brain? There are bodies of neurons in the grey matter, and because they're so close to each other, it appears darker. The brain has around 10^{11} of neurons and around 19,587 synapses, plus, it's around 78.5% water, not the 20% we were told back in the day. The white matter is formed by axons and these axons, because they are surrounded by myelin, appear lighter. The brain is called the cerebrum, the spinal cord is the medulla spinalis. We can further divide the brain into other parts – an important one is- is the medulla oblongata, also called the bulb. Up next is the pons, mesencephalon, diencephalon and telencephalon. The

last part is the cerebellum – it alone contains 69.3 billion neurons. It might be interesting to note that of these two words for the brain, cerebrum comes from Latin, while encephalon is Greek. In this picture, we can see the sagittal view that shows us a part of the neocortex and nuclei. The white matter is con- constituted of pathways that go through the axons, all the way into the next neuron. If we observe their direction, we can name them – for example the pathway from the cortex to the spinal cord would be tractus corticospinalis, and you spell that t-r-a-c-t-u-s c-o-r-t-i-c-o-s-p-i-n-a-l-i-s.

So, how is the CNS created? A neural groove emerges from the external ectoderm, and it deepens to form a tube. A part of that detaches to form the neural crest. Other important structures include the alar plate, which receives signals and as such is somato- is somatosensory, and the basal plate, which is motor.

Now let's look at various brain diseases. We're going to talk about Alzheimer's disease, Parkinson's disease, epilepsy and multiple sclerosis, respectively. Alzheimer's is connected to dementia. We know that in 2010, there were 35 million people in the world suffering from dementia. In the next 20 years, that number should double, meaning that by 2030, we are talking about 65 million people with dementia, by 2050 about 130- and by 2050 about 113 million – 2/3 of those living in developed countries. There are exactly 8,103,926,441 people living in the world today, but these numbers are growing, and it’s estimated that there are going to be 9.7 billion people by 2050. The numbers might not sound so big then, but it’s still a large percentage. Additionally, treatment fees have also been steadily rising. In 2019, insurance companies paid an average of 54,520 crowns per patient, which made a yearly total of 1,480,116,000 crowns. Alzheimer's disease is the most common cause of dementia, specifically in 75% of all dementia cases.

But what exactly is Alzheimer's disease? It is, among other things, a short-term memory disorder – memory loss does not come as a natural part of growing old. Also, it’s not just old people who suffer from this disease, younger people can have it too. Many also believe that the disease- that the disease is not terminal. The disease destroys brain cells, the brain degenerates. These changes mostly affect the neurotransmitters in the brain in charge of storing memory traces, namely acetylcholine and glutamate. Based on these, we have developed medicines, called cognition-

enhancing drugs, to treat patients, to bost-though- I apologize, through both acelo- acetylcholineestera- acetylcholinesterase inhibitors and NMDA receptor antagonists. They can halt the course of the disease for about 1 to 5 years. Using biomarkers, we can observe other changes, such as beta-amyloid, which is stored in senile plaques.

Now, the disease was named after Alois Alzheimer, who first described it in his patient August-Auguset Dieter in 1901. He then described the neuropathological picture in 1907, but it was only in 1910 that Alzheimer's disease was mentioned by Atiquzzama Khan in Munich. In fact, Dieter suffered from a rare variant: the disease usually manifests after the age of 60, but she had it much earlier due to genetic causes, which happens in 2-3% of patients. The incidence does- the incidence does increase with age, with 1-5% of people affected around 65, and up to 1/2 of the population at the age of 90. In our country, there are 724 people who are over 100 years old, which means that at least 362 of them suffer from this disease.

Now, there's one other person who is less talked about but is no less important: Octa- Octavia Velazquéz. Actually, the disease should be called Alzheimer-Velazquéz's disease, because around that time, in 1897, she'd already described around 15 patients with the disease. Unfortunately, these didn't include Mary Ackland, Mitch Banner and Sandy Duncan, because they weren't from San Francisco or Chicago, they came from the small village called Lavenham near Norwich.

Now, if you are interested in learning more, please contact me by email, on Douyin or call me at /country code/ 420 602 875 483, or my Spanish colleague from Córdoba at country code 34 662 00 57 36.

Příloha č. 6: Text originálu – španělština

Buenos días, me llamo Kiefer Schneider, es un nombre alemán y se escribe «k-i-e-f-e-r s-c-h-n-e-i-d-e-r». Estudié aquí cuando la RDA todavía existía, y ahora enseño anatomía. Ahora mismo vivo en Barcelona, estudio en la UB y hago PhD. en neurociencias, también me dedico a la psicología, la biología, la sociología y las nuevas tecnologías. Durante la próxima semana y media les daré clases sobre el sistema nervioso central. Sustituyo al profesor Felipe Manceñido, pero no se preocupen, justo después de mí vendrá Patricia Clares Naveros. Con ella, aprenderán

sobre el sistema respiratorio, el sistema urinario y reproductor, el corazón y el sistema linfático. Juntos, también hablarán del SIDA.

Cuando tratamos de comprender los sistemas corporales, es útil recurrir a las metáforas. La última vez, hablaron del sistema gastrointestinal, que funciona como un tubo que desmenuza los alimentos y los convierte en moléculas que luego absorbe. El cerebro es un poco diferente, no se trata de una sencilla hidrólisis de ésteres. Durante la existencia de la URSS, sabíamos poco de él, pero ahora, en la era de la IA, lo comparamos con ella: consideramos que recoge montones de datos y hace predicciones basándose en ellos. Fue Franco Díaz Martín quién creó este modelo. Parece que sólo usamos 1/10 de la capacidad de nuestro cerebro.

Fue posible llegar a entender cómo funciona el cerebro gracias a una cosa común, aburrida y poco llamativa: la tinción histológica. Gracias a ella logramos teñir el cerebro y ver las neuronas: por cierto hay más de 1.676, a diferencia de lo que se pensaba. Y gracias a esta tinción especial, creada por Anthony Goldie y Lukasz Kachal, vimos que no son células comunes, sino que están conectadas, con un número de conexiones comprendido entre 1.545 y 20.000. Más adelante, Juan Luis Jorge desarro- desarrolló el método en la universidad de Madrid, en colaboración con un científico chino llamado Yu Xin, que trabajaba en la universidad de Hangzhou. Si avanzamos unos 200 años, llegamos a la optogenética, ¿saben qué es?

Trabajamos con ratones modificados genéticamente, les damos genes para que sus proteínas, sus transmisores, reaccionen a la luz. Si les iluminamos con una luz de una cierta longitud de onda, de unos 550 nanómetros, se activará la neurona y así seremos capaces de controlar la población de neuronas que expresa el gen dado. A propósito, 1 nanómetro equivale a 10^{-9} de un metro. También podemos utilizar eso en humanos, les damos un sensor EEG eléctrico que registra su actividad cerebral. Lo hacen dentro del HjerneProjekt, dirigido por Alberto Jiménez Lopez, director del instituto Hjärnfonden de la ciudad de Östersund. De esta manera, las personas pueden controlar un brazo robótico – pueden ver vídeos al respecto en las cadenas BBC o MTV. También utilizan esta tecnología el FBI, la Interpol y la NASA. UNICEF tiene su propio proyecto, Ujana Salama, al que se asignaron 5.656 millones de euros. De esa suma, cobramos 25,8 millones para nuestro proyecto local y 1,75 millones se dedican a apoyar la investigación. También esperamos una subvención de la UNESCO.

Pero ¿qué constituye el cerebro? La sustancia gris se compone de cuerpos de neuronas, y como están tan cerca las unas de las otras, parece más oscura. El cerebro tiene unas 10^{11} neuronas y unas 19.587 sinapsis, el 78,5% es agua, no el 20% como se afirmaba antes. La sustancia blanca está formada por axones y estos axones, dado que están rodeados de mielina, parecen más claros. El cerebro se denomina cerebrum, la médula espinal es medulla spinalis. A continuación, dividimos el cerebro en otras partes: una importante es la medulla oblongata, también llamada bulbo raquídeo. Luego, están el pons, el mesencephalon, el diencephalon y el telencephalon. El último es el cerebellum, que contiene 69.300 millones de neuronas. Fíjense en que hay dos palabras: cerebrum, que viene del latín, y encephalon, que es griego. En esta imagen, podemos ver la vista sagital, que permite observar parte del neocórtex y los nuclei. La sustancia blanca está compuesta por vías que van por los axones hasta la neurona siguiente. Podemos nombrarlas según su dirección, por ejemplo la vía que va del córtex a la médula espinal sería tractus corticospinalis, lo que se deletrea «t-r-a-c-t-u-s c-o-r-t-i-c-o-s-p-i-n-a-l-i-s».

Ahora bien, ¿cómo nace el SNC? Empieza cuando un surco neural emerge del ectodermo externo y se profundiza conformándose así un tubo. Parte de él se desprende para formar la cresta neural. Entre otras estructuras importantes figuran la placa alar, que recibe señales y, como tal, es somatosensorial, y la placa basal, que es motora.

Veamos ahora las distintas enfermedades del cerebro. En clase vamos a hablar de la enfermedad de Alzheimer, la enfermedad de Parkinson, la epilepsia y la esclerosis múltiple sucesivamente. La enfermedad de Alzheimer está relacionada con la demencia. En 2010 había 35 millones de personas en el mundo que padecían demencia. En los 20 años siguientes, esa cifra se duplicará, lo que significa que en 2030 estaremos hablando de unos 65 millones de personas con demencia, y en 2050 de unos 113 millones, 2/3 de ellas situadas en países desarrollados. Hoy en día viven exactamente 8.103.926.441 personas en el mundo, pero las cifras no dejan de aumentar y se estima que habrá 9.700 millones de personas para el año 2050. Si bien el número de personas afectadas por la enfermedad no parece enorme en comparación con esta cifra, es un porcentaje importante. Además, las tasas de tratamiento también están aumentando sin parar: en el año 2019, las compañías de seguros pagaron un promedio de 54.520 coronas por un paciente, por lo que el gasto total anual ha aumentado a 1.400.116.000 coronas. La enfermedad de Alzheimer es

la causa más común de demencia; concretamente representa el 75% de todos los casos de demencia.

Pero ¿qué es la enfermedad de Alzheimer? Entre otras cosas, se trata de un trastorno de la memoria a corto plazo – la pérdida de memoria no es, en ningún caso, una parte natural del envejecimiento. Al mismo tiempo, no son sólo las personas mayores las que padecen esta enfermedad: también afecta a personas jóvenes. Además, muchos creen que la enfermedad no es mortal. La enfermedad destruye las células cerebrales, el cere- y el cerebro se degenera, los cambios afectan sobre todo a los neurotransmisores cerebrales encargados de almacenar las huellas de la memoria, en concreto la acetilcolina y el glutamato. Basándonos en ellos, hemos desarrollado medicamentos, llamados potenciadores cognitivos, para tratar a los pacientes, mediante inhibidores de la acetilcolinesterasa y antagonistas de los receptores NMDA. Estos medicamentos logran detener el desarrollo de la enfermedad entre 1 y 5 años. Gracias a los biomarcadores, podemos observar otros cambios, como el beta-amiloide, que se almacena en las placas seniles.

La enfermedad debe su nombre a Alois Alzheimer, que la describió por primera vez en una paciente, Auguset Dieter, en el año 1901. Luego describió el cuadro neuropatológico en 1907, pero no fue hasta 1910 cuando Atiquzzama Khan mencionó la enfermedad en Múnich. De hecho, Dieter padecía una variante poco frecuente. La enfermedad suele manifestarse después de cumplir los 60 años, pero ella la desarrolló mucho antes por causas genéticas, lo que ocurre en un 2 o 3% de los casos. La incidencia también aumenta con la edad, con un- entre 1-5% de afectados en torno a los 65 años, y hasta 1/2 de la población a los 90 años. En nuestro país, hay 724 personas mayores de 100 años, lo que significa que al menos 362 de ellas padecen esta enfermedad.

Hay una persona de la que se habla menos, pero que no por ello es menos importante: Oscar Fischer. En realidad, la enfermedad debería llamarse enfermedad de Alzheimer-Fischer, porque en 1897, ya había descrito a unos 15 pacientes con esta enfermedad. Por desgracia, Juan Santiago Muñez, Francisco Javier Ruíz y Lucía Dolores Bustamante no formaron parte de ellos, porque no eran de Salamanca o Bilbao, sino que vivieron en un pequeño pueblo de Sorihuela del Guadalimar, cerca de Jaén.

Si les interesa saber más, contáctenme por correo electrónico, en Douyin o llámenme al 00 34 6-6-2 0-0 5-7 36, o a mi colega checo de Brno al 0-0-4-2-0 6-0-2 8-7-5 4-8-3.

Příloha č. 7: Text automatických titulků – čeština

Legenda: světle modrá = malá chyba v úpravě, tmavě modrá = malá chyba v rozpoznání, žlutá = standardní chyba v úpravě, tmavě zelená = standardní chyba v rozpoznání, červená = vážná chyba v rozpoznání , 0 = vynecháno slovo

Dobré ráno. Já jsem Kefr. Schneider. To je německy. A píše se to kádí Efers. Chm. Mhm. Kdysi jsem tu taky studoval. To ještě existovala Ndr. A teď. To učím anatomii. Aktuálně pobývám v Londýně na Usil, kde dělám Phd v neurovědách. Zaměřuje se také na psychologii, biologii, sociologii a nové technologie. Bude tady s vámi mít následující týden a půl přednášky na centrální nervový systém. Zaskakuji za profesora Emila Bartošku. Ale nebojte se, hned po mě přijde Lucie Bohdalová. S ní budete probírat dýchací systém. Močové a pohlavní ústrojí. Srdce a lymfatický systém také společně proberete Aids.

Když se snažíme pochopit systémy v těle je dobré si vytvořit metaforu. Minule jsme mluvili o gastrointestinálním systému. Tam máme trubici, která rozmělňuje jídlo a přetavuje ho na molekuly, které vstřebává mozek je druhý není to žádná lehká hydrolíza Ester za dob ssr jsme o tom tolik nevěděli ale teď v době ej si k ní mozek přirovnáváme. Myslíme, že sbírá spoustu dat a na jejich základě predikuje s tímto modelem. Konkrétně přišel Jaromír Doležal. Podle všeho používáme jen tak asi jednu desetinu kapacity našeho mozku, že jsme vůbec pochopili, jak mozek funguje, nám umožnila taková nenápadná nudná obyčejná věc, histologické barvení, díky kterému jsme byli schopní. Nabarvit mozek a vidět neurony. Těch je určitě víc než tisíc šest set sedmdesát šest. Jak jsme se kdysi domnívali.

A díky tomuto speciálnímu barvení, které vymysleli Antoni Goldy a Lukáš Kachal, jsme viděli, že to nejsou jen tak ledajaké buňky, ale že mezi sebou mají nějaké spoje. Někde mezi tisíci pěti sty, čtyřiceti pěti a dvaceti tisíci spoji. Dále tuto metodu pak rozvinul Juan Weiss Horgez univerzity v Madridu, ve spolupráci s čínským vědcem jménem Youshin, který působí na univerzitě v hangrou když přeskočím, tak dvě, stě let tak jsme se dostali až k obtogenetice. Slyšeli jste o tom někdy?

Pracujeme s, geneticky upravenými myšmy, kterým dáme nějaké geny. Takže jejich proteiny, jejich transmitery reagují na světlo, když na něj pak posvítíme světlem o určité vlnové délce kolem pěti set padesáti nanometrů. Tak sedím, ten Euron aktivuje a my jsme tím schopní ovládat

populace neuronů, které exprimují. Ten daný gen. Mimochodem, jeden nanometr je deset na mínus devátou metrů. Můžeme to použít i u lidí. Dámy, jim elektrický Eeg senzor, který snímá jejich mozkovou aktivitu. Dělá se to v rámci projektu. Jeden projekt, který vede Honza Šedivý ředitel Institutu Jan Fonden ve městě Ssterreson. Lidé tak mohou ovládat třeba robotickou paži. Můžete se podívat na videa nám Bbc případně na americké televizi Mtv. Tato technologie se dá uplatnit taky. I třeba u Fbi v Interpolu nebo Nasa Unicet má vlastní projekt u Jana saláma, na který bylo vyhrazeno kolem pěti celých šesti set padesáti šesti miliard Euro z toho dvacet pět celých osm milionů čerpáme my v rámci tuzemského projektu a další jedna celá sedmdesát pět milionů je věnováno na podporu výzkumu. A dále čekáme také na Grand z Unesco.

Z čeho se ale mozek skládá? Že hmotě jsou děla neuron. A protože jsou u sebe těsně, tak to vypadá tmavší. Mozek má asi deset na jedenáctou neuronů a kolem devatenácti tisíc osmi set padesáti sedmi syna psi sedmdesát osm celých pět procenta tvoří voda, ne dvacet procent, jak se kdysi říkalo v bílé hmotě jsou axony a ty axony, když mají kombinace myjeli tak to vypadá světleji mozek se řekne cerebrum. Mícha se řekne medula spinalis. Mozek. Pak dělíme ná další části. Důležitá je medula oblongáta. Říká se jí také prodloužená mícha. Pak tam máme Ponce. 0 0 0 Chybí. Pak jen ce rebelům. Jen v něm je šedesát devět celých tři miliard neuronů. Všimněte si, že máme dvě slova: cerebrum, což je latinsky a encefallon. To je řecky. Tady na obrázku vidíme zgitální pohled. Díky tomu můžeme vidět část neokortexu a uvnitř hukley bílou hmotu tvoří dráhy, které vedou přes axony až do dalšího neuronu podle jejich směru je také pojmenováváme například dráha z kortexu do míchy by byla traktus. Corrtico spinalis. To se píše jako Tra. C T, Us. Mezera, čort i. C, o, sppina El is.

A z čeho vůbec vzniká Cns z vnějšího ektodermu vznikne neurální brázda a ta se zanoří, až vytvoří trubici. Část z toho se oddělí a vznikne z toho neurální lišta. Další důležité struktury jsou alární ploténka, která přijímá signál, tedy je Soma to senzitivní a bazální ploténka která je motorická pojďme se teď podívat na různá onemocnění mozku.

Budeme postupně v rámci hodin mluvit o Alzheimechově nemoci. Parkinsonově nemoci. Epilepsii a roztroušené skleróze. Alzheimer. Pardon. Alzheimerova nemoc je spojená s demencí. Víme, že v roce dva tisíce deset. Bylo na světě třicet pět milionů lidí trpících demencí. Za dalších dvacet let se ten počet zdvojnásobí. To znamená, že v roce dva tisíce třicet mluvíme asi o

šedesáti pěti milionech lidí s demencí. Dva tisíce padesát už asi sto třináct milionů dvě třetiny z nich ve vyspělých zemích. Jen dnes na světě žije přesně osoba miliard sto tři milionů. Devět set dvacet šest tisíc. Čtyři sta čtyřicet jedna lidí, ale čísla rostou a odhaduje se, že tou dobou bude na planetě žít asi devět celých sedm miliard lidí, takže daný počet se nezdá jako příliš velký ale stále se jedná o vysoké procento navíc poplatky za léčbu také neustále rostou v roce dva tisíce devatenáct zaplatili pojišťovny průměrně za pacienta. Padesát čtyři tisíc pět set dvacet korun. Celkem za celý rok se tak částka vyšplhala na jednu miliardu. Čtyři sta osmdesát milionů. Sto šestnáct tisíc korun mě nejčastější příčinou demence je právě Alzheimerova nemoc. Konkrétně v sedmdesáti pěti procentech všech případů demence.

Co to ale je. Alzheimerova nemoc. Je to mimo jiné porucha krátkodobé paměti. Ztráta paměti je zcela jistě nepatří přirozeně ke stáří zároveň touto nemocí. Netrpí pouze staří lidé. Může ji mít i třicátník. Mnozí se také domnívají, že nemoc není smrtelná onemocnění ničí mozkové buňky. Mozek degeneruje. Mění se hlavně nervové přen ašeče v mozku pro ukládání paměťové stopy. Konkrétně acetylcholinu a glutamátu i jejich základě jsme vytvořili přípravky. Takzvaná 0, která slouží k léčbě pacientů. Jednak inhibitory a cedlcholinsterázy. Jednak antagonisté, Nmda receptorů dokáží pozastavit průběh nemoci na zhruba jeden až pět let. Pomocí biomarkeů pozorujeme i další změny. Například Beta. A melot, který se ukládá v seniorních placích.

Nemoc. Dostala jméno podle, Aloise Alzheimera, který 0 poprvé popsal u pacientky Ogus Dieter v roce Devatenáct Set jedna devatenáct set sedm pak popsal neuropatologický obraz ale teprve devatenáct set deset a ty Kuzamakám v Mnichově zmiňuje Alzheimerovou nemoc. Petr přitom trpěla vzácnou variantou. Obvykle se nemoc projevuje po šedesáti, ale ona ji měla geneticky mnohem dříve. To je u dvou až tří 0 pacientů. Četnost výs, četnost výskytu také roste s věkem kolem šedesáti pěti let je postiženo jedno až pět procent lidí ve věku devadesát let až polovina populace u nás žije sedm set dvacet čtyři lidí kterým je nad sto let, to znamená že minimálně tři sta šedesát dva z nich trpí touto nemocí.

O tom se mluví méně, ale je neméně důležitý. Je Oskar? Fišer. Tudíž ona by se ta nemoc měla jmenovat. Alzheimer. Fišerova nemoc, protože on už tou dobou v roce. Osmnáct set devadesát sedm popsal kolem patnácti pacientů s touto nemocí bohužel Marie Stejskalová bohumila

skalová ani Karel bouček nebyli jedni z nich protože nebyly, z prahy ani z brna pocházely z malé vesnice Plešovice, u které se poblíž nachází Dalešická přehrada.

Pokud vás zajímá více. Ozvěte se na mail. Pošlete mi Cz v aplikaci doin nebo na číslo plus čtyři, dva, nula, šest, nula, dva, osm, sedm, pět, čtyři, osm, tři, případně na číslového španělského kolegy z kordoby, plus třicet čtyři, šest, šest, dva, nula, nula, pět, sedm, tři, šest.

Příloha č. 8: Text automatických titulků – angličtina

Hello everyone. My name is Keeper Schneider and you spelled that KIEF ER. SCH NEID ER. I'm German, but I studied here when the NDR still existed and now I'm back in teaching anatomy. I live in London and I'm a PhD student at UCL in neuroscience. But I also dabble in psychology, biology, sociology, and new technology. I'll be teaching you for a week and a half about the central nervous system. I'm subbing in for Professor John Fowler. When I session is over, you'll have Lucy Morgan and with her you'll be looking at the respiratory urinary and reproductive systems. As well as the heart. And I apologize as well as the heart and lymphatic system. You'll also be discussing AIDS.

Now, it's good to use metaphors to try to understand the human body. It looks like you've been talking about the gastrointestinal system. With that, basically there's a tube that takes food and makes it into molecules that it then absorbs. The brain works a bit differently. It makes us or hydrolysis look pretty easy in comparison to be honest. Now back when the USSR still existed. We didn't know much about the brain. Now, in the era of AI, we can draw parallels between it and the brain. Essentially, it collects tons of data and based on that, it makes predictions. Henry Fletcher is the one who came up with that specific model. Oh, and it seems we only is about 1 10th of our brain capacity.

Now, the fact that we even understand how the brain works was made possible thanks to a single boring mundane thing. Histological staining. Because of this, we can mark the brain and actually see neurons. As opposed to what we'd, as opposed to, as opposed to what we'd previously believed. There's certainly more than 1,676 of them. And thanks to this special staining, which was created by Anthony Goldie and Lukashko. We saw that there that they are no ordinary cells, but that they're connected through somewhere between 1,545 and 20,000 connections. This

method was then further developed by Juan Louise Jorge from the University of Madrid in cooperation with the Chinese scientist Yu Shin who works out of a university in Hangzhou. Now, let me skip about 200 years so we can get to Optogenetics. Here you guys ever heard about it?

Well, we work with genetically modified mice. We give them genes so that their proteins, their transmitters react to light. Then if we shine a light of a certain wavelength, We're on 550 nanometers. 0 Them it'll activate the neuron and thus we're able to control the neuronal population. Expresses that specific gene. Oh, by the way, one nanometer is equal to 10. To the negative 9th power of a meter. We can also do this with people. We give them an electric EEG sensor that records their brain activity. This is what is being done at the Hannah project led by Judith Fairfield, the director of the Institute, Hyung Fonda in the city of Istaland. Using this technique, people can control a robotic arm. You can watch some videos on BBC or MTV. The technology may also be used by the FBI, Interpol, and NASA. UNICEF is the funding of UNICEF is funding its own project called U Yanis Lama with 5.6 5 6 billion euro. Out of that, we drew 24, we drew 25.8 million for our local project and another 1.7 5 million has been earmarked for research support. We're also winning 0 Grant, from UNESCO.

But what exactly constitutes the brain? They're bodies of neurons in the gray matter. And because they're so close to each other, it appears darker. The brain has around 10 to the 11th power of neurons and around 19,587 synapses. Plus. It's around 78.5% water. Not the 20% we were told back in the day. The white matter is formed by axons and these axons because they are surrounded by myelin appear lighter. The brain is called the cerebrum. The spinal cord is the medulla spinalis. We can further divide the brain into other parts. An important one is it is the medulla oblangata, also called the bulb. Up next is the ponds of Mesencevalon. D and Telencephalon. The last part of the cerebellum, it alone contains 69.3 billion neurons. It might be interesting to note that of these 2 words for the brain, cerebral comes from Latin, wellencephalon is Greek. In this picture, we can see the sagittal view that shows us a part of the neocortex and nuclei. The white matter is constituted of pathways that go through the accents all the way to the next neuron. If we observe their direction, we can name them. For example, the pathway from

the cortex to the spinal cord would be tractus corticospinalis. And you spell that TRA CT US. CORT ICOS PINA LIS.

So how is the CNS created? A neural groove emerges from the external ectoderm and it deepens to form a tube. A part of that detaches to form the neural crest. Other important structures include the aller plate, which receive signals and as such is, is someatosensory. And the basal plate, which is motor.

Now, let's look at various brain diseases. We're gonna talk about Alzheimer's disease, Parkinson's disease, epilepsy, and multiple sclerosis, respectively. Alzheimer's is connected to dementia. We know that in 2,010 . There were 35 million people in the world suffering from dementia. And the next 20 years that number should double, meaning that by 2,030 we're talking about 65 million people with dementia. By 2,050 about 130 and by 2,050 about a hundred 13,000,002 thirds of those living in developed countries. They're exactly 8,103,926,441 people living in the world today. But these numbers are growing and it's estimated that there are going to be 9.7 billion people by 2,050. The numbers might not sound so big then. But it's still a large percentage. Additionally, Treatment fees have also been steadily arising. In 2,019 insurance companies paid an average of 54,520 crowns per patient. Which made a yearly total of 1,480,116,000 fr. Allzheimer's disease is the most common cause of dementia. Specifically in 75% of all dimension cases.

But what exactly is Alzheimer's disease? It is, among other things, a short-term memory disorder. Memory loss does not come as a natural part of growing old. Also, it's not just all people who suffer from this disease. Younger people can have it too. Many also believe that the disease is not terminal. The disease destroys brain cells. The brain degenerates. These changes mostly affect the neurotransmitters in the brain in charge of storing memory traces. Namely, acetylcholine and glutamate. Based on these, we have developed medicines called cognition-enhancing drugs to treat patients to boast that though that I apologize. Through both acetal, That's that I. Settle collinar stories inhibitors and NMDA receptor antagonists. They can halt the course of the disease for about one to 5 years. Using biomarkers, we can observe other changes such as beta amyloid, which is stored in C.

Now, the disease was named after Al Aloyis Alzheimer, who 1st described it in his patient August, August, that, in 19 0 1. He then described the neuropathological picture in 19 0 7, but it was only in 1,910 that Alzheimer's disease was mentioned by 0 Khan in Munich. In fact, Dieter suffered from a rare variant. The disease usually manifests after the age of 60, but she had it much earlier data genetic causes, which happens in 2 to 3% of patients. The incidents does that the incidents does increase with age with one to 5% of people affected around 65 and up to one half of the population at the age of 90. In our country there are 724 people who are over 100 years old which means that at least 362 of them suffer from this disease.

Now, there's 1 other person. Who is less talked about but is no less important. Octavian Los Angeles. Actually, the disease should be called Alzheimer's, Alzheimer's disease, because around that time at 1897, she had already described around 15 patients with the disease. Unfortunately, these didn't include Mary Ackland, Mitch Banner, and Sandy Duncan, because they weren't from San Francisco or Chicago. They came from a small village. Called Lebanon near Norwich.

Now, if you're interested in learning more, please contact me by email, on 0, or call me at Country code 4 2 0 6 0 2 8 7 5 4 8 3 or my Spanish colleague from Cordoba. Country code (346) 620-5736.

Příloha č. 9: Text automatických titulků – španělština

Buenos días me llamo kefer schneider es un hombre alemán y se escribe k sc n i d e r estudié aquí cuando la rda todavía existía y ahora enseñó anatomía ahora mismo vivo en barcelona estudio en la uve y hago un pHD en neurociencias también me dedico a la psicología la biología la sociología y las nuevas tecnologías durante la próxima semana y media les daré clases sobre el sistema nervioso central sustituyo al profesor felipe manceñido pero no se preocupen justo después de mí vendrá patricia clares naveros con ella aprenderán sobre el sistema respiratorio el sistema urinario y reproductor el corazón y el sistema linfático juntos también hablarán del sida

Cuando tratamos de comprender los sistemas corporales es útil recurrir a las metáforas la última vez hablaron del sistema gastrointestinal que funciona como un tubo que desmenuza los alimentos y los convierte en moléculas que luego orden el cerebro es un poco diferente no se trata de una sencilla hidrólisis de esteres durante la existencia de la urss sabíamos poco de él pero

ahora en la era de la ya lo comparamos con ella consideramos que recoge montones de datos y hace predicciones basándose en ellos fue franco díaz martín quien creó este modelo parece que solo usamos un décimo de la capacidad de nuestro cerebro

Fue posible llegar a entender cómo funciona el cerebro gracias a una cosa común aburrida y poco llamativa la tinción histológica gracias a ella logramos teñir el cerebro y ver las neuronas por cierto hay más de 1 676 a diferencia de lo que se pensaba y gracias a esta atención especial creada por anthony golddy y lucas kahn vimos que no son células comunes sino que están conectadas con un número de conexiones comprendido entre 1 545 y 20 Más adelante juan luis jorge desarrolló el método en la universidad de madrid en colaboración con un científico chino llamado you shin que trabajaba en la universidad de Hangzhou 0 se avanzamos unos 200 años llegamos a la optogenética saben qué es

trabajamos con ratones modificados genéticamente les damos genes para que sus proteínas sus transmisores reaccionen a la luz si les iluminamos con una luz de una cierta longitud de onda de unos 550 anómetros se activará la neurona y así seremos capaces de controlar la población de neuronas que expresa el gendado a propósito un nanómetro equivale a 10 elevado a menos 9 de un metro también podemos utilizar eso en humanos les damos un sensor eje eléctrico que registra su actividad cerebral lo hacen dentro del giorne project dirigido por alberto jiménez lópez director del instituto jar fonden de la ciudad de osterund de esta manera las personas pueden controlar un brazo robótico pueden ver vídeos al respecto en las cadenas bbc o mtv También utilizan esta tecnología el FBI la interpol y la nasa unicef tiene su propio proyecto huyano salama al que se asignaron 5 656 000 000 de euros de esa suma cobramos 25,8 000 000 para nuestro proyecto local y 1 coma 75 000 000 se dedican a apoyar la investigación también esperamos una subvención de la unesco 0

Pero qué constituye el cerebro la sustancia gris se compone de cuerpos de neuronas y como están tan cerca las unas de las otras parece más oscura el cerebro tiene unas 10 elevado a 11 neuronas y unas 19 587 sinapsis ese 78,5 0 es agua no el 20 0 como se afirmaba antes la sustancia blanca está formada por axones y estos axones dado que están rodeados de mielina parecen más claros el cerebro se denomina cerebro la médula espinal es médula espinalis a continuación dividimos el cerebro en otras partes una importante es la médula oblongata también llamada bulbo raquídio

luego están el pons el mesencefalón el dien cefalón y el telencefalón el último es el cerebelum que contiene 69 300 000 000 de neuronas fíjense en que hay 2 palabras cerebrrum que viene del latín y en cefalon que es griego en esta imagen podemos ver la vista sagital que permite observar partes del neocórtex y los nucley la sustancia blanca está compuesta por vías que van por los axones hasta la neurona siguiente podemos nombrarlas según su dirección por ejemplo la vía que va del córtex a la médula espinal sería tractus corticoespinalis lo que se deletrea trcactvs c o r t i c o sp i n a l i s

Ahora bien cómo nace el snc empieza cuando un surco neural emerge del ectodermo externo y se profundiza conformándose así un tubo parte de él se desprende para formar la cresta neural entre otras estructuras importantes figuran la placa lar que recibe señales y como tal es somatosensorial y la placa basal que es motora

Veamos ahora las distintas enfermedades del cerebro en clase vamos a hablar de la enfermedad de alzheimer la enfermedad de parkinson la epilepsia y la esclerosis múltiples sucesivamente La enfermedad de alzheimer está relacionada con la demencia en 2 010 había 35 000 000 de personas en el mundo que padecían demencia en los 20 años siguientes esa cifra se duplicará lo que significa que en 2 030 estaremos hablando de unos 65 000 000 de personas con demencia y en 2 050 de unos 113 000 002 tercios de ellas situadas en países desarrollados hoy en día viven exactamente 8 103 926 441 personas en el mundo pero las cifras no dejan de aumentar y se estima que habrá 9,7 000 000 de personas para el año 2 050 Si bien el número de personas afectadas por la enfermedad no parece enorme en comparación con esta cifra es un porcentaje importante además las tasas de tratamiento también están aumentando sin parar en el año 2 019 las compañías de seguros pagaron un promedio de 54,5 coronas por un paciente por lo que el gasto total anual ha aumentado a 1 400 116 000 coronas la enfermedad de alzheimer es la causa más común de demencia concretamente representa el 75 0 de todos los casos de demencia

Pero qué es la enfermedad de alzheimer entre otras cosas se trata de un trastorno de la memoria a corto plazo la pérdida de memoria no es en ningún caso una parte natural del envejecimiento al mismo tiempo no son solo las personas mayores las que padecen esta enfermedad también afecta a personas jóvenes además muchos creen que la enfermedad no es mortal la enfermedad destruye las células cerebrales y el cerebro se degenera los cambios afectan sobre todo a los

neurotransmisores cerebrales encargados de almacenar las huellas de la memoria en concreto la acetilcolina y el glutamato basándonos en ellos hemos desarrollado medicamentos llamados potenciadores cognitivos para tratar a los pacientes mediante inhibidores de la acetilcolina de esterasa y antagonistas de los receptores nma Estos medicamentos logran detener el desarrollo de la enfermedad entre 1 y 5 años Gracias a los biomarcadores podemos observar otros cambios como el betamiloide que se almacena en las placas seniles

La enfermedad debe su nombre a 0 alzheimer que la describió por primera vez en una paciente august peter en el año 1 901 luego describió el cuadro neuropatológico en 1 907 pero no fue hasta 1 910 cuando attic zamakan mencionó la enfermedad en múnich de hecho dicha parecía una variante poco frecuente la enfermedad suele manifestarse después de cumplir los 60 años pero ella la desarrolló mucho antes por causas genéticas lo que ocurre en un 2 o 3 0 de los casos La incidencia también aumenta con la edad con un entre 1 y 5 de afectados en torno a los 65 años y hasta la mitad de la población a los 90 años en nuestro país hay 724 personas mayores de 100 años lo que significa que al menos 362 de ellas padecen esta enfermedad

hay una persona de la que se habla menos pero que no por ellos menos importante se llama oscar fisher en realidad la enfermedad debería llamarse enfermedad de alzheimer fisher porque en 1 897 ya había descrito a unos 15 pacientes con esta enfermedad Por desgracia juan santiago muñiz francisco javier ruiz y lucía dolores bustamante no formaron parte de ellos porque no eran de salamanca o bilbao sino que vivieron en un pequeño pueblo de sorihuela del guadalimar cerca de jaén si les interesa saber más contáctenme por correo electrónico en douin o llámenme al 0 34 6 6 2 0 5 7 36 o a mi colega checo de horno al 0 0 4 2 6 0 2 8 7 5 4 8 3

Příloha č. 10: Kategorie: čeština – jména, názvy a zkratky

Legenda: modrá = malá chyba, žlutá = standardní chyba, červená = vážná chyba

Domácí jména a názvy

Původní verze	**Automatické titulky**
Londýně	Londýně
Emila Bartošku	Emila Bartošku
Lucie Bohdalová	Lucie Bohdalová
Jaromír Doležal	Jaromír Doležal
Lukáš Kachal	Lukáš Kachal
Madridu	Madridu
Honza Šedivý	Honza Šedivý
Mnichově	Mnichově
Marie Stejskalová	Marie Stejskalová
Bohumila Lískalová	bohumila skalová
Karel Bouček	Karel bouček
Prahy	prahy
Brna	brna
Plešovice	Plešovice
Dalešická přehrada	Dalešická přehrada

Cizí jména a názvy

Původní verze	**Automatické titulky**
Kiefer Schneider	Kefr. Schneider
Anthony Goldie	Antoni Goldy
Juan Luis Jorge	Juan Weiss Horgez
Yu Xin	Youshin
Hangzhou	hangrou
HjerneProjekt	Jeden projekt
Hjärnfonden	Jan Fonden
Östersund	Ssteresson
Ujana Salama	u Jana saláma
Aloise Alzheimera	Aloise Alzheimera
Auguset Dieter	Ogus Dieter
Dieter	Petr
Atiquzzama Khan	a ty Kuzamakám

Oscar Fischer	Oskar? Fišer
Alzheimer-Fischerova	Alzheimer. Fišerova
Douyin	Doin
Córdoby	kordoby

Domácí zkratky a vyhláskování

Původní verze	Automatické titulky
NDR	Ndr.
SSSR	ssr
EEG	Eeg
CNS	Cns
NMDA	Nmda
SZ	Cz
k-i-e-f-e-r s-ch-n-e-i-d-e-r	kádí Efers. Chm. Mhm.
t-r-a-c-t-u-s /mezera/ c-o-r-t-i-c-o-s-p-i-n-a-l-i-s	Tra. C T, Us. Mezera, čort i. C, o, sppina, El is.

Cizí zkratky

Původní verze	Automatické titulky
UCL	Usil
PHD	Phd
AI	ej
BBC	Bbc
MTV	Mtv
FBI	Fbi

Příloha č. 11: Kategorie: angličtina – jména, názvy a zkratky

Legenda: modrá = malá chyba, žlutá = standardní chyba, červená = vážná chyba

Domácí jména a názvy

Původní verze	Automatické titulky
London	London
John Fowler	John Fowler
Lucy Morgan	Lucy Morgan
Henry Fletcher	Henry Fletcher
Anthony Goldie	Anthony Goldie
Madrid	Madrid
Judith Fairfield	Judith Fairfield
Munich	Munich
Mary Ackland	Mary Ackland
Mitch Banner	Mitch Banner
Sandy Duncan	Sandy Duncan
San Francisco	San Francisco
Chicago	Chicago
Lavenham	Lebanon
Norwich	Norwich

Cizí jména a názvy

Původní verze	Automatické titulky
Kiefer Schneider	Keeper Schneider
Lukasz Kachal	Lukashko
Juan Luis Jorge	Juan Luis Jorge
Yu Xin	Yu Shin
Hangzhou	Hangzhou
HjerneProjekt	Hannah project
Hjärnfonden	Hyung Fonda
Östersund	Istaland
Ujana Salama	U Yanis Lama
Alois Alzheimer	Al Aloyis Alzheimer
Auguset Dieter	August, that
Dieter	Dieter

Atiquzzama Khan	- Khan
Octavia Velazquéz	Octavian Los Angeles
Alzheimer-Velasquéz‘	Alzheimer’s, Alzheimers‘s
Douyin	-
Córdoba	Cordoba

Zkratky

Původní verze	Automatické titulky
NDR	NDR
PhD	PhD
UCL	UCL
USSR	USSR
AI	AI
EEG	EEG
BBC	BBC
MTV	MTV
FBI	FBI
CNS	CNS
NMDA	NMDA
k-i-e-f-e-r s-ch-n-e-i-d-e-r	KIEF ER. SCH NEID ER.
t-r-a-c-t-u-s c-o-r-t-i-c-o-s-p-i-n-a-l-i-s	TRA CT US. CORT ICOS PINA LIS.

Příloha č. 12: Kategorie: španělština – jména, názvy a zkratky

Legenda: modrá = malá chyba, žlutá = standardní chyba, červená = vážná chyba

Domácí jména a názvy

Původní verze	Automatické titulky
Barcelona	barcelona
Felipe Manceñido	felipe manceñido
Patricia Clares Naveros	patricia clares naveros
Franco Díaz Martín	franco díaz martín
Juan Luis Jorge	juan luis jorge
Madrid	madrid
Alberto Jiménez Lopez	alberto jiménez lópez
Múnich	múnich
Juan Santiago Muñez	juan santiago muñiz
Francisco Javier Ruíz	francisco javier ruiz
Lucía Dolores Bustamante	lucía dolores bustamante
Salamanca	salamanca
Bilbao	bilbao
Sorihuela del Guadalimar	sorihuela del guadalimar
Jaén	jaén

Cizí jména a názvy

Původní verze	Automatické titulky
Kiefer Schneider	kefer schneider
Anthony Goldie	anthony golddy
Lukasz Kachal	lucas kahn
Yu Xin	you shin
Hangzhou	hangzhou
HjerneProjekt	giorne project
Hjärnfonden	jar fonden
Östersund	osterund
Ujana Salama	huyano salama
Alois Alzheimer	- alzheimer
Auguset Dieter	august peter
Dieter	dicha

Atiquzzama Khan	attic zamakan
Oscar Fischer	oscar fisher
Alzheimer-Fischer	alzheimer fischer
Douyin	douin
Brno	horno

Zkratky

Původní verze	Automatické titulky
RDA	rda
PhD	pHD
UB	uve
URSS	urss
IA	ya
EEG	eje
BBC	bbc
MTV	mtv
FBI	FBI
SNC	snc
NMDA	nma
k-i-e-f-e-r s-ch-n-e-i-d-e-r	k sc n i d e r
t-r-a-c-t-u-s c-o-r-t-i-c-o-s-p-i-n-a-l-i-s	trcactvs c o r t i c o sp i n a l i s

Příloha č. 13: Kategorie: všechny jazyky – akronymy, čísla, termíny a cizí termíny, výčty a negace

Legenda: modrá = malá chyba, žlutá = standardní chyba, červená = vážná chyba

Akronymy

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
AIDS	Aids	AIDS	sida
Interpolu	Interpolu	Interpol	interpol
NASA	Nasa	NASA	nasa
UNICEF	Unicet	UNICEF	unicef
UNESCO	Unesco	UNESCO	unesco

Čísla

Jednotky

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
1 nanometr	jeden nanometr	one nanometer	un nanómetro
2 slova	dvě slova	2 words	2 palabras
2-3 % pacientů	dvou až tří pacientů	2 to 3% of patients	un 2 o 3 de
1-5 % lidí	jedno až pět procent lidí	one to 5% of people	entre 1 y 5 de afectados
1 až 5 let	jeden až pět let	one to 5 years	entre 1 y 5 años

Desítky

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
20 let	dvacet let	20 years	20 años
60	šedesáti	60	60 años
65 let	šedesáti pěti let	65	65 años
90 let	devadesát let	90	90 años
15 pacientů	patnácti pacientů	15 patients	15 pacientes
20%	dvacet procent	20%	20
75 % všech případů	sedmdesáti pěti procentech všech případů	75% of all dimension cases	75 de todos los casos
78.50%	sedmdesát osm celých pět procenta	78.5% water	78,5 es agua

Stovky

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
200 let	dvě, stě let	200 years	200 años
550 nanometrů	pěti set padesáti nanometrů	550 nanometers	550 anómetros
724 lidí	sedm set dvacet čtyři lidí	724 people	724 personas
100 let	sto let	100 years	100 años
362 z nich	tři sta šedesát dva z nich	362 of them	362 de ellas

Tisíce

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
1,676	tisíc šest set sedmdesát šest	1,676 of them	1 676
1,545	tisíci pěti sty, čtyřiceti pěti	1.545	1 545
20 000 spoji	dvaceti tisíci spoji	20,000 connections	20
19 857 synapsí	devatenácti tisíc osmi set padesáti sedmi syna psí	19,587 synapses	19 587 sinapsis
54 520 korun	padesát čtyři tisíc pět set dvacet korun	54,520 crowns	54,5 coronas

Miliony

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
35 milionů lidí	třicet pět milionů lidí	35 million people	35 000 000 de personas
65 milionech lidí	šedesáti pěti milionech lidí	65 million people	65 000 000 de personas
25,8 milionů	dvacet pět celých osm milionů	25.8 million	25,8 000 000 para
1,75 milionů	jedna celá sedmdesát pět milionů	1.7 5 million	1 coma 75 000 000

113 milionů	sto třináct milionů	130 and by 2,050 about a hundred 13,000,002 thirds !!	113 000 002 tercios

Miliardy

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
9,7 miliard	devět celých sedm miliard lidí	9.7 billion people	9,7 000 000 de personas
5,656 miliard euro	pěti celých šesti set padesáti šesti miliard Euro	5.6 5 6 billion euro	5 656 000 000 de euros
69,3 miliard neuronů	šedesát devět celých tři miliard neuronů	69.3 billion neurons	69 300 000 000 de neuronas
8 103 926 441 lidí	osoba miliard sto tři milionů. Devět set dvacet šest tisíc. Čtyři sta čtyřicet jedna lidí	8,103,926,441 people	8 103 926 441 personas
1 480 116 000 korun	jednu miliardu. Čtyři sta osmdesát milionů. Sto šestnáct tisíc korun	1,480,116,000 fr	1 400 116 000 coronas

Roky

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
2010	dva tisíce deset	2,010.	2 010
2030	dva tisíce třicet	2,030.	2 030
2050	Dva tisíce padesát	2,050.	2 050
2019	dva tisíce devatenáct	2.019	2 019
1901	Devatenáct Set jedna	19 0 1	1 901
1907	devatenáct set sedm	19 0 7	1 907
1910	devatenáct set deset	1,910.	1 910
1897	Osmnáct set devadesát sedm	1897	1 897

Speciální

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
1/10 kapacity	jednu desetinu kapacity	1 10th of our brain capacity	un décimo de la capacidad
10^{-9} metrů	deset na mínus devátou metrů	10. To the negative 9th power of a meter	10 elevado a menos 9 de un metro
10^{11} neuronů	deset na jedenáctou neuronů	10 to the 11th power of neurons	10 elevado a 11 neuronas
2/3 z nich	dvě třetiny z nich	hundred 13,000,002 thirds !!	113 000 002 tercios de ellas
½ populace	polovina populace	one half of the population	la mitad de la población
.+420 602 875 483	plus čtyři, dva, nula, šest, nula, dva, osm, sedm, pět, čtyři, osm, tři	Country code 4 2 0 6 0 2 8 7 5 4 8 3	0 0 4 2 6 0 2 8 7 5 4 8 3
.+ 34 662 00 57 36	plus třicet čtyři, šest, šest, dva, nula, nula, pět, sedm, tři, šest	Country code (346) 620-5736	0 34 6 6 2 0 5 7 36

Termíny

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
Centrální nervový systém	Centrální nervový systém	Central nervous system	sistema nervioso central
Dýchací systém	Dýchací systém.	Respiratory	sistema respiratorio
Močové a pohlavní ústrojí	Močové a pohlavní ústrojí.	Urinary and reproductive systems	sistema urinario y reproductor
Lymfatický systém	Lymfatický systém	Lymphatic system	sistema linfático
Gastrointestinálním systému	Gastrointestinálním systému	Gastrointestinal system	sistema gastrointestinal
Molekuly	Molekuly	Molecules	moléculas
Hydrolýza esterů	hydrolíza Ester	Us or hydrolysis	hidrólisis de esteres
Histologické barvení	Histologické barvení	Histological staining	tinción histológica

Neuron, neuronu, Neurony, Neuronů (3x)	neurony, neuronů (3x), neuronu, neuron	Neuron (2x), neurons (4x)	neurona (2x), neuronas (4x)
Buňky (2x)	Buňky (2x)	Cells (2x)	células (2x)
Optogenetice	Obtogenetice	Optogenetics	optogenética
Geny, gen	Geny, gen	genes, gene	gendado, genes
Proteiny	Proteiny	Proteins	proteínas
Transmitery	Transmitery	Transmitters	transmisores
Vlnové délce	Vlnové délce	Wavelength	longitud de onda
Populace neuronů	Populace neuronů	Neuronal population	población de neuronas
Exprimují	Exprimují	Expresses	expresa
Elektrický EEG senzor	Elektrický Eeg senzor	Electric EEG sensor	sensor eje eléctrico
Mozkovou aktivitu	Mozkovou aktivitu	Brain activity	actividad cerebral
Šedé hmotě	Že hmotě	Gray matter	sustancia gris
Synapsí	Syna psí	synapses	sinapsis
Bílé hmotě, bílou hmotu	Bilé hmotě, Bílou hmotu	White matter (2x)	sustancia blanca (2x)
Axony (3x)	Axony (3x)	Axons (2x), accents (1x)	axones (3x)
myelin	Myjeli	Myelin	mielina
inhibitory acetylcholinesterázy	Inhibitory a cedlcholinsterázy	Acetal, That's that I. Settle collinear stories inhibitors	inhibidores de la acetilcolina de esterasa
antagonisté NMDA receptorů	Antagonisté Nmda receptorů	NMDA receptor antagonists	antagonistas de los receptores nma
Senilních placích	Seniorních placích	C.	placas seniles
Kortexu	kortexu	cortex	córtex
Mícha, míchy	mícha	Spinal cord	médula espinal
Prodloužená mícha	Prodloužená mícha	Bulb	bulbo raquídio
Sagitální pohled	Zgitální pohled	Sagittal view	vista sagital
Neokortexu	neokortexu	Neocortex	neocórtex
Ektodermu	Ektodermu	ectoderm	ectodermo
Neurální brázda	Neurální brázda	Neural groove	surco neural
Neurální lišta	Neurální lišta	Neural crest	cresta neural
Alární ploténka	Alární ploténka	Aller plate	placa lar

Somato-senzitivní	Soma to senzitivní	Someatosenasory	somatosensorial
Bazální ploténka	Bazální ploténka	Basal plate	placa basal
Motorická	Motorická	Motor	motora
Alzheimerově nemoci, Alzheimerova nemoc (3x),	Alzheimechově nemoci, Alzheimerova nemoc (3x)	Alzheimer's disease (3x), Allzheimer's disease	enfermedad de alzheimer (4x)
Parkinsonově nemoci	Parkinsonově nemoci	Parkinson's disease	enfermedad de parkinson
Epilepsii	Epilepsii	Epilepsy	epilepsia
Roztroušené skleróze	Roztroušené skleróze	Multiple sclerosis	esclerosis múltiples
Demencí (3x), Demence (2x)	Demencí (3x), Demence (2x)	Dementia (4x), dimension	demencia (5x)
Porucha krátkodobé paměti	Porucha krátkodobé paměti	Short-term memory disorder	trastorno de la memoria a corto plazo
Ztráta paměti	Ztráta paměti	Memory loss	pérdida de memoria
Degeneruje	Degeneruje	degenerates	degenera
Nervové přenašeče	Nervové přenašeče	Neurotransmitters	neurotransmisores
Paměťové stopy	Paměťové stopy	Memory traces	huellas de la memoria
Acetylcholinu	Acetylcholinu	Acetylcholine	acetilcolina
glutamátu	Glutamátu	Glutamate	glutamato
kognitiva	-	Cognition-enhancing drugs	potenciadores cognitivos
biomarkerů	Biomarkeů	Biomarkers	biomarcadores
Beta-amyloid	Beta. A melot	Beta amyloid	betamiloide
Neuropatologický obraz	Neuropatologický obraz	Neuropathological picture	cuadro neuropatológico

Cizí termíny

Původní verze (CS)	**Automatické titulky (CS)**	**Automatické titulky (EN)**	**Automatické titulky (ES)**
Cerebrum (2x)	Cerebrum (2x)	Cerebrum, cerebral	Cerebro, cerebrrum
Medulla spinalis	Medula spinalis	Medulla spinalis	médula espinalis
Medulla oblongata	Medula oblongáta	Medulla oblangata	médula oblongata
Mesencephalon	-	Mesencevalon.	mesencefalón
Diencephalon	-	D	dien cefalón

telencephalon	-	Telencephalon	telencefalón
Cerebellum	Ce rebelům	Cerebellum	cerebelum
pons	Ponce	-	pons
Encephalon	encefallon	wellencephalon	en cefalon
Nuclei	nukley	Nuclei	nucley
Tractus corticospinalis	traktus. Corrtico spinalis	Tractus corticospinalis	tractus corticoespinalis

Výčty

Původní verze (CS)	**Automatické titulky (CS)**	**Automatické titulky (EN)**	**Automatické titulky (ES)**
[psychologii, biologii, sociologii a nové technologie]	na psychologii, biologii, sociologii a nové technologie	in psychology, biology, sociology, and new technology	a la psicología la biología la sociología y las nuevas tecnologías
[dýchací systém, močové a pohlavní ústrojí, srdce a lymfatický systém]	dýchací systém. Močové a pohlavní ústrojí. Srdce a lymfatický systém	respiratory urinary and reproductive systems. As well as the heart. And I apologize as well as the heart and lymphatic system	sobre el sistema respiratorio el sistema urinario y reproductor el corazón y el sistema linfático
[pons, mezencefalon, diencefalon a telencefalon]	Ponce. 0 0 0	the ponds of Mesencevalon. D and Telencephalon.	el pons el mesencefalón el dien cefalón y el telencefalón
[Alzheimerově nemoci, Parkinsonově nemoci, epilepsii a roztroušené skleróze]	Alzheimechově nemoci. Parkinsonově nemoci. Epilepsii a roztroušené skleróze.	Alzheimer's disease, Parkinson's disease, epilepsy, and multiple sclerosis	la enfermedad de alzheimer la enfermedad de parkinson la epilepsia y la esclerosis múltiples
[Marie Stejskalová, Bohumila Lískalová ani Karel Bouček]	Marie Stejskalová bohumila skalová ani Karel bouček	Mary Ackland, Mitch Banner, and Sandy Duncan	juan santiago muñiz francisco javier ruiz y lucía dolores bustamante

Negace

Původní verze (CS)	Automatické titulky (CS)	Automatické titulky (EN)	Automatické titulky (ES)
ne 20 %	ne dvacet procent	Not the 20%	no el 20
ztráta paměti zcela jistě nepatří přirozeně ke stáří	Ztráta paměti je zcela jistě nepatří přirozeně ke	does not come as a	no es en ningún caso una parte
touto nemocí netrpí pouze staří lidé	zároveň touto nemocí. Netrpí	it's not just all people	no son solo las personas mayores las que padecen
nemoc není smrtelná	že nemoc není smrtelná	the disease is not terminal	no es mortal
to nejsou jen tak ledajaké buňky	to nejsou jen tak ledajaké buňky	are no ordinary cells	no son células comunes
počet se nezdá jako příliš veliký	počet se nezdá jako příliš velký	not sound so big then	no parece enorme
ale je neméně důležitý	ale je neméně důležitý.	but is no less important	que no por ellos menos importante
nebyli jedni z nich	nebyli jedni z nich	these didn't include	no formaron parte
protože nebyli z	protože nebyly, z	because they weren't from	no eran de