

**Univerzita Karlova**

**Filozofická fakulta**

Ústav informačních studií a knihovnictví

# **Diplomová práce**

Bc. Štěpánka Kološová

## **Konverzační design: principy designu hlasového robota pro přirozenou komunikaci s lidmi**

Conversation Design: Principles of Voicebot Design for Natural  
Communication with People

Praha 2022

Vedoucí práce: Mgr. Josef Šlerka, Ph.D.

Prohlášení:

Prohlašuji, že jsem diplomovou 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 26. července 2022

Štěpánka Kološová



### **Poděkování**

Na tomto místě bych ráda poděkovala kolegům z VOCALLS za cenné připomínky k této diplomové práci.

**Klíčová slova (česky)**

voicebot, hlasový robot, virtuální asistent, konverzační design, konverzační umělá inteligence, konverzační analýza

**Klíčová slova (anglicky):**

voicebot, voice robot, virtual assistant, conversation design, conversational artificial intelligence, conversational analysis

### **Abstrakt (česky)**

Tato diplomová práce se zabývá otázkou, jak mohou hlasoví asistenti komunikovat s lidmi přirozeně. Zkoumá, jak komunikují lidé, jak lze analyzovat konverzaci a její principy použít pro návrh virtuálních asistentů. Cílem práce je vytvoření modelu pravidel konverzačního designu, které je potřeba dodržet při tvorbě voicebotů. Tento model vychází jednak z poznatků z odborné literatury, tak také z analýzy reálných konverzací lidí s roboty.

### **Abstract (in English):**

This thesis addresses the question of how voice assistants can communicate with people naturally. It explores how people communicate, how conversations can be analyzed and its principles can be used to design virtual assistants. The aim of the thesis is to create a model of conversational design rules to follow when creating voicebots. This model is based on both knowledge from the literature and analysis of real conversations between humans and bots.

# OBSAH

<b>1 Úvod</b>	<b>12</b>
<b>2 Teoretická část</b>	<b>13</b>
2.1 Jak komunikujeme	13
2.1.1 Jak komunikují lidé	13
2.1.1.1 Co je to konverzace	13
2.1.1.2 Jak konverzaci analyzovat?	13
2.1.1.3 Anatomie konverzace	15
2.1.2 Jak komunikují voiceboti	17
2.1.2.1 Jakto, že nám voiceboti rozumí?	17
2.1.2.2 Konverzační analýza	18
2.1.2.3 Jsou voiceboti inteligentní?	18
2.1.3 Jak mohou komunikovat lidé s voiceboty	19
2.1.3.1 Proč spolu chceme komunikovat?	19
2.1.3.2 Jak spolu můžeme komunikovat?	20
2.2 Literární rešerše	22
2.2.1 Proč jsem si k literární rešerši vybrala právě tyto zdroje?	22
2.2.2 Schopnosti a dovednosti	22
2.2.3 Schopnost vést konverzaci	23
2.2.3.1 Střídání replik (turn taking)	23
2.2.3.1.1 Jak to použít v praxi?	24
2.2.3.2 Tempo	24
2.2.3.2.1 Jak to použít v praxi?	24
2.2.3.3 Konverzační značky	25
2.2.3.3.1 Jak to použít v praxi?	25
2.2.3.4 Minimalizace latence	25
2.2.3.4.1 Jak to použít v praxi?	26
2.2.4 Vybudovat důvěru	26
2.2.4.1 Důvěra	26
2.2.4.1.1 Jak to použít v praxi?	27
2.2.5 Být konzistentní	27
2.2.5.1 Mít účel a cíl	27
2.2.5.1.1 Jak to použít v praxi?	28
2.2.5.2 Konzistence a osobnost	28
2.2.5.2.1 Jak to použít v praxi?	28
2.2.6 Přizpůsobit se příjemci	29
2.2.6.1 Recipient Design	29
2.2.6.1.1 Jak to použít v praxi?	29
2.2.7 Schopnost porozumět a naslouchat	30
2.2.7.1 Naslouchání	30

2.2.7.1.1 Jak to použít v praxi?	31
2.2.7.2 Potvrzování a souhlasy	32
2.2.7.2.1 Jak to použít v praxi?	32
2.2.7.3 Porozumění	33
2.2.8 Schopnost udržet kontext	34
2.2.8.1 Kontext	34
2.2.8.1.1 Jak to použít v praxi?	35
2.2.9 Pracovat s chybou a neporozuměním	36
2.2.9.1 Error Handling	36
2.2.9.1.1 Jak to použít v praxi?	36
2.2.9.2 Opravy	37
2.2.9.2.1 Jak to použít v praxi?	38
2.2.10 Minimalizovat	39
2.2.10.1 konverzační maxima	39
2.2.10.1.1 Maximum kvality a relevance	39
2.2.10.1.2 Jak to použít v praxi?	39
2.2.10.1.3 Maximum kvality	40
2.2.10.1.4 Jak to použít v praxi?	40
2.2.10.1.5 Maximum způsobu	40
2.2.10.1.6 Jak to použít v praxi?	40
2.2.10.2 Teorie kognitivní zátěže	41
2.2.10.2.1 Jak to použít v praxi?	41
2.2.10.3 Univerzální, ale konkrétní	42
2.2.10.3.1 Jak to použít v praxi?	42
2.2.11 Komunikovat lidsky	43
2.2.11.1 Zdvořilost a ohleduplnost	43
2.2.11.1.1 Jak to použít v praxi?	43
2.2.11.2 Small talk	43
2.2.11.2.1 Jak to použít v praxi?	44
2.2.11.3 Humor	44
2.2.11.3.1 Jak to využít v praxi?	44
<b>3 Metodologie</b>	<b>46</b>
3.1 Cíl práce a výzkumná otázka	46
3.2 Metody sběru dat	47
3.2.1 Výběr zkoumaného vzorku	47
3.2.2 Etické otázky výzkumu a práce s daty	48
3.2.3 Pomocné otázky k analýze	48
3.3 Zkoumaný vzorek	49
3.3.1 Rozsah zkoumaného vzorku a diverzifikace dat	49
3.3.2 Scénář 1: Zadání platby k úhradě	50
3.3.3 Scénář 2: Virtuální zdravotní sestra	50
3.3.4 Scénář 3: Vymáhání pohledávek	50
3.4 Limity výzkumu	51

<b>4 Analytická část</b>	<b>52</b>
4.1 Analýza scénářů	52
4.1.1 Scénář 1: Zadání platby k úhradě	52
4.1.1.1 Popis scénáře	52
4.1.1.2 Schopnost vést konverzaci	52
4.1.1.3 Vybudovat důvěru	53
4.1.1.4 Být konzistentní	53
4.1.1.5 Přizpůsobit se příjemci	54
4.1.1.6 Schopnost porozumět a naslouchat	56
4.1.1.7 Schopnost udržet kontext	56
4.1.1.8 Pracovat s chybou a nepochopením	57
4.1.1.9 Minimalizovat	62
4.1.1.10 Komunikovat lidsky	62
4.1.2 Scénář 2: Virtuální zdravotní sestra	62
4.1.2.1 Popis scénáře	62
4.1.2.2 Schopnost vést konverzaci	63
4.1.2.3 Vybudovat důvěru	63
4.1.2.4 Být konzistentní	65
4.1.2.5 Přizpůsobit se příjemci	65
4.1.2.6 Schopnost porozumět a naslouchat	66
4.1.2.7 Schopnost udržet kontext	66
4.1.2.8 Pracovat s chybou a nepochopením	66
4.1.2.9 Minimalizovat	67
4.1.2.10 Komunikovat lidsky	68
4.1.3 Scénář 3: Vymáhání pohledávek	68
4.1.3.1 Popis scénáře	68
4.1.3.2 Schopnost vést konverzaci	69
4.1.3.3 Vybudovat důvěru	70
4.1.3.4 Být konzistentní	71
4.1.3.5 Přizpůsobit se příjemci	71
4.1.3.6 Schopnost porozumět a naslouchat	72
4.1.3.7 Schopnost udržet kontext	72
4.1.3.8 Pracovat s chybou a nepochopením	72
4.1.3.9 Minimalizovat	73
4.1.3.10 Komunikovat lidsky	73
4.2 Výsledky analýzy	74
4.2.1 Výzkumná otázka a hypotézy	74
4.2.2 Ověření poznatků na empirických datech	75
4.2.2.1 Schopnost vést konverzaci	76
4.2.2.2 Vybudovat důvěru	76
4.2.2.3 Být konzistentní	77
4.2.2.4 Přizpůsobit se příjemci	77
4.2.2.5 Schopnost porozumět a naslouchat	78
4.2.2.6 Schopnost udržet kontext	78



4.2.2.7 Pracovat s chybou a neporozuměním	78
4.2.2.8 Minimalizovat	79
4.2.2.9 Komunikovat lidsky	79
4.2.2.10 Vzbudit respekt	80
4.2.3 Cíl práce: Vytvoření modelu pravidel	80
4.2.3.1 Matice pravidel	81
4.2.3.1.1 Technologické požadavky	82
4.2.3.1.2 Konverzační pravidla	83
4.3 Náměty pro další výzkum	84
<b>5 Závěr</b>	<b>86</b>
<b>6 Seznam použité literatury</b>	<b>87</b>
<b>7 Seznam obrázků</b>	<b>91</b>

## Předmluva

V této diplomové práci jsem si dala za cíl zodpovědět, jak mohou roboti přirozeně komunikovat s lidmi a vytvořit model pravidel konverzačního designu. Ten bude vytvořen na základě poznatků z odborné literatury a zároveň bude podložen daty z reálných konverzací lidí s voiceboty. Metodologií této práce je kvalitativní analýza. Téma práce se nijak neodlišuje oproti zadání.

V první části práce se blíže podívám na to, jak funguje komunikace mezi lidmi. Popíšu, jak se můžeme na mezilidskou komunikaci dívat, jaký je rozdíl mezi komunikací a konverzací, jak lze konverzaci analyzovat a jak k tomu přistupuje konverzační analýza. Následuje část o tom, jak komunikují hlasoví asistenti. Zde uvedu princip fungování voicebotů, z čeho se skládají, jak dokáží rozpoznat lidskou řeč a jak s ní pracují. V poslední části kapitoly Jak komunikujeme se zaměřím na to, jak můžeme analyzovat konverzaci voicebotů, proč s nimi vůbec chceme mluvit a jak na to. Druhou částí práce je literární rešerše. Zde zkoumám přístupy a pohledy autorů, odborníků na konverzační design, UX design, hlasové technologie, lingvistiku, sociologii, psychologii a další obory. Poznatky autorů shrnuji do devíti schopností a dovedností, které by měl mít virtuální asistent.

Následuje kapitola vysvětlující metodologii, výběr zkoumaného vzorku, limity výzkumu a práci s daty. Po této kapitole je analytická část, ve které zkoumám tři scénáře virtuálních asistentů a hovory, ve kterých s nimi lidé komunikovali. Na tyto scénáře se dívám z pohledu oněch devíti dovedností a analyzuji je pomocí sady otázek. Ve výsledcích analýzy pak shrnuji, zdali se interpretací empirických dat potvrdila potřeba všech definovaných schopností a zároveň přidávám desátou, která vyplynula z analýzy hovorů. V závěru práce pak z těchto poznatků vytvářím model pravidel, který dělím na technologické požadavky a na konverzační pravidla. V úplném závěru práce pak zmiňuji náměty k dalšímu výzkumu.

Tato diplomová práce obsahuje 122 475 znaků, což je 68 normostran.



# 1 Úvod

*“The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.”<sup>1</sup>*

*Mark Weiser*

Trh hlasových technologií v posledních uplynulých letech raketově roste a virtuální asistenti jsou čím dál častěji součástí našich každodenních životů. A to nejen v rámci chytrých domácností či ovládání v autech. Svého hlasového asistenta většina z nás nosí každý den v kapse ve svém chytrém telefonu. Rozhodně ne všichni jej však využívají. Může to mít mnoho důvodů. Strach z technologie. Nedůvěra v zabezpečení dat. Špatné předchozí zkušenosti s komunikací s hlasovým asistentem. Obava, že nezvládne vyřídit náš požadavek. Že si vzájemně neporozumíme. Nedůvěra v technologii jako takovou. To, že si s asistentem neporozumíme ale není chybou v nás. Je to způsobeno buď nekvalitní technologií nebo špatným konverzačním designem, případně obojím dohromady. Jak uvádím v kapitole Proč spolu chceme komunikovat, hlasové technologie nám mohou přinášet nesporné výhody a usnadnit náš život v mnoha ohledech.

Rychlému vývoji hlasových technologií by se měl přizpůsobit pohled na ovládání zařízení. V dnešní době, kdy je přepis řeči na text na velmi vysoké úrovni, už není dle mého názoru nutné, aby se lidé učili komunikovat řečí počítače. Naopak nastává doba, kdy bychom měli my lidé naučit počítače, aby s námi komunikovali naší řečí. Přirozeně.

Tomu se věnuje poměrně nový obor konverzačního designu. Ten vychází z analýzy lidské konverzace a její principy využívá při tvorbě hlasových rozhraní počítačů a virtuálních asistentů. Aby se mohly hlasové technologie *“vplést do každodenního života a zmizet”*, podobně jak ve svém citátu popisuje počítačový vědec Mark Weiser, musí být vzájemná interakce především přirozená. Jak toho docílit? A jaká pravidla pro takovou konverzaci platí? To se budu snažit zjistit v této diplomové práci.

---

<sup>1</sup> WEISER, M. The computer for the 21st Century. IEEE pervasive computing [online]. New York: IEEE, 2002, 1(1), 19-25 [cit. 2022-07-25]. ISSN 1536-1268. Dostupné z: doi:10.1109/MPRV.2002.993141, str. 1.

## 2 Teoretická část

### 2.1 Jak komunikujeme

#### 2.1.1 Jak komunikují lidé

Lidé mají schopnost, kterou nemá žádný jiný živočišný druh. A to naučit se bez jakékoliv formální výuky jazyk, kterému jsou od narození vystaveni.<sup>2</sup> Porozumět se děti učí několik měsíců, naučit se mluvit ale trvá zhruba rok. Porozumět mluvené řeči a schopnost ji ovládat jsou tedy dvě různě složité věci.<sup>3</sup> Což platí jak pro nás lidi, tak pro počítače.

##### 2.1.1.1 Co je to konverzace

Konverzaci můžeme vnímat jako skupinovou činnost, ve které má každý účastník svá práva a povinnosti.<sup>4</sup> Můžeme na ni nahlížet jako na kooperativní hru – podobně jako třeba na šachy. V této hře nehrajeme proti svému protihráči, ale s ním. Vzájemně spolupracujeme a oba dodržujeme určitá pravidla. Zároveň se od protihráče můžeme učit a postupně rozvíjet své dovednosti.<sup>5</sup> Stejně tak si můžeme předat znalosti při konverzaci a zároveň si vytvořit nové – každý účastník dialogu je totiž *"učícím se systémem, který se díky dané zkušenosti vnitřně mění."*<sup>6</sup> A to je právě to, co odlišuje konverzaci od komunikace. Rozdíl mezi konverzací a komunikací spočívá v tom, že se v průběhu něco mění, například porozumění, hodnoty, záměr a podobně. V opačném případě se jedná pouze o výměnu zpráv.<sup>7</sup>

##### 2.1.1.2 Jak konverzaci analyzovat?

Podle autora knihy *Good Talk: How to Design Conversations that Matter* můžeme na konverzaci nahlížet z osmi úhlů. Tím prvním je cíl. Zde se můžeme ptát: Jaký má konverzace účel? Proč bychom se do ní měli zapojit? Druhým úhlem je zahájení – jak konverzaci zahájit? Třetí pohled, který můžeme zkoumat, je střídání mluvčích. Ptáme se kdo a kdy mluví a také

<sup>2</sup> ENFIELD, N. J. *How we talk: the inner workings of conversation*. New York: Basic Books, 2017, ix, 257 stran : ilustrace ; 22 cm. ISBN 978-0-465-05994-2, str. 123.

<sup>3</sup> PIERACCINI, Roberto. *The voice in the machine: building computers that understand speech*. Cambridge, Massachusetts: MIT Press, 2012, 1 online resource (355 p.). ISBN 1-280-49915-X, str. xxii.

<sup>4</sup> ENFIELD, N. J. *How we talk: the inner workings of conversation*. New York: Basic Books, 2017, ix, 257 stran : ilustrace ; 22 cm. ISBN 978-0-465-05994-2, str. 13.

<sup>5</sup> STILLMAN, Daniel. *Good talk: how to design conversations that matter*. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8, str. 32-33.

<sup>6</sup> DUBBERLY, Hugh a Paul PANGARO. *What is conversation? How can we design for effective conversation*. *Interactions Magazine*. 2009, 16.4, str. 22-28.

<sup>7</sup> PANGARO, Paul. *Questions for conversation theory or conversation theory in one hour*. *Kybernetes*. 2017, 46,9, str. 1581.

jak pozná, že je na řadě. Čtvrtým tématem ke zkoumání je otázka jak rozpoznat chybu a jak ji opravit. Pátým je “nit konverzace”. Zajímá nás, co je prvkem, který drží konverzaci pohromadě a kdo má zodpovědnost za to, že se nit nepřetrhne. Šestým úhlem je tempo, dále se pak můžeme zaměřit samotné účastníky. Posledním bodem je moc, kdy si můžeme pokládat otázku, kdo má právo průběh konverzace výrazně měnit.<sup>8</sup> Těchto osm úhlů pohledu lze využít při zkoumání dialogu mezi člověkem a robotem, jak si ukážeme později.

S pohledem, jak konverzaci zkoumat, však přišel již mnohem dříve americký sociolog Harvey Sacks. Ten v 60. letech zavedl metodou konverzační analýzy, ve které popisuje konverzaci jako “*speech-exchange system*”<sup>9</sup>, přičemž hlavní princip spočívá ve výměně mluvčích. Taková konverzace má následující vlastnosti:

1. *“Výměna mluvčích se opakuje, nebo alespoň nastává.*
2. *Většinou mluví jen jedna strana.*
3. *Výskyt více mluvčích najednou je běžný, ale netrvá dlouho.*
4. *Většina přechodů je bez větší mezery a bez většího překrývání.*
5. *Pořadí střídání mluvčích není pevně dané a mění se.*
6. *Délka repliky není pevně stanovena a mění se.*
7. *Délka rozhovoru není předem určena.*
8. *To, co strany říkají, není předem specifikováno.*
9. *Vzájemné rozdělení výměn replik není pevně stanoveno, ani předem specifikováno.*
10. *Počet stran se může měnit.*
11. *Rozhovor může být souvislý nebo přerušovaný.*
12. *Používají se techniky rozdělování tahů. Aktuální mluvčí může vybrat dalšího mluvčího (například když adresuje otázku třetí straně), strany si mohou samy vybrat, kdy začnou mluvit.*
13. *Používají se různé replikové konstrukční jednotky; např. replika může být jednoslovná, nebo může mít délku věty.*

---

<sup>8</sup> STILLMAN, Daniel. Good talk: how to design conversations that matter. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8, str. 54.

<sup>9</sup> SACKS, Harvey, Emmanuel SCHEGLOFF a Gail JEFFERSON. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Sotsiologicheskoe Obozrenie / Russian Sociological Review* [online]. 2015, 14(1), 142-202 [cit. 2022-07-10]. ISSN 1728-192X. Dostupné z: doi:10.17323/1728-192X-2015-1-142-202

14. *Existují opravné mechanismy pro řešení chyb a při porušení výměny replik; např. pokud se dvě strany ocitnou v situaci, kdy mluví současně, jedna z nich přestane mluvit, čímž dojde k opravě.*"<sup>10</sup>

Vedle konverzační analýzy a jazykových pravidel, tedy gramatiky, popisuje autor knihy *How We Talk* nepsaná "pravidla konverzace". Dívá se na ně spíše z morálního hlediska. Například pokud jedna strana položí druhé otázku, měla by na ní druhá strana odpovědět, nebo říct proč nemůže odpovědět. Pokud tato otázka byla položena třetí straně, neměla by na ni odpovědět druhá strana.<sup>11</sup> Správné střídání mluvčích může mít vliv na to, jestli lidé budou vnímat protistranu jako nezdvořilou, či přátelskou.<sup>12</sup>

Při analýze konverzace bychom se měli dívat spíše než na téma (o čem se mluví) na způsob (jak se mluví).<sup>13</sup> K tomu nám pomůže, pokud se podíváme blíže na samotnou anatomii konverzace.

### 2.1.1.3 Anatomie konverzace

Za nejmenší jednotku při analýze konverzace můžeme považovat replikovou konstrukční jednotku (TCU). Ta může být složena z jednotlivých slov, frází či celých vět. Výpověď (utterance) vždy obsahuje alespoň jednu replikovou konstrukční jednotku, zpravidla se jich v jedné výpovědi nachází více. Výpovědi jsou organizovány do sekvencí a sekvence pak můžeme organizovat do aktivit. Aktivitami se v konverzační analýze myslí řady souvisejících výroků, které mají společný cíl (například zahájení konverzace, řešení problému). Každá konverzace je ohraničená začátkem dialogu a jeho koncem a obsahuje minimálně jednu aktivitu.<sup>14</sup>

---

<sup>10</sup> SACKS, Harvey, Emmanuel SCHEGLOFF a Gail JEFFERSON. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Sotsiologicheskoe Obozrenie / Russian Sociological Review* [online]. 2015, 14(1), 142-202 [cit. 2022-07-10]. ISSN 1728-192X. Dostupné z: doi:10.17323/1728-192X-2015-1-142-202.

<sup>11</sup> ENFIELD, N. J. *How we talk: the inner workings of conversation*. New York: Basic Books, 2017, ix, 257 stran : ilustrace ; 22 cm. ISBN 978-0-465-05994-2, str. 13.

<sup>12</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. *Safety First: Conversational Agents for Health Care*. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 42.

<sup>13</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 28.

<sup>14</sup> SCHEGLOFF, Emanuel A. a Harvey SACKS. *Opening up closings*. *Semiotica*. 1973, 8(4). Dostupné z: doi:10.1515/semi.1973.8.4.289, str. 71.

Robert Moore tuto organizaci do sekvencí a aktivit znázornil v tabulce níže, kterou publikoval své knize *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. V knize vytváří vlastní rámec přirozené komunikace založený na poznatcích z konverzační analýzy.<sup>15</sup>

Turns	Sequences	Activities
A: Hello	GREETING	Opening
B: Hi!	GREETING	
A: How's it going?	WELFARE CHECK	
B: Good	WELFARE REPORT	
A: Hey, are ya going tonight?	PRE-REQUEST	Requesting a Ride
B: Yes	ANSWER	
A: Can I get a ride?	REQUEST	
B: Sure!	GRANT	
A: What time?	INQUIRY	
B: Seven o'clock	ANSWER	
A: Okay.	ACKNOWLEDGMENT	
A: Thank you so much!	PRE-CLOSING	
B: No problem!	PRE-CLOSING	Closing
A: Bye	FAREWELL	
B: Buh bye	FAREWELL	

obr. 0 – příklad analýzy konverzace<sup>16</sup>

Autor v knize dále říká, že konverzační analýza výroky charakterizuje a kategorizuje na základě typu sociální akce. Mezi ně můžeme řadit kromě výše zmíněného také informování, poskytování rad, vyprávění příběhů, hodnocení a podobně. Konverzační analytici přišli s obecně platnými koncepty, jak jsou aktivity v konverzaci uspořádány. Mluvíme o střídání mluvčích (turn-taking), sekvencových párech (adjacency pairs), uspořádání sekvencí (sequence organisation), opravách (repair) a preferenční organizaci (preference organisation). Podle autora z konverzační analýzy vyplývají tři základní principy, které si můžeme propůjčit také do konverzačního designu při tvorbě virtuálních asistentů. Třemi základními principy při tvorbě jsou přizpůsobení se příjemci (recipient design), minimalizace (minimization) a možnost opravy.<sup>17</sup>

<sup>15</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 29-31.

<sup>16</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 29-31.

<sup>17</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 30-31.



## 2.1.2 Jak komunikují voiceboti

### 2.1.2.1 Jakto, že nám voiceboti rozumí?

Jak může virtuální asistent rozumět tomu, co říkáme a adekvátně na to reagovat? Pro lepší pochopení toho, jak voiceboti fungují, je potřeba se podívat na to, z čeho se skládají. To můžeme rozdělit na tři základní prvky. Tím prvním je schopnost přepsat vstup uživatele, tedy jeho řeč, na text. V běžném životě, pokud jeden člověk nerozumí druhému, může ho poprosit, aby danou věc napsal. Ve světě virtuálních asistentů to má na starosti komponenta speech-to-text (STT).<sup>18</sup> Tento text následně analyzuje systém na zpracování přirozeného jazyka (NLP), který jej klasifikuje a zařadí do správné kategorie intentu (záměru uživatele). Každý virtuální asistent se řídí podle předem připraveného scénáře, který si můžeme představit jako rozhodovací konverzační strom a podle intentů uživatele se bot rozhoduje, jak bude v konverzaci pokračovat. Součástí tohoto stromu jsou předem připravené odpovědi voicebota (prompts). Komponenta text-to-speech (TTS) pak tuto předpřipravenou odpověď přečte syntetickým hlasem. Pojdme si nyní blíže popsat fungování nejdůležitější části, kterou je “mozek bota”, tedy NLP.

Systém na zpracování přirozeného jazyka funguje na principu textové klasifikace. K té můžeme mít dva různé přístupy. Prvním je analýza textu založená na extrakci klíčových slov a pravidlech (rule-based přístup). Druhý přístup zahrnuje využití strojového učení (machine-learning based přístup). Čím se oba přístupy liší a jaký to má vliv na konverzační design?

Rule-based přístup je zaměřený na vyhledávání klíčových slov v textu. Zásadním rozdílem oproti machine-learning based přístupu je to, že při tvorbě scénáře je potřeba přemýšlet nad tím, jaká slova lidé mohou použít při popisu svého záměru. Tedy, jaká klíčová slova má NLP najít v textu, aby danou řeč klasifikovalo do správné kategorie intentů. K tomu nám mohou pomoci regulární výrazy (např. do intentu, ve kterém je klíčové slovo zapsáno jako ^ano\$, bude NLP zařazovat jen tu řeč, ve které zaznamená pouze toto jedno konkrétní slovo) či logické operace (AND, OR, NOT, NEAR), které určují vazbu mezi jednotlivými slovy v textu.

---

<sup>18</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. xi.

Oproti tomu machine-learning based přístup pracuje s množinou dat, jak v minulosti lidé svůj záměr vyslovili. Tato data je nejprve nutné manuálně oštitkovat, aby bylo zřejmé, do jaké kategorie intentů spadají. Text následně NLP lemmatizuje (převede slova na základní tvar) a očistí od výplňových slov a znaků, aby se snížilo množství kombinací frází, na kterých se neuronová síť učí. Takto vznikne natrénovaný model, který je následně možné použít pro detekování intentu.

Díky rozpoznání řeči, jejímu zpracování, správně postavenému konverzačnímu stromu a syntetickému hlasu zvládne voicebot komunikovat s lidmi funkčně po technické stránce. Je však důležité, aby byla konverzace nejen funkční, ale také přirozená.

### 2.1.2.2 Konverzační analýza

Pokud na konverzaci člověka s robotem nahlédneme optikou konverzační analýzy, zjistíme, že vlastnosti, které platí pro mezilidskou komunikaci, pro tu s voicebotem platí jen částečně. Odlišné fungování můžeme najít hned v několika vlastnostech, které by podle Harveyho Sackse konverzace měla mít.<sup>19</sup> Přejechy mezi promluvami mohou vzhledem k technologickým omezením doprovázet mezery. Neplatí, že to, co obě strany říkají není předem specifikováno. Promluvy bota jsou totiž dopředu naskriptované. Třetím bodem, který se nedá na konverzaci s botem aplikovat, je fakt, že se počet mluvčích může libovolně měnit.<sup>20</sup> Při designu konverzačních stromů se pracuje zpravidla pouze se dvěma účastníky konverzace, tedy s voicebotem na jedné straně a s člověkem na druhé. Bylo by velmi technologicky náročné, pokud by na druhé straně byli lidé dva. Není to však nemožné. V současné době existuje pokročilá hlasová biometrie<sup>21</sup>, díky které lze rozeznat pohlaví či věk mluvčího a jednotlivé mluvčí by tak šlo od sebe odlišit.

### 2.1.2.3 Jsou voiceboti inteligentní?

V roce 1980 přišel filozof John Searle s myšlenkovým experimentem, který známe jako Argument čínské pokoje.<sup>22</sup> Představme si situaci, kdy je člověk, který neovládá čínštinu, zavřený v místnosti s čínskými texty. Tomuto člověku pokládáme otázky, on má

---

<sup>19</sup> Konkrétně jsou to body 4, 5, 8 a 10.

<sup>20</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 25.

<sup>21</sup> Zmiňme například brněnskou firmu Phonexia, která se hlasovou biometrií zabývá.

<sup>22</sup> SEARLE, John. Chinese room argument. *Scholarpedia journal* [online]. 2009, 4(8), 3100 [cit. 2022-07-25]. ISSN 1941-6016. Dostupné z: doi:10.4249/scholarpedia.3100

však k dispozici překladový slovník, na základě kterého může i bez znalosti jazyka přijít se smysluplnou odpovědí. Pro někoho, kdo se nachází mimo tuto místnost se může zdát, že člověk ovládá čínštinu perfektně. Můžeme si položit otázku – pokud by se takto choval počítač, mohli bychom jej považovat za inteligentní? Searle chtěl tímto experimentem poukázat na to, že schopnost zodpovědět otázku ještě neznamená, že jsme prokázali schopnost porozumět. A to stejné platí i u počítačů.<sup>23</sup>

Jak moc musí být voicebot inteligentní, aby dokázal komunikovat s lidmi? V otázce konverzačního designu nejde o to, aby lidé nepoznali, že hovoří se strojem. Jednak by to bylo neetické a zároveň vědomí toho, že člověk hovoří s virtuálním asistentem, může pomoci ke vzájemně lepšímu porozumění. Z analyzovaných hovorů, kterým se budu věnovat v další části diplomové práce totiž vyšlo najevo, že lidé přizpůsobují svou řeč tomu, s kým mluví. Pokud mluví se strojem, jsou stručnější, nemají tendenci se příliš rozpovídat a v některých případech mluví heslovitě. Takováto řeč je pak jednodušší na zpracování a eliminuje možné chyby. Pokud mluvíme o inteligentních virtuálních asistentech, máme na mysli technologicky kvalitní systém s dobře připraveným konverzačním scénářem, ve kterém je robot připraven reagovat na ideálně všechny požadavky klienta, které v dané konverzaci mohou nastat.

### **2.1.3 Jak mohou komunikovat lidé s voiceboty**

#### **2.1.3.1 Proč spolu chceme komunikovat?**

Aby mohli lidé v minulosti ovládat počítač, bylo potřeba se naučit mluvit v jazyce počítače. Dnes se učí počítače mluvit jazykem uživatele. Díky tomuto vývoji vznikla disciplína user experience (UX), ve které se design zaměřuje na prožitek uživatele, který komunikuje s technologií. Díky technologickému pokroku ve zpracování řeči a strojovém učení se v posledních letech objevila oblast virtuálních asistentů, kteří více či méně komunikují jako lidé.<sup>24</sup>

Výhody hlasových technologií jsou nesporné. Proč jsou hlasové technologie na vzestupu lze shrnout do následujících výhod. Tou první je rychlost. Nedávná studie ze

---

<sup>23</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. xi.

<sup>24</sup> MOORE, Robert J., Margaret H. SZYMANSKI, Raphael. ARAR a Guang-Jie. REN. Studies in Conversational UX Design. Cham: Springer International Publishing, 2018, 1 online resource (204 pages). ISBN 9783319955797. Dostupné z: doi:10.1007/978-3-319-95579-7, str. 1.

Stanfordovy univerzity ukazuje, že diktování zpráv je mnohem rychlejší, než psaní.<sup>25</sup> Mluvená řeč má potenciál být nejen nejrychlejší, ale také nejefektivnější.<sup>26</sup> Když mluvíme, předáváme druhé straně mnoho informací také díky tomu, jaký tón hlasu použijeme, zdali mluvíme rychle či pomalu, potichu, nebo nahlas, nebo v promluvě uděláme pauzu. Řeč tedy nese více informací, než psaný text. Další výhodou je praktičnost. Hlasová ovládní nám umožňují odložit zařízení z rukou, což může být při některých činnostech praktičtější i bezpečnější (například při vaření či řízení auta). Ovládní pomocí hlasu může být také intuitivnější. Řeč totiž ovládáme přirozeně a hlasová uživatelská rozhraní poskytují jednodušší onboarding v porovnání s rozhraními, ve kterých se uživatel musí nejprve zorientovat, aby věděl, jak s nimi pracovat.<sup>27</sup>

Hlasové technologie jsou všude okolo nás, je však na místě zmínit, že se nehodí pro všechny situace. Možné nevýhody s sebou nese využití na veřejných místech. Zadávání hlasových příkazů počítači v přeplněné open-space kanceláři by uživatelům způsobilo spíše diskomfort. Je potřeba brát ohled také na soukromí. Pokud s virtuálním asistentem vedeme konverzaci, ve které sdělujeme rodné číslo nebo zdravotní stav, pravděpodobně budeme takovýto dialog vést raději v soukromí.<sup>28</sup>

### 2.1.3.2 Jak spolu můžeme komunikovat?

Počítače jsou lepší než lidé ve spoustě činnostech. Například v hraní šachů, v zapamatování si neuvěřitelného množství čísel či v počítání složitých matematických vzorců. Ale nejsou lepší než my ve dvou nejběžnějších a nejpřirozenějších aktivitách, a to je mluvení a porozumění řeči.<sup>29</sup> Lidé totiž při komunikaci využívají dovednost, kterou můžeme nazvat jako presupozice a implikace, kdy posluchač porozumí významům, které však mluvčí explicitně nezmínil. Umíme číst mezi řádky. Tato dovednost je pro virtuálního asistenta

---

<sup>25</sup> RUAN, Sherry, Jacob WOBBEROCK, Kenny LIOU, Andrew NG a James LANDAY. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. Proceedings of ACM on interactive, mobile, wearable and ubiquitous technologies [online]. Ithaca: ACM, 2018, 1(4), 1-23 [cit. 2022-07-25]. ISSN 2474-9567. Dostupné z: doi:10.1145/3161187

<sup>26</sup> MARGE, Matthew, Carol ESPY-WILSON, Nigel G WARD, et al. Spoken language interaction with robots: Recommendations for future research. Computer speech & language [online]. LONDON: Elsevier, 2022, 71, 101255 [cit. 2022-07-06]. ISSN 0885-2308. Dostupné z: doi:10.1016/j.csl.2021.101255, str. 2.

<sup>27</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 1-5.

<sup>28</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 1-5.

<sup>29</sup> PIERACCINI, Roberto. The voice in the machine: building computers that understand speech. Cambridge, Massachusetts: MIT Press, 2012, 1 online resource (355 p.). ISBN 1-280-49915-X, str. xxii.

obzvláště obtížná, protože se opírá o selský rozum. Ten využíváme také při pochopení ironie či metafor, kterým má počítač také problém porozumět.<sup>30</sup>

Jak tedy mohou lidé a stroje vzájemně komunikovat? Klíčem je perfektní znalost toho, jak funguje konverzace mezi lidmi. A pravidla, která platí v mezilidské komunikaci, pak v co největší míře, tam, kam nám to technologie dovolí, využít při tvorbě uživatelských rozhraní. Mezi ně patří i ta hlasová, tedy voiceboti, o kterých je tato diplomová práce. Obor, který se věnuje využití pravidel mezilidské komunikace pro tvorbu hlasových uživatelských rozhraní, se nazývá konverzační design. Jak tato pravidla využít a k čemu slouží rozeberu v následující kapitole, kde zkoumám, jak se na principy konverzačního designu dívají autoři odborné literatury.

---

<sup>30</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 46.

## **2.2 Literární rešerše**

Jaká jsou pravidla designu hlasových robotů, aby s lidmi komunikovali přirozeně? V této kapitole shrnu nejdůležitější poznatky, se kterými přišli autoři, kteří se konverzačním designem dlouhodobě zabývají. V následující praktické kapitole tato pravidla podložím ukázkami reálných situací z hovorů a vysvětlím, proč člověk či bot postupoval právě takto.

### **2.2.1 Proč jsem si k literární rešerši vybrala právě tyto zdroje?**

V oblasti konverzační AI se pohybuji téměř tři roky, za tu dobu jsem měla možnost vyzkoušet si různé aspekty práce v tomto oboru, od rešerše trhu, přes marketing, analýzy potřeb zákazníků k nespočtu iterací produktu až po stavění konverzačních scénářů pro voiceboty. V rámci zvyšování kvalifikace jsem prošla kurzy konverzačního designu, sleduji nejnovější poznatky z oboru a mám přehled o tom, které společnosti jsou v konverzační umělé inteligenci nejdále. Zároveň jsem měla možnost zjistit, která literatura je ve studiích citovaná nejvíce či je hojně doporučována odborníky z oboru.

Vybrala jsem si knihy a studie od autorů, kteří mají s konverzačním designem letité zkušenosti. Většina z nich se pohybuje či pohybovala ve firmách, které stály u zrodu hlasových technologií, rozpoznávání řeči, syntetických hlasů a podobně (např. IBM, Google, Amazon, Nuance Communication atd) a jejich názory jsou proto podloženy roky práce i daty z reálných konverzací, případně zkoumají danou problematiku na akademické půdě. Autoři často čerpají ve své práci také z jiných oborů jako je lingvistika, psychologie či sociologie.

### **2.2.2 Schopnosti a dovednosti**

Každý z autorů, jejichž práci jsem zkoumala, se na problematiku dívá z trochu jiného pohledu. Při studiu literatury jsem také zjistila, že se témata, o kterých mluví, často překrývají a jsou vzájemně propojená. Například pokud je bot konzistentní, máme k němu větší důvěru, pokud je stručný, lépe si rozumíme a podobně. Některé aspekty konverzačního designu se týkají samotné tvorby konverzačního stromu či psaní výpovědí pro voicebota. Jiné se zaměřují spíše na technologickou stránku věci a další charakterizují očekávání lidí od chování virtuálních asistentů.

Témata jsem proto rozdělila do jednotlivých kategorií a rozhodla jsem se je pojmut tak, aby je spojovaly schopnosti a dovednosti voicebota, které jsou zásadní pro to, aby s ním

byli lidé vůbec funkčně schopni a ochotni komunikovat. Z této syntézy vyšlo celkem devět dovedností, bez kterých se virtuální asistent neobejde. Je to schopnost vést konverzaci, vybudovat důvěru, mít cíl a být konzistentní, přizpůsobit se příjemci, porozumět a naslouchat, schopnost udržet kontext, schopnost pracovat s chybou a neporozuměním, minimalizovat a být lidský. Každé dovednosti se nyní budu věnovat ve větším detailu a přidám také příklady dobré praxe.

### **2.2.3 Schopnost vést konverzaci**

Při analýze toho, jak funguje konverzace a také témat, která v literatuře zkoumají odborníci na konverzační design jsem k došla ke dovednosti, která je pro voicebota zcela nezbytná, a tou je schopnost vést konverzaci. Slovo vést je zde zásadní. Každý virtuální asistent totiž má nějaký účel a také předem naskriptovaný konverzační strom, podle kterého se řídí. Člověk však zpravidla toto pozadí voicebota nezná a je proto na virtuálním asistentovi, aby konverzaci vedl. Co je k tomu potřeba a proč?

#### **2.2.3.1 Střídání replik (turn taking)**

Konverzace vždy vyžaduje více, než jednoho účastníka. Pokud by tomu tak nebylo, nemluvili bychom o dialogu, ale o monologu. Z toho vyplývá, že je postavená na střídání se v tom, kdo kdy mluví, k čemuž se vztahuje zavedený výraz turn taking.<sup>31</sup> Se střídáním replik souvisí také párová sekvence dvojic (adjacency pairs) na sebe navazujících výpovědí. Může to být například základní sekvence otázka-odpověď, žádost-vyhovění/odmítnutí a podobně.<sup>32</sup> Znalost druhů párových sekvencí nám může pomoci při designu konverzace.

V konverzaci mezi lidmi probíhá střídání mluvčích přirozeně a dialog s voicebotem či chatbotem je na turn takingu založen také. K tomu, aby byla konverzace plynulá je zde ale mnohem více potřeba, aby bylo zřejmé, kdo mluví a kdo poslouchá. Člověk by měl intuitivně vědět, kdy se konverzace střídá a mluvčím je on. K tomu slouží otázky či pobídky, kdy bot člověku napoví, nebo jej nasměruje. Dělíme je na následující druhy: otázky s otevřeným koncem, otázky z nabídky, otázky typu ano/ne, otázky týkající se umístění, kvantifikační

---

<sup>31</sup> SACKS, Harvey, Emmanuel SCHEGLOFF a Gail JEFFERSON. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Sotsiologicheskoe Obozrenie / Russian Sociological Review* [online]. 2015, 14(1), 142-202 [cit. 2022-07-10]. ISSN 1728-192X. Dostupné z: doi:10.17323/1728-192X-2015-1-142-202, str. 1.

<sup>32</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 51-53.

otázky a instrukce.<sup>33</sup> Podle studie, která analyzovala hovory v zákaznických centrech mají otázky kladené v úvodu hovoru za cíl zjistit informaci o zákazníkovi a identifikovat problém, se kterým volá, což zároveň následně určí rozsah a trajektorii hovoru. V těle hovoru jsou pak kladeny otázky, které zjišťují podrobnosti o problému a jeho rozsah.<sup>34</sup>

#### 2.2.3.1.1 Jak to použít v praxi?

Příklady takových otázek či pobídek mohou být následující:

1. Otázky s otevřeným koncem: *“Jak Vám můžu pomoci?”*
2. Otázky z nabídky: *“Chcete balíček odeslat, na nějaký čekáte, nebo pracujete jako kurýr?”*
3. Otázky typu ano/ne: *“Přejete si pokračovat? Prosím vyslovte ano nebo ne.”*
4. Otázky týkající se umístění: *“Jaká je prosím Vaše adresa?”*
5. Kvantifikační otázky: *“Kolik lidí pojedete s odtahovou službou?”*
6. Instrukce: *“Nyní mi sdělte poznávací značku vašeho vozidla.”*

#### 2.2.3.2 Tempo

V konverzaci je potřeba udržet určité tempo. V mezilidské komunikaci můžeme vnímat tempo jako jakýsi rytmus konverzace a to jak se mění, zdali se zrychluje či zpomaluje, nám zpravidla o dialogu může mnoho sdělit.<sup>35</sup> Zrychlení výměny názorů může znamenat blížící se hádku, jiný rytmus bude mít třeba terapeutický rozhovor.

#### 2.2.3.2.1 Jak to použít v praxi?

Z dobré praxe však vyplývá, že to pro komunikaci mezi robotem a člověkem platí jen částečně. Voicebot by měl být schopen udržet tempo konverzace. K tomu mu slouží prvky aktivního naslouchání a pobídky k tomu, aby člověk mluvil, konverzační značky, které informují, která část konverzace se právě odehrává. Pokud mluvíme o zrychlení či zpomalení, zde to tak úplně neplatí. V praxi se zpomalení syntetické řeči používá při čtení složitých prvků, jako je například email či adresa, ale tato změna tempa se netýká změny rytmu konverzace. Slouží nám jako nástroj pro lepší porozumění. Z řeči člověka můžeme však

---

<sup>33</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 139.

<sup>34</sup> SZYMANSKI, Margaret H. a Robert J. MOORE. Adapting to Customer Initiative: Insights from Human Service Encounters. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 19-32 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_2, str. 21, 24.

<sup>35</sup> STILLMAN, Daniel. Good talk: how to design conversations that matter. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8, str. 94.



zjistit reakční dobu, tedy to, jak rychle reaguje na to, co voicebot říká. V průběhu konverzace se pak této rychlosti může bot přizpůsobit. Pokud například mluví se starým člověkem, který může mít reakční dobu delší, konverzaci přizpůsobíme tak, aby měl dostatek prostoru vyjádřit to, co chce říct a nestalo se, že mu voicebot skočí do řeči.

### 2.2.3.3 Konverzační značky

Konverzační značky nám předávají informaci, kde se v konverzaci právě nacházíme a že je nám porozuměno. Zároveň pomáhají s přechodem mezi interakcemi a drží je pohromadě. Právě konverzační značky jsou častokrát to, díky čemu má při konverzaci s botem uživatel pocit z plynulé konverzace. Pokud bychom je do vytváření výpovědi bota nezahrnuli, mohl by dialog působit příliš stroze. Konverzační značky jsou časové údaje, potvrzení a poděkování a zpětná vazba.<sup>36</sup>

#### 2.2.3.3.1 Jak to použít v praxi?

V praxi mohou být konverzační značky následující:

1. Časové údaje:
  - a. *“Nejprve mi prosím sdělte registrační značku Vašeho vozidla.”*
  - b. *“Už jsme v polovině. Nyní budu potřebovat Váš rok narození.”*
  - c. *“A poslední otázka, doporučil byste tuto službu vašim známým?”*
2. Poděkování a potvrzení:
  - a. *“Děkuji. Rozuměl jsem správně, že je Vaše jméno Jan Novák?”*
  - b. *“Mám to. Teď si otevřu databázi zásilek, prosím chvíli strpení.”*
3. Zpětná vazba:
  - a. *“Výborně, to rád slyším. Je ještě něco, co byste chtěla dodat?”*
  - b. *“Mrzí mě, že se necítíte dobře. Hned Vás přepojím na sestru.”*

#### 2.2.3.4 Minimalizace latence

Vzájemné střídání se v roli mluvčího by mělo být plynulé. Jedním z omezení hlasové technologie ale může být latence, která se projevuje delší pauzou mezi ukončením promluvy člověka a začátkem promluvy voicebota, případně se objevuje mezi dvěma promluvami bota. Zpravidla ke zpoždění může docházet z důvodu špatné konektivity, při zpracování většího

---

<sup>36</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 39-40.

množství informací systémem a také pokud bot přistupuje k nějaké databázi.<sup>37</sup> Pokud je latence příliš vysoká, má to vliv na klientské hodnocení vyřízení požadavku.<sup>38</sup>

#### 2.2.3.4.1 Jak to použít v praxi?

Příliš dlouhá pauza může působit jako technická chyba. Obzvláště v telefonním hovoru se může i pár sekund ticha zdát jako poměrně dlouhá doba. Pokud tedy víme, že v nějaké části konverzace může pauza nastat, je dobré o tom předem informovat, například těmito způsoby:

1. *“Moment, ověřuji si číslo pojistné smlouvy v databázi.”*
2. *“Jen si otevřu si seznam zásilek, prosím o chvíli strpení.”*

### 2.2.4 Vybudovat důvěru

#### 2.2.4.1 Důvěra

Důvěra je klíčovou součástí jak mezilidské konverzace, tak také komunikace mezi člověkem a robotem. Pokud lidé virtuálnímu asistentovi důvěřují, konverzace jim připadá lidštější a jsou v průběhu takovéto konverzace ochotni sdělit mnohem větší množství informací, než kdyby hovořili s asistentem, kterému nevěří. K tomu, aby lidé vnímali roboty jako důvěryhodnější, nebo si k nim dokonce vybuďovali vztah, přispívá, pokud je bot schopem větší sociální interakce, než jen vést jasně ohraničenou konverzaci založenou na úkolech (např. požadavek na spuštění hudby, dotaz na aktuální počasí a podobně).<sup>39</sup>

Ve studii, která zkoumala aspekty mezilidské konverzace a porovnávala je s tou “virtuální” vyšlo najevo, že důvěra v konverzaci mezi lidmi bývá spojená se sdílením osobních informací a také v možnosti projevit zranitelnost – lidé mluvili zpravidla o emoční důvěře. V konverzacích s roboty se pak důvěra objevuje spolu s požadavkem na transparentnost, bezpečnost dat a soukromí. Lidé ji zmiňovali jako “*vstupní bránu pro další*

---

<sup>37</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 56.

<sup>38</sup> SARIKAYA, Ruhi. The Technology Behind Personal Digital Assistants: An overview of the system architecture and key components. IEEE signal processing magazine [online]. PISCATAWAY: IEEE, 2017, 34(1), 67-81 [cit. 2022-07-25]. ISSN 1053-5888. Dostupné z: doi:10.1109/MSP.2016.2617341, str. 79.

<sup>39</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. *Conference on Human Factors in Computing Systems - Proceedings* [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 3-4.

interakci”.<sup>40</sup> Na důvěryhodnost má velký vliv také osobnost bota. Pokud bude bot pro člověka familiární, bude s ním jednat snadněji.<sup>41</sup>

#### 2.2.4.1.1 Jak to použít v praxi?

Virtuální asistent by měl komunikovat transparentně. Pokud se jedná o telefonní hovor, je dobré při představení zmínit, odkud voicebot volá, případně kam se člověk dovolal. Aby nastavil už od začátku komunikace správné očekávání, je dobré, když bot na začátku uvede na pravou míru, že člověk mluví s virtuálním asistentem. V současné době jsou syntetické hlasy na takové úrovni, že je může být obtížné rozeznat od lidského hlasu (pro příklad uvádím hlas Antonín Neural od Microsoftu<sup>42</sup>) a pro starší lidi může být problém rozeznat, že nemluví s člověkem. Pokud máme k dispozici nějaká data o člověku (například jméno), může jej v rámci zvýšení důvěryhodnosti voicebot použít a konverzaci personalizovat.

Důvěru lze podpořit také tím, že při dotazu na citlivé údaje není potřeba sdělit údaj celý, ale jen jeho část. Například čtyři poslední čísla rodného čísla, pojistné smlouvy a podobně. Je dobré také myslet na to, že se člověk, který mluví s voicebotem, nemusí nacházet v situaci, kdy tyto údaje může sdělit a nabídnout jiné řešení, například zadání rodného čísla na klávesnici. Samozřejmostí je zabezpečení dat, které se skrze komunikaci s virtuálním asistentem přenáší, veškeré osobní údaje by měly být chráněny v rámci platné legislativy. Na dotaz člověka by měl být virtuální asistent schopen dotazy o zpracování dat zodpovědět.

### 2.2.5 Být konzistentní

#### 2.2.5.1 Mít účel a cíl

Konverzace mezi lidmi má buď sociální účel nebo transakční (konverzace zaměřená na shromáždění informací, které jsou potřeba k dokončení konkrétního úkolu).<sup>43</sup> Také každý dialog bota, který konverzační designer navrhuje, by měl mít nějaký účel. Zpravidla se jedná

---

<sup>40</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. *Conference on Human Factors in Computing Systems - Proceedings* [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 6.

<sup>41</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 84.

<sup>42</sup> Je možné vyzkoušet v tomto TTS generátoru:

<https://azure.microsoft.com/cs-cz/services/cognitive-services/text-to-speech/#overview>

<sup>43</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. *Conference on Human Factors in Computing Systems - Proceedings* [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 4-6.

o účel transakční, kdy člověk vybízí bota ke konkrétnímu úkolu, případně se jedná o hovor, kdy informace sbírá bot a následně s nimi pracuje. Ačkoliv může být náročnější vytvořit si sociální či emoční vazbu s robotem, existují i roboti s čistě sociálním účelem, například virtuální asistenti, kteří mají za cíl zlepšit život starých lidí, kteří jsou osamělí, nebo boti, kteří mají s uživateli terapeutickou konverzaci.

#### 2.2.5.1.1 Jak to použít v praxi?

Před začátkem návrhu konverzačního stromu je potřeba se zamyslet, jaký cíl má konverzace mít. Cest, jak se do daného cíle může člověk dostat může být několik a konverzace může mít zároveň několik různých konců. Aby byla konzistentní, měl by být jeden hlavní cíl. Pokud je virtuální asistent komerční, je stanovení cílů často navázáno na očekávání daného zadavatele a je předmětem analýzy při ověřování toho, zdali voicebot funguje správně. V takovém případě by celému procesu měla předcházet podrobná business analýza. Prvním krokem před tvorbou flow by mělo být nastínění ideálního průchodu konkrétním scénářem a na základě této happy path se mohou vytvářet další rozvětvení tak, aby odpovídaly stanovenému účelu konverzace.

#### 2.2.5.2 Konzistence a osobnost

Aby mohl být virtuální asistent dobrý partner pro konverzaci, je potřeba aby byl nejen důvěryhodný, ale také konzistentní. V případě, že bude voicebot nekonzistentní, bude uživatel zmatený a nebude ani zdaleka tolik sdílný, jako u virtuálního asistenta, kterému věří.<sup>44</sup> Stejně jako lidé komunikují různě, mohou mít odlišný styl i virtuální asistenti. Proto by měli mít osobnost, která bude jednotná po celou dobu konverzace.

##### 2.2.5.2.1 Jak to použít v praxi?

Má s člověkem mluvit muž, nebo žena? Má mu robotu tykat, nebo vykat? Jedná se o vážný hovor s připomínkou nezaplacené faktury, rozhovor s virtuálním terapeutem, který by měl člověka uklidnit, nebo asistent v aplikaci, který má uživatele primárně bavit? Nastavení osobnosti bota souvisí primárně s cílem konverzace. V případě, že voicebot zastupuje konkrétní společnost, neměl by se jeho styl komunikace odlišovat od té firemní. V praxi je dobré, když si kromě základních charakteristik (pohlaví, výška a rychlost hlasu a podobně) vytvoříme několik různých typů stejných konverzačních značek (časových údajů,

---

<sup>44</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 84.

poděkování, zpětné vazby), které budou v rámci konverzačního stylu, který odpovídá zvolené osobnosti.

## **2.2.6 Přizpůsobit se příjemci**

### **2.2.6.1 Recipient Design**

Představme si situaci, kdy člověk s velmi dobrou znalostí angličtiny komunikuje s jiným, který se jazyk začal teprve učit a nemá příliš velkou slovní zásobu. S největší pravděpodobností se člověk s větší slovní zásobou přizpůsobí levelu znalostí toho druhého. Stejný princip, tedy to, že se přizpůsobíme příjemci, se děje běžně nejen na úrovni porozumění jazyku, ale také ve výběru tématu nebo úrovně detailu, u kterého si myslíme, že bude pro druhou stranu srozumitelná.<sup>45</sup> Obdobně bychom se měli přizpůsobit příjemci také při komunikaci s asistentem.

#### **2.2.6.1.1 Jak to použít v praxi?**

V potaz bychom měli brát to, jestli virtuální asistent mluví s člověkem poprvé, nebo hovoří s uživatelem, kterého již zná. Novým uživatelům můžeme chtít sdělit více instrukcí, ty pravidelné ale nechceme zdržovat jejich opakování. V potaz bychom ale měli brát nejen počet použití, ale také frekvenci. Pokud člověk konverzuje s asistentem sice pravidelně, ale s malou frekvencí (např. jedenkrát měsíčně), možná bude chtít instrukce slyšet znova.<sup>46</sup> Principem human-centered designu je znalost uživatele a přizpůsobení se cílovému publiku. Pokud se v něm nachází uživatelé s různým stupněm expertízy, můžeme jim umožnit průchod scénářem více cestami podle jejich úrovně znalostí procesu. Lidem, kteří proces znají, umožníme projít scénář rychleji než těm, kteří jej prochází poprvé a potřebují více informací.<sup>47</sup> Je také dobré dát uživatelům možnost doptat se.

---

<sup>45</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 32.

<sup>46</sup> PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 47-48.

<sup>47</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 32.

## 2.2.7 Schopnost porozumět a naslouchat

### 2.2.7.1 Naslouchání

Jak bot naslouchá v rámci konverzace? Vezměme si jako příklad prvního chatbota Elizu, který simuloval psychologický rozhovor. Bot zde měl za úkol naslouchat pacientovi a pokládat mu otázky. Pokud se pozorně podíváme na přepis konverzace, je zde jasně vidět, že bot ve své otázce buď pokládá obecnou otázku s žádostí o uvedení příkladu (řádek 4) nebo zopakuje, co slyšel od uživatele a v otázce jen vymění zájmena. Eliza tak pokládala otázky, ze kterých má uživatel dojem, že mu rozumí, aniž by musela porozumět obsahu.<sup>48</sup>

```
1 A: Men are all alike.
2 B: In what way?
3 A: They're always bugging us about something or other.
4 B: Can you think of a specific example?
5 A: Well, my boyfriend made me come here.
6 B: Your boyfriend made you come here?
7 A: He says I'm depressed much of the time.
8 B: I am sorry to hear that you are depressed.
9 A: It's true. I am unhappy.
10 B: Do you think coming here will help you not to be unhappy?
```

*obr. 1 – detail konverzace s chatbotem Eliza<sup>49</sup>*

U současných virtuálních asistentů naslouchání funguje tak, že řeči člověka se bot snaží zjistit intent, na základě kterého se pak rozhodne, jak pokračovat dále, případně zároveň z řeči člověka vytěží nějaká data. Aby k naslouchání vůbec mohlo dojít, je potřeba, aby člověk věděl, že může mluvit a že ho bot poslouchá, což můžeme nazvat jako aktivní naslouchání.

V mezilidské komunikaci je atribut aktivního naslouchání velmi důležitý. V oboustranném interaktivním dialogu je to věnování pozornosti, projevení zájmu a ochota účastnit se dialogu – konverzace musí probíhat oběma směry. Kromě slov to může být i

---

<sup>48</sup> MOORE, Robert J. a Raphael. ARAR. Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 2.

<sup>49</sup> MOORE, Robert J. a Raphael. ARAR. Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 2.

neverbální zpětná vazba, např. oční kontakt.<sup>50</sup> Ve studii, která zkoumala to, co lidé očekávají od konverzace s robotem se zjistilo, že lidé očekávají přesné naslouchání, tedy spíše funkční rovinu naslouchání. Při komunikaci s robotem lidé požadují, aby jim robot porozuměl jasně a rychle, bez nutnosti doptávání se.<sup>51</sup>

#### 2.2.7.1.1 Jak to použít v praxi?

Na aktivní naslouchání lze použít podobné metody, jako používáme v mezilidské konverzaci. Můžeme parafrázovat, co jsme od člověka právě slyšeli, zeptat se, zdali jsme slyšeli správně nebo si potvrdit, zdali je to vše, co chtěl člověk říci.<sup>52</sup> V praxi to pak může vypadat následovně. Pro parafrázování můžeme využít věty jako *“Rozumím, že chcete řešit změnu limitu na kreditní kartě.”*, pro dotaz na to, zdali jsme slyšeli správně můžeme jednoduše přidat *“je to tak?”* na konci věty. Potvrzení, zdali je to vše, co chtěl člověk sdělit, vyřešíme odmlkou poté, co klient domluví, abychom se vyvarovali tomu, že jen udělal pauzu mezi promluvou a následně položíme otázku *“Děkuji, chcete říct ještě něco dalšího?”*

Hlasoví asistenti se oproti chatbotům musí vypořádat také s minimalizací pauz mezi dokončením řeči mluvčího a mezi tím, než bot začne hovořit. S tímto faktem je třeba počítat. Bot musí rozpoznat, že řeč člověka již skončila a může mluvit sám.<sup>53</sup>

Součástí tvorby robota je také jeho vylepšování. Měli bychom mít na paměti, že uživatel může danou věc vyslovit velkým množstvím způsobů, doporučuje také analyzovat uživatelské vstupy v čase a tyto intenty voicebota doučovat.<sup>54</sup> Každou takovou iterací vylepšujeme jeho naslouchání.

---

<sup>50</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. Conference on Human Factors in Computing Systems - Proceedings [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 4-6.

<sup>51</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. Conference on Human Factors in Computing Systems - Proceedings [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 4-6.

<sup>52</sup> STILLMAN, Daniel. Good talk: how to design conversations that matter. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8, str. 88.

<sup>53</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 54-55.

<sup>54</sup> DASGUPTA, Ritwik. Voice User Interface Design Moving from GUI to Mixed Modal Interaction. Berkeley, CA: Apress, 2018, 1 online resource (xi, 108 pages). ISBN 1-4842-4125-8. Dostupné z: doi:10.1007/978-1-4842-4125-7, str. 13.

### 2.2.7.2 Potvrzování a souhlasy

Potvrzení vstupů je důležitou součástí každého dobrého designu virtuálního asistenta. Člověk potvrzením získá pocit, že mu bylo porozuměno a zároveň lze tato předcházet situacím, kdy bot porozuměl špatně a vytváří se zde prostor pro opravu. Pro zamyšlení nad tím, které věci je potřeba potvrzovat nám podle autora pomůže zodpovědět si tyto otázky: Jaké jsou důsledky špatného potvrzení? Jakými způsoby bude systém poskytovat zpětnou vazbu? Jaký typ potvrzení je nejvhodnější?<sup>55</sup>

#### 2.2.7.2.1 Jak to použít v praxi?

Explicitní potvrzení je založeno na požadavku vyslovení souhlasu či nesouhlasu. Například při otázce *“Rozuměl jsem číslo 1 8 5 4. Je to správně?”* V takovém případě chce voicebot od člověka slyšet *“ano”*, *“ne”*, či jiný konkrétní výraz pro vyjádření souhlasu či nesouhlasu. Oproti tomu implicitní potvrzení pracuje s variantou, kdy danou informaci sdělíme v textu tak, aby byla srozumitelná, ale není potřeba ji vysloveně potvrdit. Například *“Dobře, objednávám asistenci na 18:30. Kolik lidí pojede s asistenční službou?”* Pokud člověk po této větě nijak nezareaguje, bere bot tuto informaci za implicitně potvrzenou. Uživatel má však možnost v případě nesouhlasu skočit do řeči, či danou informaci opravit hned v následujícím vstupu.<sup>56</sup>

Cathy Pearl popisuje v knize *Designing Voice User Interfaces: Principles of Conversational Experiences* popisuje pět metod, jak může virtuální asistent pracovat s jistotou. První metodou je třístupňová jistota. Při použití této metody pracujeme s procentuálním stupněm jistoty, se kterým nám systém na rozpoznávání řeči vrátil vstup člověka. Určíme i prahovou hodnotu (trashold), například 55 až 80 % a vše, co je pod touto hodnotou odmítneme jako neporozumění. Bot se tedy na otázku zeptá znovu. To, co spadá mezi uvedená procenta si necháme od člověka explicitně potvrdit. Tvzení, kterým bot rozuměl s jistotou vyšší, než je prahová hodnota, pak stačí potvrdit pouze implicitně.<sup>57</sup>

---

<sup>55</sup> PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.

<sup>56</sup> PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.

<sup>57</sup> PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.



Druhou metodou, kterou autorka v knize zmiňuje je implicitní potvrzení. Při tom využijeme část z řeči klienta v promluvě bota.<sup>58</sup> Například pokud se člověk zeptá: “*Kdy je splatnost faktury?*” bot odpoví: “*Splatnost faktury je 31. ledna.*” místo strohé odpovědi “*31. ledna*”. Pokud bot v první otázce zachytil špatný intent a člověk se ptal s jiným záměrem, je z odpovědi bota jasné, co porozuměl a na co konkrétně odpovídá. Člověk má možnost jej v následujícím konverzačním uzlu opravit.

Třetí metoda je vhodná především pro hlasové asistenty, kteří jsou například součástí domácnosti, nebo vozidla. Spočívá ve vykonání požadované akce bez hlasového potvrzení, že bot rozuměl příkazu. Například na výrok člověka “Rozsviť světla” bot světla rozsvítí světla, ale vykonání akce nepotvrdí hlasem. Tento způsob potvrzení je potřeba dělat jen tam, kde je akce okamžitá a nedochází ke zpoždění. Další variantou potvrzení beze slov, která je vhodná i pro hlasové asistenty na telefonu, může být přehraní zvukového signálu, ze kterého bude zřejmé, že se stala požadovaná akce a bot člověku rozuměl.<sup>59</sup> Může to být například zvukový signál na požadavek vrácení do hlavní nabídky, případně jiné zvuky na pozadí, jako je psaní na klávesnici, které má za cíl vyvolat dojem, že si voicebot zapisuje to, co člověk diktuje.

Čtvrtá metoda, tedy obecné potvrzení se používá se v situacích, ve kterých nepotvrzujeme konkrétní údaj, který člověk řekl, ale dáváme najevo, že rozumíme. Obecné potvrzení můžeme použít například při otevřené otázce, kdy promluva člověka stále pokračuje. (např. “*Hm*”, “*Rozumím*”, “*Povídejte,*”). Pátá metoda je vhodná jen pro multimodální virtuální asistenty, kteří mají zároveň i vizuální podobu, případně obrazovku. Jedná se totiž o vizuální potvrzení souhlasu, které můžeme využít buď samostatně, nebo jím doplnit potvrzení hlasem.<sup>60</sup>

### 2.2.7.3 Porozumění

Schopnost naslouchat se váže se schopností porozumět. Na porozumění se můžeme dívat také z pohledu pospolitosti, vzájemného porozumění. V mezilidské komunikaci je vzájemné porozumění nejen o snaze pochopit, co druhý říká, ale také o tom najít společnou řeč. A ta se lépe hledá, pokud protistranu známe. U komunikace s robotem je porozumění

---

<sup>58</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.

<sup>59</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.

<sup>60</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 24-32.

spíše jednosměrné. Společnou řeč můžeme u komunikace s robotem pojímat jako personalizaci, kdy si virtuální agent pamatuje informace o uživateli, aby konverzaci přizpůsobil. Je to však jen jednosměrný proces, protože je zpravidla pouze ze strany robota a není spoluvytvářený.<sup>61</sup> To, že má bot s člověkem společnou řeč, může také vyjádřit tím, že potvrdí, že rozumí tomu, co člověk právě řekl.<sup>62</sup>

## **2.2.8 Schopnost udržet kontext**

### **2.2.8.1 Kontext**

Věda, která je na pomezí lingvistiky a filosofie a při studiu jazyka zkoumá kontext vět, se nazývá pragmatika. Zkoumá jednotlivé výpovědi z pohledu toho, co jimi chtěl řečník sdělit a v jakém kontextu.<sup>63</sup> Abychom docílili opravdové konverzační úrovně, ne jen na sérii jednoduchých příkazů, měli by si boti být schopni pamatovat si minulost. Klíčové znaky dobré mezilidské komunikace jsou totiž podle autorky knihy *Designing Voice User Interface* hlavně tyto tři: uvědomění si souvislostí (věnování pozornosti člověku a okolí), vzpomínky na předchozí interakce a výměna vhodných otázek.<sup>64</sup> Teorie *Theory of Common Ground*<sup>65</sup> zase říká, že předpokladem toho, aby si dva konverzující rozuměli je to, že musí sdílet stejné znalosti. Právě proto je důležité, aby virtuální asistent o člověku něco věděl. Kontext může voicebot získat analýzou dat z minulých konverzací a pamatováním si dat z probíhající konverzace, v takovém případě mluvíme o kontextu diskurzu.<sup>66</sup> Také je možné využít toho, co o daném člověku lze zjistit ze systému zadavatele v případě, že je bot součástí například centra zákaznické péče. V potaz je potřeba brát také kontext “*času, místa, prostorové*

---

<sup>61</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. Conference on Human Factors in Computing Systems - Proceedings [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 4-6.

<sup>62</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 43.

<sup>63</sup> HARRIS, Randy Allen. Voice interaction design crafting the new conversational speech systems. San Francisco, CA: Morgan Kaufmann Publishers, 2005, 1 online resource (619 p.). ISBN 1-281-01658-6, str. 78.

<sup>64</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 5-8.

<sup>65</sup> CLARK, H. H. a S. E. BRENNAN. Grounding in communication. Perspectives on socially shared cognition. American Psychological Association, 1991, 127-149. Dostupné z: doi:10.1037/10096-006

<sup>66</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 5-8.

*konfigurace místa, kde se uživatel a systém nacházejí, aktuální události a kulturní normy.*<sup>67</sup>  
Například současní chatboti jsou v průměru schopni udržet kontext konverzace z 85 %.<sup>68</sup>

Voicebot by měl být také schopen koreference, tedy schopnosti reagovat na jiné označení téhož. Dané téma by si měl voicebot být schopen zapamatovat a pak jej využít v průběhu hovoru. Například pokud se člověk nejprve zeptá, zdali může zaslat zásilku do zahraničí a následně položí dotaz, jaké jsou její maximální rozměry, měl by virtuální asistent vědět, že se druhý dotaz týká zásilky, aniž by bylo ve větě toto klíčové slovo přímo zmíněno.<sup>69</sup>

V mezilidské komunikaci si často vyměňujeme kontext nevědomky a zároveň i my sami máme někdy problém pochopit kontext. Aby si voicebot informace “zapamatoval”, vytváří si kontextový řetězec, do kterého si postupně sesbírané informace od klienta zařazuje do kontextového rámce. V potaz by měl brát tyto typy kontextu: fyzický kontext (kde se klient právě nachází a co dělá?), emocionální kontext (jak se člověk cítí? je nějaký problém, který může bot vyřešit?) a konverzační kontext (o čem člověk právě mluví? bavíme se stále o tom stejném, nebo se konverzace někam posunula?).<sup>70</sup>

#### 2.2.8.1.1 Jak to použít v praxi?

Pokud umíme předpovědět, proč daný člověk zahajuje konverzaci s virtuálním asistentem, velmi to zvýší zákaznickou zkušenost i pravděpodobnost, že člověk požadavek s voicebotem vyřídí úspěšně. Vezměme si jako příklad situaci, kdy si člověk zablokuje přístup do svého internetového bankovníctví a klikne na telefonní číslo v aplikaci, na kterém je nasazený virtuální asistent. O klientovi v tomto případě víme osobní údaje, známe kontext hovoru a jeho momentální situaci. Můžeme si domyslet, že bude našťvaný nebo nervózní a tomuto předpokladu můžeme konverzaci přizpůsobit. Začátek hovoru tedy může vypadat

---

<sup>67</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 44.

<sup>68</sup> RADZIWIŁL, Nicole a Morgan BENTON. Evaluating Quality of Chatbots and Intelligent Conversational Agents. Software quality professional [online]. Milwaukee: American Society for Quality, 2017, 19(3), 25 [cit. 2022-07-25]. ISSN 1522-0540, str. 11.

<sup>69</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 49-52.

<sup>70</sup> DASGUPTA, Ritwik. Voice User Interface Design Moving from GUI to Mixed Modal Interaction. Berkeley, CA: Apress, 2018, 1 online resource (xi, 108 pages). ISBN 1-4842-4125-8. Dostupné z: doi:10.1007/978-1-4842-4125-7, str. 13.

třeba takto: *“Dobrý den, pane Nováku, jsem tady, abych Vám pomohl. Voláte kvůli zablokovanému heslu?”* Zde zároveň využíváme i princip minimalizace, ke kterému se dostanu později – přestože slovní spojení *“zablokované heslo”* může mít v různých kontextech různé významy a může se týkat většího spektra návazných akcí, klient v této situaci s největší pravděpodobností ví, že voicebot mluví o heslu do internetového bankovníctví.

## **2.2.9 Pracovat s chybou a nepochopením**

### 2.2.9.1 Error Handling

U technologie rozpoznávání řeči musíme vždy počítat s určitou chybovostí. Jak ukazuje studie provedená na 10 000 uživatelů virtuálních asistentů, chyba ve špatném přepisu bývá nejčastějším důvodem nespokojenosti lidí s virtuálními asistenty.<sup>71</sup> Abychom s ní dokázali pracovat, je potřeba znát fakt, že nám rozpoznávače řeči nevrací jen to, co člověk doopravdy řekne, ale připojí k tomu množství interpretací toho, co si daný systém na rozpoznání řeči myslí, že člověk řekl a s jakou pravděpodobností. Tuto pravděpodobnost pojmenujme jako jistotu (confidence) a právě ta je zásadním kritériem pro práci s možnou chybou.<sup>72</sup>

#### 2.2.9.1.1 Jak to použít v praxi?

Je velmi důležité, jak v konverzačním designu budeme řešit chyby, nepochopení, nebo reakci na ticho. V mezilidské komunikaci máme více nástrojů, jak dát najevo nepochopení (například gestem, pohledem a podobně), u virtuálních asistentů však máme možnosti omezené.

Může se stát situace, kdy systém nedetekuje žádnou řeč i přesto, že člověk něco řekl. V takovém případě můžeme člověka upozornit (*“Bohužel jsem nic neslyšel, můžete mi prosím říct vaše jméno a příjmení?”*), případně neříkat nic a dát člověku možnost zopakovat své vyjádření znovu. V prvním případě systém přiznává chybu, druhý případ je přirozenější pro lidskou komunikaci a lidé ho jako chybu nemusí vnímat. Další chybnou situací může být, že systém zachytil zvukový signál, ale nedokázal rozpoznat řeč. Zde se můžeme zachovat

<sup>71</sup> SARIKAYA, Ruhi. The Technology Behind Personal Digital Assistants: An overview of the system architecture and key components. IEEE signal processing magazine [online]. PISCATAWAY: IEEE, 2017, 34(1), 67-81 [cit. 2022-07-25]. ISSN 1053-5888. Dostupné z: doi:10.1109/MSP.2016.2617341, str. 79

<sup>72</sup> PIERACCINI, Roberto. The voice in the machine: building computers that understand speech. Cambridge, Massachusetts: MIT Press, 2012, 1 online resource (355 p.). ISBN 1-280-49915-X, str. 241.

podobně, jako v případě výše. Příklad, kdy voicebot porozuměl, ale udělal špatnou akci, zjistíme zpravidla při zpětné analýze již proběhlých hovorů. Řešením je doplnění kategorií intentů a správných reakcí na ně. Obecně je potřeba s určitou chybovostí při rozpoznávání vždy počítat. Předcházet takovýmto situacím lze lepší analýzou toho, co by člověk mohl říct a zdali se některá klíčová slova neobjevují v jiné kategorii intentů.<sup>73</sup>

Předejít neporozumění se dá také prací s podkladovými daty (například databází SPZ či jmen), které bot v reálném čase porovná s tím, co člověk řekl, případně asistovaným učením rozpoznávače řeči, kterému řekneme, že pokud například při diktování čísel slyším “šéf”, je to “šest”. Pokud si člověk s botem vícekrát neporozumí, je dobré nabídnout mu jiný způsob komunikace, například možnost zadat číslo smlouvy na klávesnici telefonu a potvrdit křížkem, případně přepojení na živého operátora.

Může se také stát, že člověk na druhé straně má vadu řeči či koktá nebo mluví nesrozumitelně (může být například pod vlivem alkoholu), případně nemáme kvalitní nástroj na rozpoznávání řeči. Dokola opakující se omluvy bota, že řeč člověka nezachytil by byly nepříjemné, vyřešit to můžeme třeba tím, že otázky postavíme tak, aby člověk odpovídal ano/ne, případně nabídneme právě přepojení na operátora. Bot by měl zároveň zamezit možnému zacyklení při opakovaném neporozumění, například omezením opakovaného průchodu daným konverzačním uzlem.

### 2.2.9.2 Opravy

Jedním z rysů lidské konverzace je možnost opravit jakékoliv předešlé tvrzení za pochodu, aniž by to výrazně ovlivnilo průběh konverzace. Tyto opravy se dějí v průměru každých 84 sekund.<sup>74</sup> Může se stát, že nezachytíme, co řečník řekl, nebo naopak řečník neporozumí, co jsme řekli my. Přirozená konverzace nabízí možnosti, jak tyto problémy v naslouchání nebo porozumění napravit. V takovém případě má daná oprava prioritu před pokračováním v předchozím tématu. Při dialogu s robotem by měla být oprava možná v jakékoliv části konverzace a mělo by být možné opravit problémy s mluvením, posloucháním a porozuměním. Při designu virtuálních asistentů je potřeba brát v potaz také to, že lidé mají

---

<sup>73</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 41-47.

<sup>74</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 59.

tendenci být efektivní a struční a zpravidla použijí jen tolik slov, kolik je potřeba k porozumění a čekají, zdali to zafungovalo. Pokud ne, tak teprve výrok doplní.<sup>75</sup>

#### 2.2.9.2.1 Jak to použít v praxi?

V mezilidské komunikaci platí, že pokud uděláme chybu, je mnohem lepší přiznat, že jsme chybu udělali, než oznámit, že se chyba stala.<sup>76</sup> Stejně tak to platí pro virtuální asistenty. V praxi můžeme sdělit *“Omlouvám se, teď jsem Vám nerozuměl. Můžete to prosím zopakovat?”*, tedy chybu přiznat. Zároveň je součástí sdělení i prosba o nápravu. Pokud by se porozumění nepovedlo ani napodruhé, je na místě nabídnout jiné řešení, například: *“Dneska není můj den, číslo zásilky jsem nezachytil. Přepojím Vás nyní na mého lidského kolegu, který s Vámi vše vyřeší.”*

Posluchač může přeslechnout mluvčího, který mluví příliš rychle, nebo se nachází ve velmi hlučném prostředí. V takovém případě se nabízí otázka *“Pardon, můžete to říct ještě jednou?”* Další z možných oprav může být přerušování mluvčího při neporozumění či chybnému údaji. V tomto případě chce posluchač znát vysvětlení, nebo si jen potvrdit, co slyšel (*“Říkal jste ve tři hodiny?”*) nebo chybu opravit (*“Jsem Elena, ne Irena”*). Stejně jako v konverzaci s ostatními lidmi, musíme počítat, že stejně instinktivně bude člověk mluvit i s hlasovým asistentem.<sup>77</sup> Nabízí se možnost skákání do řeči, kdy uživatel může vstoupit s opravou přímo do řeči voicebota, čímž ji přeruší. Tato možnost je však velmi citlivá na hlučné prostředí, ve kterém se uživatel může nacházet. Aby skákání do řeči mohlo fungovat, je potřeba, aby voicebot poslouchal řeč člověka zatímco sám mluví. Je zde však riziko, že může zachytit řeč někoho jiného, či hluk chybně vyhodnotit jako reakci člověka a ve své promluvě se zastavit.

---

<sup>75</sup> MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 36-38.

<sup>76</sup> STILLMAN, Daniel. *Good talk: how to design conversations that matter*. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8, str. 139.

<sup>77</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 59.

## 2.2.10 Minimalizovat

### 2.2.10.1 konverzační maxima

V 90. letech přišel britský filosof Paul Grice s kooperativním principem, který říká, že úspěch konverzace závisí na tom, že obě strany spolupracují. Spolu s ním popsal také čtyři pravidla pro úspěšnou komunikaci, které nazval 4 konverzační maxima. Aby konverzace fungovala, je podle Grice potřeba dodržet maximum kvality, maximum kvantity, maximum relevance a maximum způsobu.<sup>78</sup> Jak tato pravidla můžeme aplikovat na svět virtuálních asistentů?

1. Maximum kvality: Bot by měl mluvit pravdivě.
2. Maximum kvantity: Bot by měl sdělit jen tolik informací, kolik je nutné pro porozumění.
3. Maximum relevance: Bot by měl vždy mluvit k věci.
4. Maximum způsobu: Bot by měl mluvit tak, aby mu protistrana rozuměla.

#### 2.2.10.1.1 Maximum kvality a relevance

Kooperační princip i tato čtyři maxima můžeme vnímat jako pokyny pro psaní výpovědí pro virtuální asistenty. Maximum kvantity nám říká, že bychom měli sdělit přesně tolik informací, kolik je potřeba a maximum relevance to, že mají být informace relevantní. To se však může lišit na základě toho, s kým zrovna mluvíme - co je pro někoho příliš informací, může být pro jiného málo, případně mohou být nerelevantní. Konverzaci bychom proto měli přizpůsobit příjemci, případně zvolit zlatou střední cestu.<sup>79</sup>

#### 2.2.10.1.2 Jak to použít v praxi?

V praxi se můžeme setkat s korigováním maxima kvantity i u vstupů člověka. Pokud na otázku *“Jak Vám mohu pomoci?”* odpoví člověk bankovnímu asistentovi *“chci změnit limit”*, může se bot doptat: *“Chcete změnit limit na kartě nebo v internetovém bankovníctví?”* Pokud by odpověděl pouze *“limit”*, může ho bot pobídnout: *“Potřebuji slyšet více informací.”* V případě dlouhé promluvy by mohla výpověď bota znít: *“Prosím, sdělte mi Váš požadavek v jedné větě.”*

---

<sup>78</sup> GRICE, H. P. (H. Paul). *Studies in the way of words*. Cambridge (Mass.): Harvard University Press, 1991, vi, 394 s. ISBN 0-674-85271-0, str. 28.

<sup>79</sup> HARRIS, Randy Allen. *Voice interaction design crafting the new conversational speech systems*. San Francisco, CA: Morgan Kaufmann Publishers, 2005, 1 online resource (619 p.). ISBN 1-281-01658-6, str. 80.

### 2.2.10.1.3 Maximum kvality

Toto maximum se týká hlavně etiky. Fakt, že bude bot sdělovat pouze pravdu a podložená tvrzení velmi souvisí s nastavením důvěrného vztahu.<sup>80</sup> Maximum kvality se také týká správného nastavením očekávání s čím může virtuální asistent pomoci. V opačném případě se totiž může konverzace stát matoucí a frustrující.<sup>81</sup> To by mělo být nastaveno již v úvodu hovoru.<sup>82</sup>

### 2.2.10.1.4 Jak to použít v praxi?

Co se týče dodržení maxima kvality při nastavení očekávání, neměli bychom používat obecnou otázku *“Jak Vám mohu pomoci?”*, pokud bot umí rozeznat pouze jeden, či dva intenty. V případě, že zvládne pouze několik úkolů, můžeme dát člověku vybrat z nabídky: *“Dobrý den, jsem virtuální asistentka Eva z cestovní agentury ABC a umím zarezervovat letenku či hotel. S čím Vám mohu pomoci?”* Případně otázku vynechat a rovnou přejít k výzvě. *“Dobrý den, vítejte v Rozvozové službě. Jsem virtuální asistent a rád Vám řeknu, kde je Vaše zásilka. Nejprve budu potřebovat její číslo. Máte jej prosím u sebe?”*

### 2.2.10.1.5 Maximum způsobu

Tomu, co bot říká, by mělo být rozumět. Proto je dobré se vyhnout nejednoznačným vyjádřením, málo používaným slovům, dvojsmyslům, zkratkám a neplést různé věci dohromady. Komunikovat jasně, přímočaře, konkrétně a uspořádaně.<sup>83</sup>

### 2.2.10.1.6 Jak to použít v praxi?

Nejednoznačné vyjádření může znít například takto: *“Je to všechno nebo chcete ještě něco dodat?”* V případě, že by člověk odpověděl *“ano”*, není zřejmé, co tím člověk myslel a jak bude v tomto případě bot pokračovat. Lepším řešením je zeptat se pouze na jednu z těchto dvou otázek *“Je to všechno?”* případně *“Chcete ještě něco dodat?”*. Uživatelům bychom také měli dát možnost doptat se na to, čemu nerozumí. Pokud například potřebuje bot ke

---

<sup>80</sup> HARRIS, Randy Allen. Voice interaction design crafting the new conversational speech systems. San Francisco, CA: Morgan Kaufmann Publishers, 2005, 1 online resource (619 p.). ISBN 1-281-01658-6, str. 82.

<sup>81</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. xi.

<sup>82</sup> SZYMANSKI, Margaret H. a Robert J. MOORE. Adapting to Customer Initiative: Insights from Human Service Encounters. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 19-32 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_2, str. 31.

<sup>83</sup> HARRIS, Randy Allen. Voice interaction design crafting the new conversational speech systems. San Francisco, CA: Morgan Kaufmann Publishers, 2005, 1 online resource (619 p.). ISBN 1-281-01658-6, str. 82-84.



změně dodavatele energie číslo EAN, může to vypadat následovně: Robot: “*Prosím, nadiktujte mi číslo vašeho odběrného místa, takzvaný kód EAN.*” Člověk: “*Co to je? Kde ho najdu?*” Robot: “*Kód EAN zjistíte z faktury Vašeho současného dodavatele elektřiny. Jedná se o osmnáctimístné číslo, které bývá uvedeno v pravé horní části faktury. Jakmile jej najdete, vyslovte prosím pokračovat.*”

### 2.2.10.2 Teorie kognitivní zátěže

V souvislosti s konverzačními maximy nelze nezmínit také kognitivní zátěž, což je mentální úsilí, které je potřeba vyvinout, abychom udrželi nějakou informaci. Při tvorbě virtuálních asistentů a jejich textů je potřeba myslet na to, že chceme u uživatelů udržet tuto kognitivní zátěž co nejnižší, aby bylo odpovídání rychlé a bez námahy. Komunikace s voicebotem může být na přetížení informacemi zvláště náchylná, protože oproti chatové komunikaci člověk nemůže text znovu projít očima, zároveň se mohou lidé při odpovídání cítit pod tlakem, protože ví, že na odpověď mají omezený čas a mohou mít tendenci odpovídat hned, jakmile uslyší výzvu k akci. Pokud je sdělení přetíženo informacemi, může se stát, že lidé neporozumí všemu, co slyšeli, nebo jim odpověď trvá příliš dlouho, což může voicebot vyhodnotit jako mlčení.<sup>84</sup> Také lidská konverzace má tendenci být co nejvíce efektivní, a to nejen z pohledu stručnosti ale také minimalizací dodatečných dotazů. To přirozeně přizpůsobujeme tomu, s kým mluvíme pomocí rozpoznávacích znaků. Pokud bych mluvila o své spolužačce, kterou protistrana zná, můžu v konverzaci zmínit pouze jméno a bude jasné, o kom mluvím. Pokud by ji však protistrana neznala, přidám ke jménu rozpoznávací znak a pojmenuji ji jako “*spolužačka Kristýna*”, aby bylo zřejmé o koho jde bez potřeby se doptávat na další dotazy.<sup>85</sup>

#### 2.2.10.2.1 Jak to použít v praxi?

Vysoké kognitivní zátěži lze předejít stručností a co nejpřesnějším sdělením, paralelní strukturou při vyjmenovávání (podstatná jména a slovesa by se při vyjmenovávání neměla míchat dohromady), konverzačními značkami, které oddělí jednotlivá vyjádření, správným pořadím slov a myšlenek (to nejdůležitější sdělení nebo výzva k akci by měla být vždy na konci), dodržování pravidla tří, kdy by měl bot vyjmenovat maximálně tři položky a také k

---

<sup>84</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 150-153.

<sup>85</sup> MOORE, Robert J. a Raphael. ARAR. Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0, str. 35-36.

tomu můžeme využít prozodii řeči (pauzy, důraz a podobně).<sup>86</sup> Minimalizovat bychom měli také možné další dotazy pomocí rozpoznávacích znaků.

### 2.2.10.3 Univerzální, ale konkrétní

Maximum kvality a relevance i teorie kognitivní zátěže nám napovídají, že bychom měli dobře zvážit, kterými slovy výpověď bota vyjádříme. Autorky knihy *Conversations with Things* toto pravidlo nazvaly schopností být *“univerzální, ale konkrétní”*.<sup>87</sup> Hlavním cílem je přímočarost a význam by měl být univerzálně srozumitelný, bez dvojsmyslů. Zároveň bychom měli udržet unikátní tón komunikace, který se vztahuje na osobnost bota.<sup>88</sup> I pro člověka by mělo být zřejmé kdy má mluvit a hlavně co může říct, přičemž platí, že by měl bot uživateli sdělit příklad spíše než obecné instrukce, pro lidi je totiž přirozenější věc zopakovat a doplnit si své údaje, než se řídit obecnou instrukcí.<sup>89</sup>

#### 2.2.10.3.1 Jak to použít v praxi?

Obecná instrukce může znít takto: *“Prosím řekněte mi své datum narození ve formátu den, měsíc a rok”*. Pokud ji chceme nahradit konkrétní, můžeme ji přepsat na: *“Prosím řekněte mi své datum narození, například 1. ledna 1991.”*

Když mluvíme s ostatními, máme tendenci zrcadlit jejich postoj, gesta a také styl řeči. Zrcadlení se může týkat výslovnosti, slovníku, hlasitosti nebo třeba akcentu. Na tohle je potřeba myslet i při tvorbě virtuálních asistentů. Pokud bude bot stručný při pokládání otázky, pak s největší pravděpodobností také člověk odpoví stručně. Pokud navrhne bota, který se nejprve široce představí a následně se zeptá na méně určitou otázku, bude mít člověk tendenci se více rozpovídat.<sup>90</sup>

---

<sup>86</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 153-170.

<sup>87</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 134.

<sup>88</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 134.

<sup>89</sup> PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 19-20.

<sup>90</sup> DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 63.

## 2.2.11 Komunikovat lidsky

### 2.2.11.1 Zdvořilost a ohleduplnost

V této sekci budu vycházet z paradigmatu, že na psychosociální jevy jako je zdvořilost, osobnost či lichocení reagují lidé při komunikaci s počítačem stejným způsobem, jako při mezilidské interakci.<sup>91</sup> Definovat zdvořilost může být poměrně náročné. To, jak ji vnímáme, totiž ovlivňuje způsob, jakým jsme byli vychováni a také kultura.

#### 2.2.11.1.1 Jak to použít v praxi?

Obecně jako zdvořilé chování můžeme považovat to, kdy je někdo milý, nápomocný a snaží se vyvarovat konfliktu. Neexistuje žádný univerzální seznam zdvořilostních pravidel a je dobré uvědomit si, že zdvořilost není jen o slovech “*prosím*” a “*děkuji*”, ale také o pozdravu či rozloučení, trpělivosti s uživatelem, vhodnosti vykání a tykání a podobně. Tato pravidla je potřeba nastavit na základě cílové skupiny, která s botem bude komunikovat a jejího vnímání zdvořilosti.<sup>92</sup> Konkrétně rozloučení má ještě další přidanou hodnotu – lidé ví, že je konverzace u konce a mohou zavěsit, či konverzaci jinak ukončit.<sup>93</sup> Co se týče designu bota tak, aby působil ohleduplně, můžeme toho docílit pomalejším tempem konverzace a kladením menšího množství otázek.<sup>94</sup>

### 2.2.11.2 Small talk

Jak uvádím výše, konverzace by měla mít svůj účel, který může být buď sociální nebo transakční. V mezilidské komunikaci je small talk poměrně běžnou záležitostí a odbočení od tématu dotazem na to, kolik je hodin, nebo poznámkou o počasí nejspíše nikoho nepřekvapí. Proto bychom i při tvorbě konverzačního stromu měli myslet na to, že může nastat situace, kdy člověk vybočí z ideální cesty průchodu scénáře a zeptá se na otázku, kterou můžeme zařadit mezi small talk.

---

<sup>91</sup> NIJHOLT, Anton, Oliviero STOCK, Alan DIX a John MORKES. Humor modeling in the interface. CHI '03 Extended Abstracts on human factors in computing systems [online]. ACM, 2003, 1050-1051 [cit. 2022-07-12]. ISBN 9781581136371. Dostupné z: doi:10.1145/765891.766143, str. 1.

<sup>92</sup> DEIBEL, Diana a Rebecca EVANHOE. Conversations with Things: UX Design for Chat and Voice. Rosenfeld Media, 2021. ISBN 1-933820-26-8, str. 73-74.

<sup>93</sup> PEARL, Cathy. Designing voice user interfaces: principles of conversational experiences. O'Reilly Media, 2016. ISBN 978-1-491-95541-3, str. 53.

<sup>94</sup> BENNETT, Gregory A. Conversational Style: Beyond the Nuts and Bolts of Conversation. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 161-180 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_8, str. 175.

#### 2.2.11.2.1 Jak to použít v praxi?

Voicebot při naslouchání v daném konverzačním uzlu ví, na které vstupy uživatele (intenty) by měl zareagovat, ty ostatní spadnou do nerozpoznaných intentů. Vzhledem k tomu, že small talk se může vyskytnout kdykoliv v průběhu celé konverzace, je dobré stanovit si určité “globální intenty”, které robot zaslechne kdykoliv v průběhu celé konverzace. Reakce na tyto intenty by měla odpovídat personě bota a po té, co robot zareaguje, by se měl vrátit zpět do uzlu, ze kterého odbočil. Globální intenty mohou být následující:

1. požadavek na počkání
2. dotaz na čas
3. dotaz na náladu
4. požadavek na vtip
5. dotaz, zdali je virtuální asistent opravdu robot a podobně

#### 2.2.11.3 Humor

Ve studii zaměřené na aspekty dobré komunikace účastníci zmiňovali humor jako prvek který konverzacím přidává další rozměr a v některých případech může být hybatelem konverzace. Zcela zásadní je ale relevance humoru pro danou konverzaci, humor musí zapadat do kontextu. Při otázce na důležitost humoru při interakci s virtuálním asistentem jej účastníci výzkumu popisovali jako pozitivní rys, ne však jako nutně nezbytnou věc, jako u lidské komunikace.<sup>95</sup> Humor v konverzaci nemá pouze zábavnou roli. Buduje důvěru, pomáhá regulovat konverzaci a plní sociální roli. Může být spontánní, ale také použit záměrně a pokud tyto poznatky z mezilidské interakce převedeme do konverzaci mezi člověkem a robotem, tak právě záměrným použitím humoru ve vhodný moment můžeme zabránit frustraci.<sup>96</sup>

##### 2.2.11.3.1 Jak to využít v praxi?

S humorem je potřeba pracovat velmi citlivě. Měl by se odvíjet od cílové skupiny uživatelů, kteří budou s virtuálním asistentem hovořit a také by měl zapadat do osobnosti bota. U robota, jehož účel je vymáhání nesplacených pohledávek budou humorné

---

<sup>95</sup> CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. Conference on Human Factors in Computing Systems - Proceedings [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705, str. 4-6.

<sup>96</sup> NIJHOLT, Anton, Oliviero STOCK, Alan DIX a John MORKES. Humor modeling in the interface. CHI '03 Extended Abstracts on human factors in computing systems [online]. ACM, 2003, 1050-1051 [cit. 2022-07-12]. ISBN 9781581136371. Dostupné z: doi:10.1145/765891.766143, str. 1

poznámky působit těžkopádně, až ironicky a mohly by uživatele naštvat. Tam, kde nám to persona bota dovolí, můžeme humorem snižovat frustraci při více nerozpoznaných intencích za sebou: *"Jejda, dneska mi to ale vůbec nejde. Raději vás přepojím na živého kolegu."* Pokud v konverzaci očekáváme vtip z druhé strany, tedy od člověka a dokážeme jej správně vyhodnotit, můžeme pro onu podporu humoru využít zvukovou nahrávku smíchu, která bude následovat jako reakce na onu vtipnou poznámku.

### 3 Metodologie

Pro tuto diplomovou práci jsem jako metodu výzkumu zvolila kvalitativní analýzu. Ta se dá charakterizovat tvrzením, že se jejích výsledků nedosahuje pomocí kvantifikace, tedy jakoukoliv statistickou metodou.<sup>97</sup> V literární rešerši jsem nejprve použila metodu syntézy, ve které jde o složení jednotlivých částí do celku a o popsání principů či charakteristik, které dané části spojuje.<sup>98</sup> Pomocí syntézy jsem popsala devět schopností, které by měl virtuální asistent mít, aby byl schopen komunikace s lidmi.

Na základě těchto teoretických poznatků budu v analytické části interpretovat empirická data reálných hovorů. V této části využiji metodu indukce, kdy *“z pravidelnosti zkoumaných událostí odvozujeme obecné pravidlo o určité pravidelnosti platné pro další události na jiném místě nebo v jiném čase.”*<sup>99</sup> V analyzovaných scénářích a hovorech budu hledat situace, které potvrzují či vyvrací poznatky, které jsem získala při tvorbě literární rešerše o zkoumaném tématu. Výzkumný proces, kterým budu data zkoumat, bude popisný. Tento proces umožňuje vytvořit konkrétní obraz ze *“specifických podrobností situace, jevu nebo vztahů.”*<sup>100</sup> Nejprve však popíšu cíl výzkumu, výzkumnou otázku a hypotézy.

#### 3.1 Cíl práce a výzkumná otázka

Cílem práce je vytvořit model pravidel konverzačního designu. K tomu nám bude sloužit teoretické ukotvení (poznatky z odborné literatury) a následné ověření poznatků na příkladu scénářů virtuálních asistentů a reálných hovorů s lidmi. S cílem práce je úzce spjata také výzkumná otázka, která zní:

1. Jaká jsou pravidla konverzačního designu, aby byla komunikace lidí s roboty co nejvíce přirozená?

---

<sup>97</sup> STRAUSS, Anselm L., Stanislav JEŽEK a Juliet M. CORBIN. Základy kvalitativního výzkumu: postupy a techniky metody zakotvené teorie. Boskovice: Albert, 1999, 196 stran : tabulky. ISBN 80-85834-60-X, str. 10.

<sup>98</sup> HENDL, Jan. Kvalitativní výzkum: základní teorie, metody a aplikace. Čtvrté, přepracované a rozšířené vydání. Praha: Portál, 2016, 437 stran : ilustrace ; 22 cm. ISBN 978-80-262-0982-9, str. 35.

<sup>99</sup> HENDL, Jan. Kvalitativní výzkum: základní teorie, metody a aplikace. Čtvrté, přepracované a rozšířené vydání. Praha: Portál, 2016, 437 stran : ilustrace ; 22 cm. ISBN 978-80-262-0982-9, str. 36.

<sup>100</sup> HENDL, Jan. Kvalitativní výzkum: základní teorie, metody a aplikace. Čtvrté, přepracované a rozšířené vydání. Praha: Portál, 2016, 437 stran : ilustrace ; 22 cm. ISBN 978-80-262-0982-9, str. 38.

Hypotéza k této výzkumné otázce je následující:

1. Existuje soubor pravidel, který je potřeba dodržet při tvorbě virtuálního asistenta. Tato pravidla jsou čistě konverzačního charakteru a vychází z konverzační analýzy a z chování lidí v mezilidské komunikaci.
2. K této výzkumné otázce se vážou tyto dvě podotázky.
  - a. Jsou lidé schopni komunikovat s robotem?
  - b. Jsou lidé ochotni komunikovat s robotem?

Hypotézy k těmto otázkám jsou následující:

1. Lidé jsou schopni plnohodnotné komunikace s virtuálním asistentem.
2. Lidé jsou ochotni vyřídit svůj požadavek s robotem.

## **3.2 Metody sběru dat**

### **3.2.1 Výběr zkoumaného vzorku**

Ke kvalitativní analýze jsem vybrala 3 scénáře a hovory, ve kterých lidé s voicebotem mluví po telefonu. Telefonní hovory jsou ideálními daty pro analýzu toho, jaké vzorce lidé při konverzaci využívají. Telefonická interakce je totiž oprostěna od neverbální komunikace a je omezená pouze na využití řeči – gramatiky, lexikonu a prozodie.<sup>101</sup>

K analýze jsem využila hovory, které proběhly skrze platformu VOCALLS, která nabízí řešení automatizace hlasové komunikace. V rámci webové aplikace je možné poslechnout si jak samotnou nahrávku hovoru, tak také zobrazit textový přepis a metadata o samotném hovoru, jako jsou vytěžená data, číselně vyjádřená jistota, na základě které se bot v daném uzlu rozhodnul zareagovat, doba promluvy, doba latence a podobně. Právě textové přepisy jsou ideální pro analýzu, protože jsou očištěny od hluku okolí a je možné se rychle zorientovat v konverzaci. Jsou také ideální k využití pro ilustraci dané situace v diplomové práci, jelikož jsou textového charakteru. Při analýze jsem využila i možnost si hovor poslechnout z uložené nahrávky, abych například zjistila, zda nedošlo k chybě, kdy člověk

---

<sup>101</sup> SZYMANSKI, Margaret H. a Robert J. MOORE. Adapting to Customer Initiative: Insights from Human Service Encounters. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 19-32 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_2, str. 20.

mluvil, ale voicebot jeho řeč nezachytil. Kromě hovorů jsem analyzovala také samotné scénáře. K analýze jsem zvolila ty, které jsou po technické stránce plně funkční a ve kterých lidé s voicebotem pravidelně komunikují či komunikovali.

### **3.2.2 Etické otázky výzkumu a práce s daty**

Veškeré hovory, ve kterých se vyskytují osobní údaje (jména, rodná čísla, čísla bankovních účtů a podobně) jsou pro účely této diplomové práce anonymizovány. Skrze data není možné identifikovat žádného klienta, který s virtuálním asistentem hovořil. Neuvádím ani názvy společností, ve kterých jsou daní virtuální asistenti nasazeni. V popisu scénáře používám pouze obor, ve kterém podnikají.

### **3.2.3 Pomocné otázky k analýze**

K jednotlivým dovednostem jsem vytvořila set otázek, na které se zaměřuji při analýze.

1. Schopnost vést konverzaci
  - a. Provádí bot člověka hovorem? Co k tomu využívá?
  - b. Jaké konverzační značky využívá?
  - c. Dochází v hovorech k vyšší latenci?
2. Vybudovat důvěru
  - a. Je bot transparentní? Je zřejmé, že člověk hovoří s virtuálním asistentem a za jakou společnost bot vystupuje?
  - b. Je při sdělování dat potřeba vyslovit celý údaj?
  - c. Informuje bot o způsobu zabezpečení dat?
3. Být konzistentní
  - a. Jaký je účel konverzace? Transakční nebo sociální?
  - b. Zmiňuje bot na začátku konverzace, jaký je její cíl?
  - c. Komunikuje voicebot konzistentně? Jakou má osobnost?
4. Přizpůsobit se příjemci
  - a. Přizpůsobuje se bot stupni expertízy klienta?
  - b. Má klient možnost doptat se na detail?
5. Schopnost porozumět a naslouchat
  - a. Jakým způsobem bot naslouchá a je to pro člověka zřejmé, že porozuměl?
  - b. Jsou ve scénáři globální intenty?



- c. Jaký typ potvrzování bot využívá?
- 6. Schopnost udržet kontext
  - a. Jaké má bot o volajícím informace a jak s nimi pracuje?
  - b. Jak pracuje s kontextem?
- 7. Pracovat s chybou a nepochopením
  - a. Jak bot pracuje s chybou a nepochopením?
  - b. Jaký typ dat bot od člověka získává a co se stane, pokud bot nepochopí?
- 8. Minimalizovat
  - a. Komunikuje bot v rámci konverzačních maxim? Komunikují lidé v rámci maxim nebo je musí bot korigovat?
  - b. Dává někde člověku vybrat z nabídky? Jsou někde nejednoznačná vyjádření?
- 9. Komunikovat lidsky
  - a. Jakým stylem bot komunikuje s klientem?
  - b. Je zdvořilý? Je zde možnost small talku? Využívá humor?

### 3.3 Zkoumaný vzorek

#### 3.3.1 Rozsah zkoumaného vzorku a diverzifikace dat

U každého scénáře jsem si prošla konkrétní konverzační strom, ideální cestu i veškeré rozvětvení scénáře. Zároveň jsem analyzovala třicet hovorů náhodně vybraných z množiny posledních 140 proběhlých hovorů ke každému scénáři. Analyzovala jsem jak hovory, které byly označeny příznakem “*listened to end*”, ve kterých bot s klientem došel do konce konverzace a zavěsil, tak také hovory, které měly příznak “*picked*”, které ukončil v průběhu hovoru člověk.

Conversation log							Search	Q	⌂
Phone	State	Time	Direction	Duration	Actions				
[redacted]	picked	27. 7. 2022 10:51	IN	3min 4s					
[redacted]	listened to end	27. 7. 2022 10:02	IN	3min 26s					
[redacted]	picked	27. 7. 2022 10:00	IN	15s					
[redacted]	listened to end	27. 7. 2022 09:49	IN	1min 55s					

obr. 2 – detail aplikace

Vzhledem k tomu, že k datům přistupuji skrze jedinou platformu, uvědomuji si riziko zkeslení výsledků analýzy a proto zkoumaná data co nejvíce diverzifikuji. A to konkrétně využitím hovorů s různými syntetickými hlasy, odlišnými enginey pro převod řeči na text, či výběrem hovorů s různým kontextem. Tento detail ke každému scénáři uvádím níže.

### **3.3.2 Scénář 1: Zadání platby k úhradě**

téma scénáře	zadání platby k úhradě
obor klienta	bankovníctví
STT engine	NEWTON Technologies, Google
TTS voice name	cs-CZ-AntoninNeural (Microsoft)
typ hovoru	příchozí (inbound)
ID scénáře	771

### **3.3.3 Scénář 2: Virtuální zdravotní sestra**

téma scénáře	virtuální zdravotní sestra
obor klienta	zdravotnictví
STT engine	Google
TTS voice name	cs-CZ-VlastaNeural (Microsoft)
typ hovoru	příchozí (inbound)
ID scénáře	1551

### **3.3.4 Scénář 3: Vymáhání pohledávek**

téma scénáře	vymáhání pohledávek
obor klienta	věřitelé
STT engine	Google, Microsoft
TTS voice name	cs-CZ-AntoninNeural (Microsoft)

typ hovoru	odchozí (outbound)
ID scénáře	872

### 3.4 Limity výzkumu

Na tomto místě bych ráda uvedla limity výzkumu, které mohou ovlivnit jeho výsledky. Definice limitů může zároveň sloužit jako podklad pro další výzkum v této oblasti.

Hlavním limitem tohoto výzkumu je fakt, že analyzuji dovednosti voicebota pouze na základě nahrávek a přepisu textů. Pro plné pochopení toho, zdali virtuální asistent tuto schopnost opravdu má, by bylo vhodné analyzovat také zpětnou vazbu konkrétních lidí, kteří v daných hovorech s asistentem hovoří. Jsem si vědoma toho, že některé schopnosti a dovednosti mají převážně funkční charakter (například schopnost pracovat s chybou či naslouchat) a je možné jejich funkčnost ověřit skrze přepis hovoru či nahrávku. Nedílnou součástí dalších dovedností je však i subjektivní pocit konkrétního uživatele, například u schopnosti vybudovat důvěru či dovednosti komunikovat lidsky.

Vzhledem k tomu, že jsem si zvolila scénáře, které jsou plně funkční a ve kterých lidé již komunikovali či stále komunikují, je pravděpodobnější, že v nich bot již má konkrétní dovednosti, a proto je možné, že je v analýze spíše potvrdím, než vyvrátím. Pokud bych analyzovala scénáře, které jsou teprve ve fázi vývoje, lépe by se na nich interpretovaly situace, kdy s voicebotem lidé nebyli schopni komunikovat právě z toho důvodu, že by mu nějaká ze schopností chyběla.

Posledním limitem výzkumu může být také omezení vybraného vzorku pouze na konverzace, které proběhly přes telefon. Jsem si vědoma toho, že v otázce cíle konverzace bude mít většina z nich spíše transakční charakter, než sociální.

## 4 Analytická část

### 4.1 Analýza scénářů

V této části diplomové práce budu analyzovat vybrané scénáře. U každého z nich nejprve popíšu, jak probíhá ideální cesta scénářem, jaký má cíl a proč byl vytvořen. Následně analyzuji scénář i proběhlé hovory za pomoci sady otázek, které zkoumají danou dovednost či schopnost virtuálního asistenta. Součástí analýzy jsou výřezy z konkrétních konverzací, na kterých lze interpretovat poznatky získané v teoretické části práce, ale i konkrétní konverzační uzly ze scénáře.

#### 4.1.1 Scénář 1: Zadání platby k úhradě

##### 4.1.1.1 Popis scénáře

Tento scénář byl vytvořen pro automatizaci hlasové komunikace v bance pro typy rutinních hovorů, ve kterých klienti volají do banky, aby zadali jednorázový příkaz k úhradě. Jedná se o úkon, na který není potřeba lidský operátor, proto jej člověk zvládne vyřídit s virtuálním asistentem. Scénář je poměrně komplexní a bot provádí klienta od výběru typu účtu, ze kterého chce platbu poslat, po zadání všech potřebných či volitelných údajů, jako je číslo účtu příjemce, kód banky, částky, data splatnosti, variabilního, konstantního a specifického symbolu a zprávy pro příjemce či odesílatele. Na závěr vytěžená data rekapituluje a požaduje explicitní souhlas *“ano souhlasím”*, aby platbu zadal k úhradě. Následně nabízí možnost zadat další platbu a pokud klient o další platbu nemá zájem, voicebot se loučí a scénář končí.

##### 4.1.1.2 Schopnost vést konverzaci

Voicebot provádí klienta procesem krok po kroku. Využívá k tomu otázky s nabídkou (*“Vidím, že u nás máte dva účty, běžný a spořicí. Ze kterého si přejete provést úhradu?”*), otázky ano/ne, kvalifikační otázky i instrukce (*“Nyní mi sdělte kód nebo název banky”*). Lze tedy jasně rozpoznat, kdy skončila řeč bota a mluvčím se stává člověk. Voicebot klienta neinformuje, v jaké části se klient zrovna nachází a nemusí být zřejmé, kolik otázek ještě do konce procesu zbývá. Virtuální asistent využívá konverzační značky typu časové údaje (*“Nejprve mi prosím řekněte, zda číslo účtu obsahuje předčíslí,”*) poděkování a potvrzení.

23.04.2021 12:48:16: Robot voice - "Je to správně?"

23.04.2021 12:48:20: Human voice - "ano"

23.04.2021 12:48:21: Robot voice - "Děkuji za potvrzení."

23.04.2021 12:48:23: Robot voice - "Nyní mi prosím sdělte částku, kterou si přejete poslat."

*obr. 3 – detail hovoru (scénář 1)*

Ke zvýšené latenci ve scénáři nedochází, protože bot žádná data neporovnává s databází. Při výběru, ze kterého účtu chce klient platbu poslat se voicebot neptá na číslo účtu, ale naopak přichází se sdělením, které typy účtů u klienta vidí a stačí mu odpověď “běžný”, “*spořicí*”, případně souhlas či nesouhlas, pokud má účet pouze jeden. Číslo účtu tak není potřeba porovnat s databází. Stejně tak vytěžené údaje zaslá voicebot do systému banky při skončení hovoru skrze API, ani zde proto není riziko latence. Při poslechu nahrávek z analyzovaných hovorů jsem žádné větší pauzy mezi promluvami nezaznamenala.

#### 4.1.1.3 Vybudovat důvěru

Voicebot se na začátku představuje jako virtuální bankéř, nezmiňuje však konkrétní společnost, ani zde není možnost se na ni doptat. Vzhledem k tomu, že se jedná o příchozí hovor, předpokládá se, že klient je obeznámen s tím, kam volá. Ze stejného důvodu bot neinformuje o zabezpečení dat, ani o tom, že je hovor nahráván. Voicebot v tomto případě neklade velký důraz na vytvoření důvěry hned na začátku hovoru. Jeho způsob komunikace se však nedá označit ani jako nedůvěryhodný. Bot komunikuje jednoznačně, jasný je i cíl konverzace. Dá se tedy říci, že důvěru virtuální asistent nemusel v hovorech budovat, protože vychází z předpokladu, že ji k němu uživatelé již vybudovanou mají a v analyzovaných hovorech jsem nezaznamenala situaci, ve které by bot působil nedůvěryhodně. Co se týče sdělování údajů, je vždy potřeba vyslovit celý údaj, což je s ohledem na proces zadání platby logické.

#### 4.1.1.4 Být konzistentní

Účel konverzace je transakční. Na začátku konverzace bot zmiňuje, že umí pouze zadat jednorázovou platbu, čímž vysvětluje i cíl konverzace. Bot komunikuje seriózně, jak je pro bankovní prostředí typické, a konzistentně. Nepoužívá hovorové výrazy, ale sám jim rozumí.

**29.04.2021 21:30:47:** Robot voice - "Nyní mi sdělte kód nebo název banky."

**29.04.2021 21:30:52:** Human voice - "komerčka"

**29.04.2021 21:30:54:** Robot voice - "Děkuji"

**29.04.2021 21:30:55:** Robot voice - "Rozumím správně, že kód banky je 0 - 1 - 0 - 0, -- to znamená komerční banka, je to tak?"

*obr. 4 – detail hovoru (scénář 1)*

#### 4.1.1.5 Přizpůsobit se příjemci

V tomto scénáři nejsou stanoveny odlišné průchody dle stupně expertízy, je však možnost doptat se, pokud klient něčemu nerozumí, a to konkrétně u předčíslení a zprávy pro příjemce a odesílatele.

**09.04.2021 15:25:34:** Robot voice - "Nejprve mi prosím řekněte, zda číslo účtu obsahuje předčíslení, to znamená číslo před pomlčkou."

**09.04.2021 15:25:41:** Human voice - "to nevím"

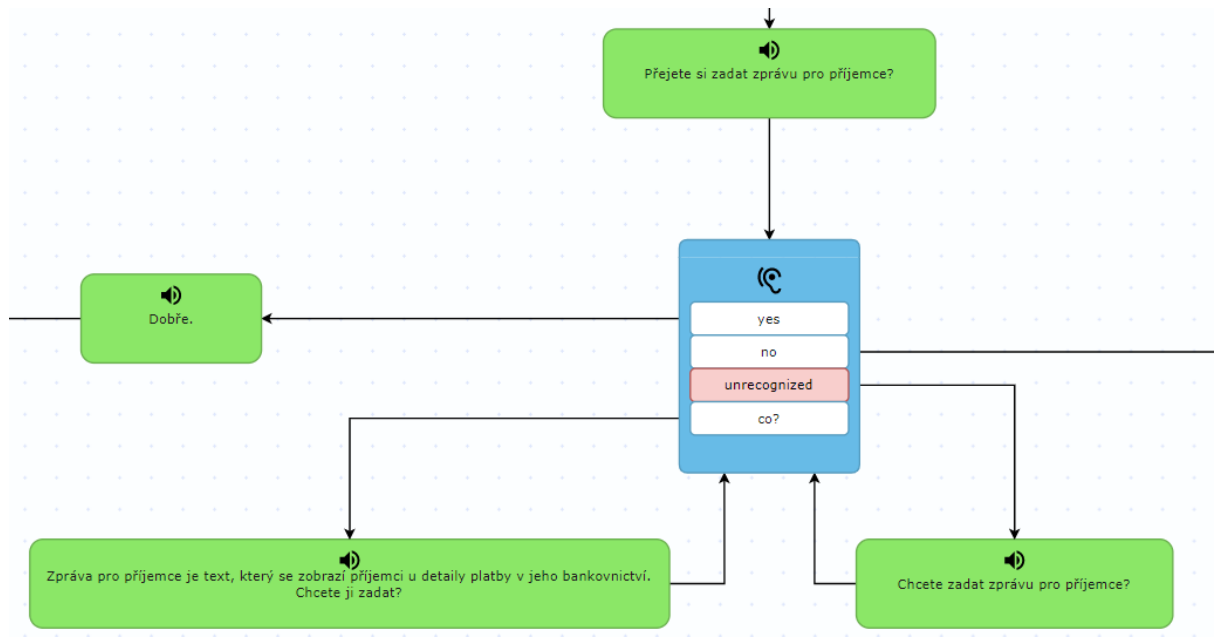
**09.04.2021 15:25:43:** Robot voice - "Předčíslení je část čísla účtu před pomlčkou. Pokud účet pomlčku nemá, znamená to, že nemá ani předčíslení. - Má tedy účet příjemce předčíslení?"

**09.04.2021 15:26:03:** Human voice - "Ne nemám"

**09.04.2021 15:26:05:** Robot voice - "Rozumím."

*obr. 5 – detail hovoru (scénář 1)*

Z analyzovaných hovorů vyšlo, že lidé v některých případech nevěděli, co je to předčíslení, naopak na to, co je zpráva pro příjemce či odesílatele se neptal žádný z volajících. Proto zde pro vysvětlení přidávám výřez z konverzačního stromu, na kterém je vidět konkrétní uzel a reakce voicebota. V zelených buňkách se nachází to, co bot říká, modrá buňka naopak značí naslouchání a zobrazuje kategorie intencí, na které bot dokáže reagovat. V každé kategorii jsou pak konkrétní klíčová slova a gramatiky (soubor klíčových slov vyjadřující stejný intent).



obr. 6 – detail konverzačního stromu (scénář 1)

Voicebot se přizpůsobuje příjemci také při instrukci k diktování zprávy, která je vidět v hovoru níže. Bot zmiňuje, že zpráva musí být krátká, nesdělí však maximální délku. Tato informace zazní až v případě, že by člověk při diktování objem slov překročil.

19.04.2021 19:26:29: Robot voice - "Myslete na to, že zpráva musí být krátká. - Můžete diktovat."

19.04.2021 19:26:37: Human voice - "[REDACTED]  
[REDACTED]  
[REDACTED]"

19.04.2021 19:26:49: Robot voice - "Je mi líto, ale zpráva je příliš dlouhá. Může obsahovat zhruba 20 slov. Zkuste ji prosím přeformulovat a o něco zkrátit. Můžete diktovat."

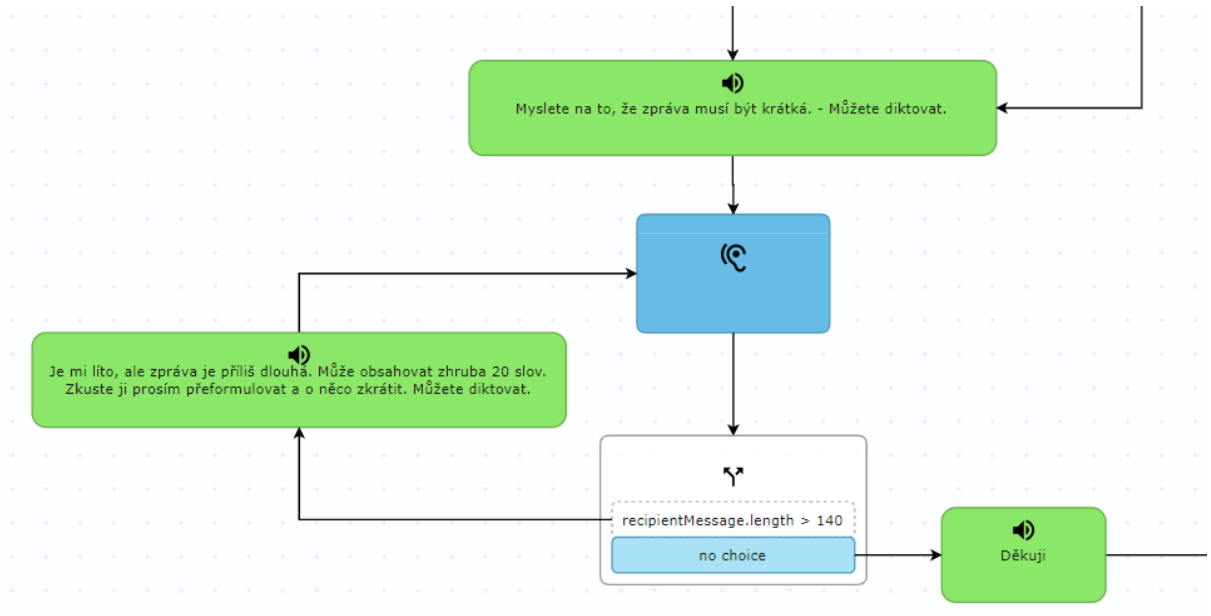
19.04.2021 19:27:05: Human voice - "[REDACTED]"

19.04.2021 19:27:09: Robot voice - "Děkuji"

obr. 7 – detail hovoru (scénář 1)

Zpráva, kterou lze vložit k odeslané platbě, může mít maximálně 140 znaků. Pro běžného člověka je však obtížné odhadnout, kolik slov se do 140 znaků vejde. Aby člověk sdělení porozuměl bez přemýšlení, bot napoví, že se jedná zhruba o 20 slov, což je lépe představitelné. Na obrázku níže je pak vidět konverzační logika k tomuto uzlu. Jakákoliv řeč

člověka, která je zachycena v modré STT buňce je uložena do proměnné a bot pomocí javascriptové funkce porovnává počet znaků.



obr. 8 – detail konverzačního stromu (scénář 1)

#### 4.1.1.6 Schopnost porozumět a naslouchat

Ve všech případech bot využívá explicitní souhlas a s člověkem si potvrzuje každý zadaný údaj. Ve scénáři se také nachází globální intent pro zrušení platby. Člověk žádost o zrušení celého procesu může vyslovit kdykoliv v průběhu scénáře, nezávisle na tom, v jakém uzlu se právě nachází. Stejně tak bot reaguje na žádost o přepojení na operátora, prosbu o počkání (lze vidět na obrázku č. 13) a dotaz na zopakování.

#### 4.1.1.7 Schopnost udržet kontext

Bot s největší pravděpodobností zná jméno klienta, nikde s ním však nepracuje. Má ale informace o typech účtu, které má klient u dané banky otevřené. S touto informací pracuje hned na úvod hovoru, kdy se jej ptá, z jakého účtu chce platbu odeslat. Pracuje také s kontextem vytěžených dat v průběhu hovoru a při závěrečné rekapitulaci ověřuje pouze ty údaje, které si klient přál zadat. Při tom využívá javascriptové funkce, kdy v případě, že je proměnná, do které si bot uložil daný údaj prázdná (např. variabilní symbol = {VS}), přeskočí ji a pokračuje ve čtení další proměnné, aby řeč působila jednoduše.



```

if (VS != "") {
|   recapitulationMessage += format(" - Variabilní symbol je - $1.", interleave(VS, ' - '));
}

if (KS != "") {
|   recapitulationMessage += format(" - Konstantní symbol je - $1.", interleave(KS, ' - '));
}

if (SS != "") {
|   recapitulationMessage += format(" - Specifický symbol je - $1.", interleave(SS, ' - '));
}

```

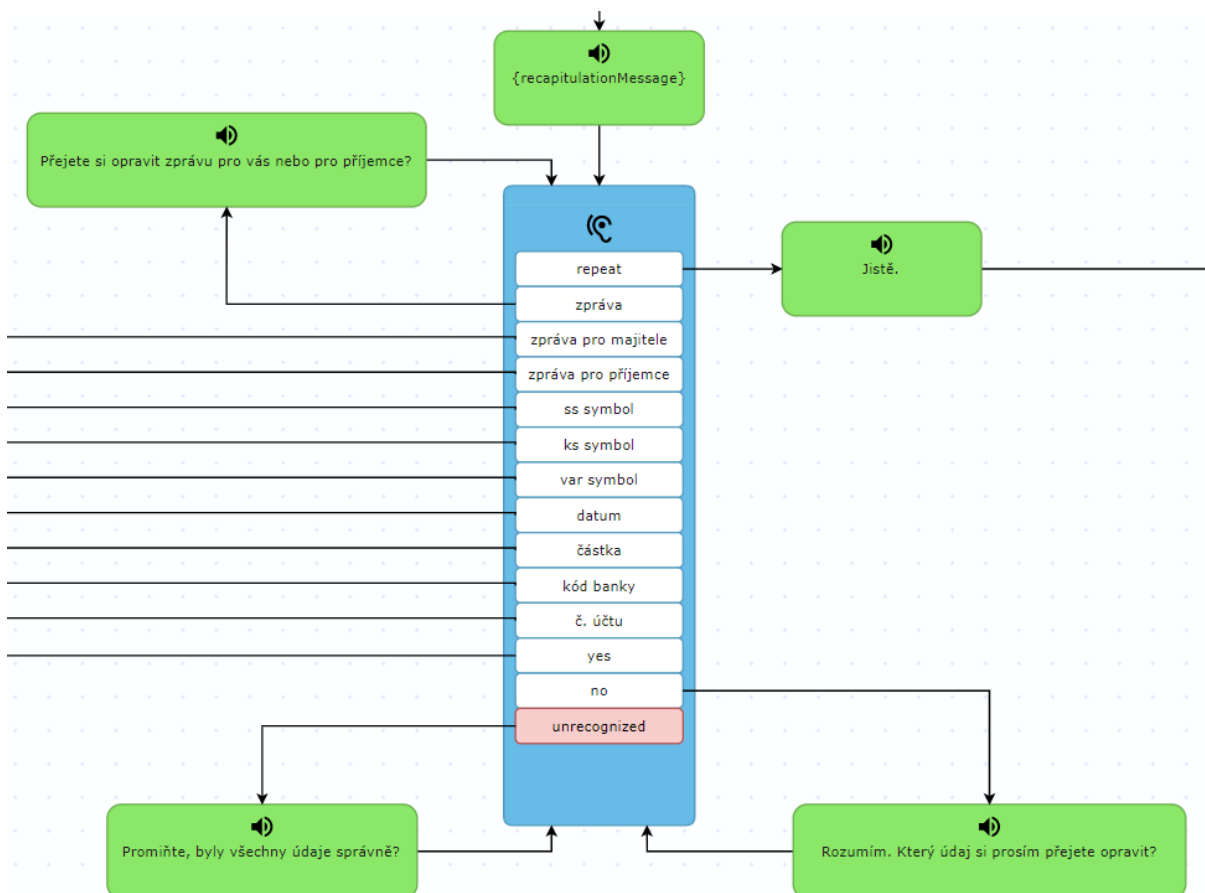
obr. 9 – detail skriptu (scénář 1)

Voicebot nijak nezohledňuje situační či emocionální kontext klienta.

#### 4.1.1.8 Pracovat s chybou a neporozuměním

Při tomto procesu je kladen velký důraz na přesné porozumění, proto si každý ze zadaných údajů bot s člověkem potvrzuje a v případě neporozumění nabídne možnost opravy. Na konci procesu před závěrečným potvrzením zadání platby voicebot rekapituluje veškeré zadané údaje, které si v průběhu hovoru uložil do proměnných. V tomto případě je celá odpověď naskriptovaná a vložena do proměnné *{recapitulationMessage}*.

Zatímco voicebot diktuje, stále klienta poslouchá, přičemž reaguje na konkrétní intenty, jako je variabilní symbol, číslo účtu, zpráva a podobně. Zároveň zachycuje slova obsažená v gramatice ano či ne, tedy slova, která vyjadřují souhlas či nesouhlas. Člověk tak může skočit do řeči kdykoliv v průběhu promluvy bota. Konkrétní údaje k opravě mají v tomto případě větší váhu, než samotný nesouhlas, proto v případě, že člověk řekne “*Ne, konstantní symbol je špatně.*”, převáží slovo “*konstantní*” slovo “*ne*” a voicebot zareaguje nabídkou opravit právě konstantní symbol.



obr. 10 – detail konverzačního stromu (scénář 1)

V hovoru na obrázku č. 11 chtěl člověk opravit zprávu pro příjemce. Engine na rozpoznávání řeči nezachytil slovo “zprávu”, případně ji člověk vůbec nevyslovil. Aby bot zvládnul zařadit intent, stačilo mu k tomu pouze klíčové slovo “pro příjemce”, na základě kterého rozpoznal záměr klienta.

**21.04.2021 15:26:48:** Robot voice - "Odesíláme platbu [ ] korun z Vašeho spořicího účtu. - Číslo účtu příjemce je - [ ] - pomlčka - [ ] - lomeno - [ ] - Variabilní symbol je - [ ] - Specifický symbol je - 1 - 3. - Zpráva pro příjemce je - [ ] [ ] - Zpráva pro vás je - [ ] - -- Bylo vše v pořádku?"

**21.04.2021 15:27:32:** Human voice - "ne"

**21.04.2021 15:27:34:** Robot voice - "Rozumím. Který údaj si prosím přejete opravit?"

**21.04.2021 15:27:40:** Human voice - "pro příjemce"

**21.04.2021 15:27:42:** Robot voice - "Rozumím. Pojdme tedy zprávu upravit."

**21.04.2021 15:27:45:** Robot voice - "Myslete na to, že zpráva musí být krátká. - Můžete diktovat."

*obr. 11 – detail hovoru (scénář 1)*

Ve dalším hovoru byla rekapitulace vytěžených dat v pořádku, člověk se ale rozhodl celou transakci zrušit. Tento příklad jsem zde dala záměrně, protože jde na něm vidět nedokonalost technologie při rozpoznávání řeči. STT engine totiž nesprávně přepsal klientovo "Ne, chci to zrušit" na "nechci to zrušit". Úloha konverzačního designu je přemýšlet i nad těmito situacemi a podle toho scénář tvořit. Bot se v tomto případě správně zeptal, zdali si klient přeje platbu zrušit, ten to potvrdil.

**26.04.2021 13:08:58:** Robot voice - "Nyní spolu raději všechny platební údaje zkontrolujeme. Pozorně poslouchejte, ať nám nic neunikne."

**26.04.2021 13:09:07:** Robot voice - "Odesíláme platbu [REDACTED] korun a [REDACTED] haléřů z Vašeho běžného účtu. - Číslo účtu příjemce je [REDACTED] - lomeno - [REDACTED] - Zpráva pro příjemce je - [REDACTED] . -- Bylo vše v pořádku?"

**26.04.2021 13:09:46:** Human voice - "Ano bylo"

**26.04.2021 13:09:48:** Robot voice - "Děkuji. Potřebuji ještě vaše závěrečné potvrzení. Pokud souhlasíte se zadáním platby k úhradě, vyslovte prosím "ano souhlasím"."

**26.04.2021 13:10:01:** Human voice - "nechci to zrušit"

**26.04.2021 13:10:04:** Robot voice - "Přejete si zadání platby zrušit?"

**26.04.2021 13:10:07:** Human voice - "ano"

*obr. 12 – detail hovoru (scénář 1)*

Na dalším z analyzovaných hovorů je možné poukázat na situaci, kdy bot klientovi neporozuměl. Nejednalo se však o chybu technologie, která by nesprávně přepsala klientovu řeč. Zde můžeme vidět pochybení konverzačního designu, kdy při tvorbě scénáře nebylo slovo “dobrý” zařazeno do gramatiky souhlasu. Voicebot jednoduše nevěděl, jak má na řeč člověka reagovat a proto se na otázku zeptal znovu.

**22.04.2021 9:39:04:** Robot voice - "Odesíláme platbu [redacted] korun z Vašeho spořicího účtu. - Číslo účtu příjemce je - [redacted] - lomeno - [redacted]. - Variabilní symbol je - [redacted]. - Konstantní symbol je - [redacted] - Specifický symbol je - [redacted] - Zpráva pro příjemce je - [redacted]. - Zpráva pro vás je - [redacted]. -- Bylo vše v pořádku?"

**22.04.2021 9:39:10:** Human voice - "moment"

**22.04.2021 9:39:11:** Robot voice - "Dobře, počkám. Až budete chtít pokračovat, vyslovte prosím pokračovat."

**22.04.2021 9:39:19:** Human voice - "pokračovat"

**22.04.2021 9:39:21:** Robot voice - "Odesíláme platbu [redacted] korun z Vašeho spořicího účtu. - Číslo účtu příjemce je - [redacted] - lomeno - [redacted]. - Variabilní symbol je - [redacted]. - Konstantní symbol je - [redacted] - Specifický symbol je - [redacted] - Zpráva pro příjemce je - [redacted]. - Zpráva pro vás je - [redacted]. -- Bylo vše v pořádku?"

**22.04.2021 9:39:22:** Human voice - "dobrý"

**22.04.2021 9:39:24:** Robot voice - "Promiňte, byly všechny údaje správně?"

**22.04.2021 9:39:30:** Human voice - "ano"

*obr. 13 – detail hovoru (scénář 1)*

V případě, že by člověk slovo “dobrý” zopakoval, doptal by se robot alternativní otázkou, ve které směřuje člověka k odpovědi ano/ne, čímž snižuje možnost chyby.

Text	?	Promiňte, byly všechny údaje správně?
Alternative text	?	Odpovězte prosím ano nebo ne. Bylo vše správně?

*obr. 14 – textace (scénář 1)*

V jiném případě vyřešil bot opakované neporozumění přepojením na operátora.

29.04.2021 21:46:37: Robot voice - "Rozuměl jsem - [redacted] -  
[redacted] - Je to prosím správně?"

29.04.2021 21:46:48: Human voice - "ne"

29.04.2021 21:46:49: Robot voice - "Bohužel jsem nezvládl číslo  
účtu rozpoznat a proto Vás přepojím na mého živého kolegu.  
Vyčkejte na lince."

*obr. 15 – detail hovoru (scénář 1)*

#### 4.1.1.9 Minimalizovat

V některých případech by mohl bot komunikovat stručněji (např. v hovoru na obrázku č. 4), kdy zopakuje jak název banky, tak její kód. Předpokládám, že je v tomto případě stručnost upozaděna, aby se zamezilo nepochopení. Lidé v hovorech mluvili také s dodržением maxim. Korigovat řeč klienta bylo potřeba pouze v případě, že přesáhl délku diktované zprávy (viz. obrázek č. 7)

#### 4.1.1.10 Komunikovat lidsky

Virtuální asistent komunikuje zdvořile, využívá poděkování, je nápomocný a za každé potvrzení správnosti údaje poděkuje. Humor ani small talk se zde nevyskytuje, voicebot nezvládne reagovat na otázky typu "Jak se máš?".

29.04.2021 14:34:32: Robot voice - "Co pro Vás můžu udělat?"

29.04.2021 14:34:36: Human voice - "jak se máš"

29.04.2021 14:34:38: Robot voice - "S tímto Vám zatím  
nedokážu poradit, na to je machr můj lidský kolega."

*obr. 16 – detail hovoru (scénář 1)*

### **4.1.2 Scénář 2: Virtuální zdravotní sestra**

#### 4.1.2.1 Popis scénáře

Virtuální sestra slouží k vyřizování hovorů s požadavky pacientů, aby se lékaři a sestry mohly věnovat své práci a lidé zároveň nemuseli čekat na lince dlouhé minuty na přepojení. Takto mohou svůj problém sdělit virtuálnímu asistentovi, který zprávu následně předá do ordinace. Díky tomu, že může paralelně vyřídit velké množství hovorů, výrazně zvyšuje dovolatelnost. Zdravotní sestry navíc dostanou informace ve strukturované podobě a

při zpětném kontaktování pacienta se mohou na hovor připravit. Na začátku hovoru bot zjistí, zdali se jedná o registrovaného pacienta, následně se ptá na jméno, číslo pojištěnce a na důvod hovoru. Před dotazem proč pacient volá, však pokládá otázku, zdali se jedná o život ohrožující problém a podle toho se řídí. Na závěr ověří kontaktní údaje a informuje pacienta, že vše předává do ordinace, odkud se mu ozvou. V případě urgentního požadavku přímo do ordinace přepojuje.

#### 4.1.2.2 Schopnost vést konverzaci

Scénář není příliš rozvětvený, bot k vedení hovoru využívá uzavřené (*"A jste prosím naším již registrovaným pacientem?"*) i otevřené otázky a instrukce (*"Ted' mi prosím stručně sdělte, kvůli čemu voláte."*). Využívá také časových konverzačních značek (*"A poslední otázka,"*), aby byla konverzace konzistentní a člověk věděl, kde se nachází (*"A ted' k Vašemu požadavku."*). Voicebot pacienta vede postupně od zjištění obecných informací, po dotaz na urgenci požadavku, nadiktování problému a následné ověření kontaktních údajů, na které je možné se ozvat zpět. Při diktování nechává pacientům dostatečný prostor na to, aby svůj problém popsali a následně se ptá, zdali chtějí něco dodat. Jiný průchod scénářem není možný, i přesto je potřeba, aby bot konverzaci vedl k jednomu ze dvou konců hovoru (předání informací a zavěšení, předání informací a přepojení).

Co se týče rizika vyšší latence, voicebot v tomto scénáři nepřistupuje k žádné databázi, s mírnou pauzou je však možné se setkat při popisování problému. Voicebot si musí být jistý, že člověk již domluvil, aby mohl položit otázku *"Přejete si ještě něco dodat?"* Na tomto místě se proto objevuje pauza mezi koncem promluvy člověka a začátku promluvy bota.

#### 4.1.2.3 Vybudovat důvěru

Virtuální sestra je transparentní. Informuje, že je hovor nahráván a pro neregistrované pacienty přidává informaci o tom, kde lze nalézt informace o zpracování osobních údajů. Z hovoru je zřejmé, že pacient hovoří s virtuálním asistentem, nezmiňuje však jméno ordinace. Při sdělování čísla pojištěnce žádá bot o nadiktování celého údaje, ale v případě, že jej klient sdělit nechce, nemá pojištění nebo se číslo nepodaří rozpoznat, hovor nekončí.

**19.07.2022 13:28:50:** Robot voice - "Pro české občany je číslo pojištění stejné jako rodné číslo bez lomítka, najdete ho na kartičce pojištění. Cizinci mají své číslo pojištění také na kartičce. Až budete připraveni diktovat, vyslovte prosím můžu diktovat."

**19.07.2022 13:29:06:** Human voice - "No tak já to nemám No je to v pytlí"

**19.07.2022 13:29:11:** Robot voice - "Číslo pojištění tedy dohledáme později."

*obr. 17 – detail hovoru (scénář 2)*

**14.07.2022 15:42:55:** Robot voice - "Pro české občany je číslo pojištění stejné jako rodné číslo bez lomítka, najdete ho na kartičce pojištění. Cizinci mají své číslo pojištění také na kartičce. Až budete připraveni diktovat, vyslovte prosím můžu diktovat."

**14.07.2022 15:43:10:** Human voice - "nechci"

**14.07.2022 15:43:12:** Robot voice - "Číslo pojištění tedy dohledáme později."

*obr. 18 – detail hovoru (scénář 2)*

Prvkem budování důvěry je informace, že pacientovi bot pomůže předat jeho požadavek do ordinace. Tato informace zazní jak na začátku hovoru, tak na jeho konci. V některých hovorech jsem se naopak setkala s prvky nedůvěry, kdy lidé nebyli ochotni s robotem hovořit, buď vůbec a zavěsili, nebo se neochota projevila pouze na začátku a následně člověk hovorem úspěšně prošel. Toto interpretuji jako neschopnost voicebota vybudovat si respekt. Dalším faktorem, který zde může hrát roli je také to, že na linku volá velké množství starších lidí, jako tomu bylo v hovoru níže. Ti mohou být ze změny, kdy na ně místo zdravotní sestry promluví robot, zmatení, protože nemusí být na komunikaci s hlasovými asistenty zvyklí.

**26.07.2022 16:19:04:** Robot voice - "Prosím, nejprve mi řekněte, zda voláte za sebe, nebo jménem někoho jiného."

**26.07.2022 16:19:14:** Human voice - "Co to je za blbost ale"

*obr. 19 – detail hovoru (scénář 2)*



#### 4.1.2.4 Být konzistentní

Ačkoli se může zdát, že jsou v konverzaci prvky sociálního účelu konverzace, má účel čistě transakční. Bot hned na začátku hovoru jasně informuje, co je cílem hovoru a nastavuje očekávání, s čím může pomoci, tedy vyslechnout pacientův požadavek a předat jej do ordinace, případně na zdravotní sestru pacienta rovnou přepojit. Komunikuje konzistentně a ohleduplně (*“Nevadí, číslo pojištěnce dohledáme později.”* nebo *“Přeji ať je vše brzy v pořádku.”*) Voicebot komunikuje tak, aby bylo jasné, že je maximálně nápomocný. Jak uvádím na příkladu výše, nastavení očekávání v tomto případě nemusí být dostatečné, pokud je cílová skupina zvykla na jiný způsob komunikace. V rámci přizpůsobení se příjemci může na takovéto reakce voicebot poskytnout detailnější informace o tom, jak bude hovor probíhat a přesvědčit volající, že daný požadavek zvládne vyřídit stejně kvalitně, jako člověk.

#### 4.1.2.5 Přizpůsobit se příjemci

Virtuální sestra se přizpůsobuje příjemci hned v několika ohledech. Pro neregistrované pacienty sdělí informaci o zpracování osobních údajů, v případě život ohrožujícího problému přepojuje do ordinace. Klient se v průběhu scénáře má možnost doptat na detail, co znamená požadavek *“urgentní z důvodu vážného ohrožení zdraví nebo života”*. V ostatních případech bot vždy vysvětlí, co daným požadavkem myslí (například kde najde pacient číslo pojištěnce) nebo kam se obrátit, v případě, že by se zdravotní stav zhoršil.

**20.07.2022 9:58:33:** Robot voice - "A teď k Vašemu požadavku. Je prosím urgentní z důvodu vážného ohrožení zdraví nebo života?"

**20.07.2022 9:58:41:** Human voice - "No nevím to nedokážu posoudit"

**20.07.2022 9:58:48:** Robot voice - "Pokud nevíte jistě, zkuste si prosím představit, že bychom se Vám ozvali zpět například až zítra nebo pozítří. Bylo by Vaše zdraví vážně ohroženo?"

**20.07.2022 9:58:58:** Human voice - "ne"

*obr. 20 – detail hovoru (scénář 2)*

Při analýze hovorů jsem zaznamenala, že měli někteří pacienti tendenci přizpůsobit se voicebotovi a při diktování vzkazu mluvili velmi zřetelně, až roboticky.

#### 4.1.2.6 Schopnost porozumět a naslouchat

Ve scénáři je výrazný prvek aktivního naslouchání – zatímco člověk diktuje vzkaz, bot na pozadí přehrává zvuk psaní na klávesnici, aby bylo zřejmé, že zprávu zaznamenává. Na konci promluvy pacienta se ptá “*Přejete si ještě něco dodat?*”, pokud by chtěl pacient pokračovat v promluvě. Ve scénáři nejsou globální intenty, není možné například zažádat o přepojení na sestru. Voicebot využívá explicitní souhlasy.

#### 4.1.2.7 Schopnost udržet kontext

Virtuální asistent pracuje s kontextem klienta. Ví, že se nachází v nepříjemné situaci, proto se snaží být co nejvíce nápomocný a sdělit co nejvíce informací, které by člověk v dané situaci mohl potřebovat.

**22.07.2022 13:26:10:** Robot voice - "Dobře. A poslední otázka. Můžeme Vás kontaktovat na tomto telefonním čísle, ze kterého voláte?"

**22.07.2022 13:26:19:** Human voice - "ano"

**22.07.2022 13:26:21:** Robot voice - "Děkuji, informace ihned předávám do ordinace. V případě zhoršení zdravotního stavu volejte znovu do ordinace nebo záchrannou službu na čísle 155. Pokud se stav nijak nezhorší, vyčkejte prosím alespoň jeden pracovní den. Nekontaktujte nás znovu ve stejné věci, Váš požadavek už byl zaznamenán. Zatím Vám přeji hezký den. Nashledanou."

**22.07.2022 13:26:45:** Hung up

*obr. 21 – detail hovoru (scénář 2)*

Co se týče získaných údajů, voicebot nepracuje se jménem klienta, pouze s telefonním číslem pacienta při ověření kontaktního údaje. Zároveň se řídí podle toho, zdali je případ urgentní, či nikoliv.

#### 4.1.2.8 Pracovat s chybou a nepochopením

V tomto scénáři se pracuje s nepochopením odlišně, než tomu bylo ve scénáři výše. Na rozdíl od diktování čísla účtu a částky zde voicebot neklade takový důraz na přesnost porozumění, jako spíše na rychlost vyřízení a předání požadavku do ordinace. Klient zde diktuje jméno, číslo pojištěnce a volnou řečí popisuje zdravotní stav. V případě, že se některý z údajů nepodaří rozpoznat, nebo jej pacient nepotvrdí, voicebot nabídne jednu možnost

opravy a pokud se nepodaří ani napodruhé, pokračuje v hovoru dále s informací, že si daný údaj ověří zdravotní sestra později.

K nepřesnému přepisu řeči může docházet při diktování problému, se kterým člověk volá. Nepřesnosti, které jsem objevila v analyzovaných hovorech, se však týkaly pouze jednotlivých slov (např. “*píše A*” místo “*PCR*”, “*PC artest*” místo “*PCR test*” nebo “*kuny*” místo “*kovid*”). Většina z nich neměla zásadní vliv na pochopení kontextu celé zprávy. U některých vyžadovalo pochopení větší představivost, jako například hovor níže.

**19.07.2022 11:44:09:** Robot voice - "Teď mi prosím stručně sdělte, kvůli čemu voláte. Následně se Vás pokusím přepojit do ordinace."

**19.07.2022 11:44:18:** Human voice - "Pozitivní test na kuny, prosím"

**19.07.2022 11:44:26:** Robot voice - "Přejete si ještě něco dodat?"

**19.07.2022 11:44:30:** Human voice - "ne"

*obr. 22 – detail hovoru (scénář 2)*

#### 4.1.2.9 Minimalizovat

Virtuální asistent není ve svých promluvách příliš stručný, avšak ostatní konverzační maxima dodržuje. Mluví transparentně a nepoužívá nejednoznačné formulace. Všechny informace, na které by se mohl pacient doptat, dopředu vysvětlí. Týká se to například čísla pojištěnce, kdy voicebot informuje nejen, že je pro české občany identické, jako rodné číslo, hovoří také o tom, kde jej najít. Tato instrukce je i pro cizince. V analyzovaných hovorech jsem zaznamenala, že i přes toto vysvětlení někteří lidé diktovali název pojišťovny.

**15.07.2022 13:26:34:** Robot voice - "Pro české občany je číslo pojištěnce stejné jako rodné číslo bez lomítka, najdete ho na kartičce pojištěnce. Cizinci mají své číslo pojištěnce také na kartičce. Až budete připraveni diktovat, vyslovte prosím můžu diktovat."

**15.07.2022 13:26:58:** Human voice - "VZP"

*obr. 23 – detail hovoru (scénář 2)*

Či v tomto případě:

**04.07.2022 13:45:31:** Robot voice - "Pro české občany je číslo pojištěnce stejné jako rodné číslo bez lomítka, najdete ho na kartičce pojištěnce. Cizinci mají své číslo pojištěnce také na kartičce. Až budete připraveni diktovat, vyslovte prosím můžu diktovat."

**04.07.2022 13:45:47:** Human voice - "Jsem pojištěný u všeobecné zdravotní pojišťovny"

*obr: 24 – detail hovoru (scénář 2)*

#### 4.1.2.10 Komunikovat lidsky

Zvolený ženský hlas i promluvy voicebota působí uklidňujícím dojmem, Bot má za cíl být co nejvíce nápomocný, proto komunikuje lidsky a zdvořile. Vzhledem k situaci, ve které pacienti volají zde však není příliš velký prostor pro práci se small talkem či s humorem.

### **4.1.3 Scénář 3: Vymáhání pohledávek**

#### 4.1.3.1 Popis scénáře

Poslední analyzovaný scénář je z oblasti věřitelů. Rozhodla jsem se jej analyzovat právě z toho důvodu, že je osobnost bota odlišná od předchozí virtuální sestry, proto do analýzy přináší další diverzifikaci dat. Virtuální asistenti, kteří působí v oblasti vymáhání pohledávek mohou mít samozřejmě různé osobnosti. Ty se odvíjí od stavu dluhu, se kterým se klientům připomínají a také od firemní komunikace dané společnosti. Voicebot může volat s připomenutím blížící se splátky (tzv. pre-collection), s upomínkou nezaplacené platby (tzv. soft collection), nebo řešit inkaso pohledávek (tzv. hard collection), tedy mimosoudní vymáhání. Od toho se pak může odvíjet i styl komunikace.

V tomto případě se jedná se o odchozí hovor, ve kterém voicebot volá lidem, kvůli inkasu pohledávek. Nejprve je potřeba ověřit, zdali mluví se správným člověkem, jinak není možné v hovoru pokračovat. Toto ověření probíhá potvrzením jména a dotazem na rok narození. Následně bot informuje o výši dluhu a domlouvá se s klientem na datu splacení a výši splátky. Hovor je poměrně komplexní a má několik možných závěrů – může se jednat o

omyl, klient nemusí mít čas na hovor a bot se s ním domluví na jiném termínu hovoru. Ideálním koncem je situace, kdy mezi botem a člověkem proběhne dohoda. Může se však stát, že hovor skončí tím, že dohoda neproběhla, nebo se chce domluvit na individuálních splátkách, v takovém případě končí hovor přepojením na operátora.

#### 4.1.3.2 Schopnost vést konverzaci

Vzhledem k rozvětvení scénáře je potřeba, aby bot dokázal konverzaci vést a směřovat k jednomu z možných konců. V případě ideální cesty, tedy domluvení se na splacení dluhu, voicebot vede konverzaci pomocí kladení uzavřených otázek na výši splátky a termínu. V případě, že se na zaplacení s klientem dohodne, zopakuje mu v závěru, na čem se společně dohodli.

**01.07.2022 12:33:53:** Robot voice - "Prosím odpovězte ano nebo ne, zdali uhradíte celý dluh jednorázově."

**01.07.2022 12:33:59:** Human voice - "momentálně nemám"

**01.07.2022 12:34:01:** Robot voice - "A zvládnete uhradit najednou alespoň třetinu dlužné částky, což je █████ korun?"

*obr. 25 – detail hovoru (scénář 3)*

Voicebot ověřuje rok narození klienta, sděluje mu jeho jméno, částku i společnost, u které dluží. K latenci však v hovoru nedochází, protože jsou tyto informace uloženy do mezipaměti (cache) voicebota před tím, než hovor vytočí. Není tak potřeba data v průběhu hovoru ověřovat v databázi. Virtuální asistent využívá také konverzační značky (*"Abychom mohli pokračovat,"*) a spíše než poděkování využívá ve scénáři potvrzení.

**22.07.2022 18:50:01:** Robot voice - "Můžeme se domluvit, že to uhradíte do čtrnácti dní?"

**22.07.2022 18:50:05:** Human voice - "ano"

**22.07.2022 18:50:07:** Robot voice - "Dobře. Dohodli jsme se tedy, že třetinu dlužné částky splatíte do dvou týdnů."

*obr. 26 – detail hovoru (scénář 3)*

#### 4.1.3.3 Vybudovat důvěru

Téma dluhů může být pro spoustu lidí velmi citlivé. Je proto potřeba, aby byl voicebot co nejvíce transparentní. Bot se na úvod hovoru představí, sdělí z jaké společnosti volá a tyto informace zopakuje také při dotazu kdykoliv v průběhu hovoru.

**15.07.2022 16:57:46:** Robot voice - "Abychom mohli pokračovat, potreboval bych znát rok vašeho narození."

**15.07.2022 16:57:51:** Human voice - "A kdo volal vůbec"

**15.07.2022 16:57:54:** Robot voice - "Jsem virtuální operátor [redacted] ze společnosti [redacted]."

*obr. 27 – detail hovoru (scénář 3)*

Dalším prvkem k vybudování důvěry a transparentnosti je oslovení klienta jménem, aby bylo zřejmé, že je hovor směřován přímo jemu. Transparentní je také ve sdělení výši dluhu, společnosti, u které dluh vznikl a také v informacích, co může následovat, pokud klient dluh nezaplatí. Na tyto informace je možné se v průběhu hovoru kdykoliv doptat. Virtuální asistent také hned v první promluvě sděluje, že je hovor monitorován.

**22.07.2022 18:47:30:** Robot voice - "Dobrý den, u telefonu [redacted] z [redacted]. Na začátek vás upozorním, že je hovor monitorován."

**22.07.2022 18:47:39:** Robot voice - "Mluvíím prosím s panem [redacted]?"

**22.07.2022 18:47:48:** Human voice - "ano"

*obr. 28 – detail hovoru (scénář 3)*

Při sdělení data narození je možné říct pouze rok. V některých případech však lidé v analyzovaných hovorech datum odmítli sdělit, voicebot tak nemohl pokračovat v hovoru.

**01.07.2022 13:19:23:** Robot voice - "Vše vám vysvětlím během dvou minut. Nejprve potrebuji znát rok vašeho narození."

**01.07.2022 13:19:32:** Human voice - "a já potrebuji jako vědět nemám nebudu dávat svoje"

*obr. 29 – detail hovoru (scénář 3)*

Tyto hovory interpretuji jako neschopnost voicebota vybudovat důvěru klienta, který odmítl sdělit svůj rok narození. To zcela jistě souvisí s nastavením očekávání a cíle konverzace. V toto konkrétním případě je však obtížné vybudovat důvěru, pokud není možné s cílem konverzace pracovat, aniž by voicebot klienta neověřil pomocí jména a data narození.

22.07.2022 18:40:56: Robot voice - "Slyšíme se? Prosím váš rok narození, ať můžeme pokračovat."

22.07.2022 18:41:09: Robot voice - "Bohužel se mi nepodařilo ověřit vaši totožnost. Hovor musím ukončit."

22.07.2022 18:41:15: Robot voice - "Kontaktujte nás prosím na čísle [REDAKCE]."

*obr. 30 – detail hovoru (scénář 3)*

#### 4.1.3.4 Být konzistentní

Účel konverzace je transakční a hlavním cílem je vytvořit dohodu. Bot cíl hovoru nesdělí hned na úvod, protože informaci o výši dluhu ani o jeho existenci nesmí bot sdělit nikomu jinému, než dlužníkovi. Bot komunikuje spíše stroze, projevuje se to například tím, že spíše než poděkování používá potvrzení (viz hovor na obrázku č. 26).

#### 4.1.3.5 Přizpůsobit se příjemci

V tomto scénáři voicebot nepracuje s žádným stupněm expertízy či s frekvencí hovorů, se kterou daný klient může projít scénářem. Opakovaný hovor by mohl nastat v případě, že by daný klient měl více dluhů, které by převzala tato inkasní společnost. Voicebot však se všemi prochází stejný scénář. Voicebot hovoří víceméně stručně, klient má však v průběhu scénáře možnost doptat se na veškeré potřebné informace.

Voicebot se umí přizpůsobit klientům, kteří se chtějí domluvit na individuální splátky a ty přepojuje na operátora. Zároveň se přizpůsobuje také otevírací době call centra a pokud přepojení není možné, hovor ukončí se zprávou, kdy se má klient na dané číslo ozvat zpátky. Dalším přizpůsobením se klientovi je respektování změny klientova názoru v průběhu scénáře. Toto přizpůsobení souvisí s kontextem scénáře a mluvím o něm níže.

#### 4.1.3.6 Schopnost porozumět a naslouchat

Virtuální asistent v hovorech využívá především explicitní potvrzení a v případě dohody na závěr hovoru zopakuje, na čem se s klientem dohodl. Důležité je správné porozumění roku narození hned na začátku scénáře, protože v opačném případě by hovor nemohl pokračovat dále.

#### 4.1.3.7 Schopnost udržet kontext

Voicebot má o klientovi tyto informace: jméno, dlužnou částku a firmu, ve které dluží. S těmito informacemi v průběhu hovoru pracuje. Zároveň pracuje s informací, v jakém stavu se nachází klientova půjčka, a pokud se klient na splacení dluhu odmítne domluvit, informuje jej, jaké kroky budou následovat. Poté mu znovu nabízí domluvit se na splátce.

Voicebot tedy nabízí klientovi rozmyslet si původní rozhodnutí. Pokud klient odpoví ano, voicebot "zapomíná" předchozí klientovo rozhodnutí a prochází s ním scénář znovu. Pracuje zde tedy s kontextem toho, že je klient v situaci, kdy neví, co může nesplacením dluhu nastat a dává mu možnost přehodnotit předchozí tvrzení.

#### 4.1.3.8 Pracovat s chybou a neporozuměním

S možností chyby pracuje voicebot hned na začátku hovoru, kdy počítá s variantou, že se dovolal na nesprávné číslo a jedná se o omyl. V analyzovaných hovorech jsem se pak setkala s občasným neporozuměním při diktování data narození, kdy v některých případech STT engine přepsal číslo do formátu, který bot nevyhodnotil jako datum. V tomto případě se bot informoval člověka, že špatně rozuměl a požádal jej o sdělení data znovu.

22.07.2022 18:14:27: Robot voice - "Abychom mohli pokračovat, potřeboval bych znát rok vašeho narození."

22.07.2022 18:14:33: Human voice - "54█"

22.07.2022 18:14:36: Robot voice - "Asi vám špatně rozumím, řekněte mi prosím ještě jednou rok narození."

22.07.2022 18:14:41: Human voice - "5 4 19█"

obr. 31 – detail hovoru (scénář 3)



I v tomto scénáři je velký důraz na správné porozumění, domluvení se na splátce je totiž závazné. Gramatiky souhlasu či nesouhlasu proto obsahují i jiné výrazy, než pouze ano či ne. Kromě gramatiky souhlasu bot zachytává také slova jako “uhradím” či “zvládnou”. Je však potřeba dát pozor na to, aby v kontextu věty opravdu znamenaly souhlas.

#### 4.1.3.9 Minimalizovat

Komunikace voicebota je v tomto případě velmi stručná a podobné jsou také odpovědi klientů. Na většinu otázek lze odpovědět ano nebo ne. Voicebot dodržuje všechny čtyři konverzační maxima, je transparentní v důvodu hovoru i v informacích o dlužné částce. Je univerzální, ale konkrétní, při čtení informací o stavu dluhu klienta a o krocích, které by následovaly nezaplacení dluhu v dohodnutém termínu.

#### 4.1.3.10 Komunikovat lidsky

Voicebot komunikuje k věci, nedá se ale říci, že by byl příliš zdvořilý. Svou osobností je přizpůsobený cílové skupině, se kterou komunikuje a jeho hlavní cíl je domluvit se na splacení dlužné částky, nebo alespoň její části. Z tohoto důvodu zde není prostor ani pro small talk, či humor. Zároveň je i nápomocný, v případě, že klient nemá peníze ani na zaplacení navrhované výše splátky, přepojí jej na operátora, který se domluví na individuálních platbách. Bere v potaz také to, že klient, který se původně domluvit na zaplacení nechtěl, může změnit názor poté, co mu voicebot sdělí, v jakém stavu je jeho dluh a nabídne mu možnost znovu se domluvit na splátce. V některých případech totiž nezaplacení dlužné částky vede k předání tohoto dluhu do soudního vymáhání, kde se mohou náklady na dluh navýšit o desetitisíce korun. Voicebot proto informuje klienta, že je i v jeho zájmu se na splátkách domluvit a že je mu nápomocen situaci vyřešit.

## 4.2 Výsledky analýzy

### 4.2.1 Výzkumná otázka a hypotézy

Výzkumná otázka v této diplomové práci zněla:

1. Jaká jsou pravidla konverzačního designu, aby byla komunikace lidí s roboty co nejvíce přirozená?

Hypotéza k této výzkumné otázce byla následující:

1. Existuje soubor pravidel, který je potřeba dodržet při tvorbě virtuálního asistenta. Tato pravidla jsou čistě konverzačního charakteru a vychází z konverzační analýzy a z chování lidí v mezilidské komunikaci.

Tato hypotéza se potvrdila částečně. Lze konstatovat, že existuje soubor pravidel, který je potřeba při tvorbě dodržet. Soubor těchto pravidel není nikde definován. Pravidla však nejsou pouze konverzačního charakteru, ale jsou velmi úzce spjata také s technologickými požadavky. Některá pravidla přímo vychází z limitů technologie a v mezilidské komunikaci by nenastala (například kategorizace textu a lemmatizace). Jak technické požadavky, tak konverzační pravidla vychází z chování lidí v mezilidské komunikaci, protože cílem přirozené komunikace mezi člověkem a robotem je snaha co nejvíce se přizpůsobit té mezilidské, proto při tvorbě můžeme využít poznatky z konverzační analýzy. Nedá se však říci, že by se vlastností konverzace tak, jak je v rámci konverzační analýzy popsat Harvey Sacks, dala aplikovat na oba druhy konverzace (mezilidskou a tu s robotem). Jak uvádím v kapitole, kde se konverzační analýze věnuji, schopnost konverzace s voicebotem se v několika bodech zásadně odlišuje.

Odpověď na výzkumnou otázku je tedy následující. Jsou to pravidla technického i konverzačního charakteru, která jsou vzájemně propojená. V modelu pravidel, který jsem pro tuto práci vytvořila tato pravidla a požadavky shrnuji a zároveň uvádím, jaký vliv na schopnosti a dovednosti voicebota mají. Matice modelu pravidel je v kapitole níže.

K výzkumné otázce jsem doplnila také dvě podotázky a navazující hypotézy.

- a. Jsou lidé schopni komunikovat s robotem?
- b. Jsou lidé ochotni komunikovat s robotem?

Hypotézy k těmto otázkám jsou následující:

- a. Lidé jsou schopni plnohodnotné komunikace s virtuálním asistentem.
- b. Lidé jsou ochotni vyřídit svůj požadavek s robotem.

V hypotéze pod k první podotázce by bylo potřeba nejprve si definovat, co znamená plnohodnotná komunikace. V případě, že je to komunikace, která má začátek, konec a v jejím průběhu je splněn její cíl, který byl na začátku stanoven, pak je lze tuto hypotézu potvrdit. Lidé tedy jsou schopni komunikace s robotem.

Druhou hypotézu lze vzhledem k analyzovaným hovorům částečně potvrdit. Redefinovala bych ji na hypotézu a zároveň odpověď na podotázku: Někteří lidé jsou ochotni vyřídit svůj požadavek s robotem.

#### **4.2.2 Ověření poznatků na empirických datech**

V předchozích kapitolách jsem sbírala poznatky z odborné literatury a studií a následně jsem tyto poznatky ověřovala skrze interpretaci empirických dat z telefonních hovorů. Pro jednotlivá témata jsem vytvořila celkem devět dovedností, do kterých spadají: schopnost vést konverzaci, vybudovat důvěru, být konzistentní, přizpůsobit se příjemci, porozumět a naslouchat, udržet kontext, pracovat s chybou a nepochopením, minimalizovat a komunikovat lidsky. V této části práce se budu věnovat tomu, zdali na základě empirických dat mohu potvrdit či vyvrátit, že by měl voicebot mít sadu právě těchto schopností a dovedností. Při analyzování hovorů jsem se však setkala s mnoha situacemi, kdy jsem se při snaze zařadit chování bota ke konkrétní dovednosti zjistila, dané chování může spadat pod více dovedností zároveň a také že se dovednosti vzájemně ovlivňují. Díky analýze hovorů a scénářů vyvstala potřeba seznam aktualizovat o jednu další dovednost, a tou je schopnost vzbudit respekt.

#### 4.2.2.1 Schopnost vést konverzaci

Potřeba této dovednosti se potvrdila.

Ve všech analyzovaných scénářích jsem identifikovala schopnost virtuálního asistenta vést konverzaci. Boti ve všech scénářích využívali konverzačních značek, otevřených i uzavřených otázek, aby konverzaci vedli od zahájení až po její konec. Přestože jsem v této analýze neměla scénář, ve kterém by bot konverzaci nevedl, dovoluji si tvrdit, že je tato dovednost potřeba. Konverzace s voicebotem po telefonu mají totiž v drtivé většině případů transakční účel a dopředu definovaný jeden či více cílů. Hlavní cíl konverzace je znám oběma stranám, protože jej voicebot uvede na začátku hovoru v rámci nastavení správného očekávání. Neinformuje však klienta o dalších možných koncích, ani o detailu, jak bude hovor probíhat, jakými konverzačními uzly například dle expertízy klienta je potřeba projít a podobně. Vzhledem k tomu, že je voicebot jediný, kdo zná průběh konverzace, je nutné, aby dialog vedl.

#### 4.2.2.2 Vybudovat důvěru

Potřeba této dovednosti se potvrdila pouze částečně.

U scénáře pro vymáhání pohledávek jsem potřebu schopnosti zaznamenala. V některých situacích totiž lidé hovor pokládali hned na samotném začátku. Myslím si, že to bylo právě z důvodu, že voicebot nezvládl hned na počátku hovoru vybudovat důvěru. To potvrzuje také studie hlasových asistentů, ve které autoři upozorňují na to, že se neočekávané reakce (v tomto případě zavěšení) mohou objevit již po prvním obratu (promluvě voicebota). *“Hned na začátku hovoru je proto dobré sdělit, že vám tento virtuální agent může pomoci s vyřešením vašeho problému.”*<sup>102</sup> Podle mého názoru by toto sdělení, včetně akcentu pomoci, mohlo v lidech vybudovat důvěru k voicebotovi a zároveň nastavit správné očekávání, přičemž tyto dvě schopnosti jsou vzájemně propojené.

---

<sup>102</sup> SZYMANSKI, Margaret H. a Robert J. MOORE. Adapting to Customer Initiative: Insights from Human Service Encounters. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 19-32 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_2, str. 31.

Naopak u scénářů pro zadání platby k úhradě jsem se nesečkala s žádným chováním voicebota, které by značilo snahu vybudovat důvěru. U virtuální sestry jsem jako prvek důvěry označila transparentnost při sdělování informací o zpracování údajů a také o sdělení, že všechna data ihned předá do ordinace. Obecně však u botů, kteří působí v oblasti zdravotnictví, hraje důvěra velkou roli, protože scénáře mnohou obsahovat sdělování stigmatizujících informací (např. o užívání návykových látek, poradenství v léčbě duševních onemocnění či HIV).<sup>103</sup>

#### 4.2.2.3 Být konzistentní

Nelze určit, zdali se potřeba dovednosti potvrdila či vyvrátila.

Ve všech analyzovaných scénářích komunikoval bot konzistentně. Proto nelze určit, zdali se potřeba této dovednosti potvrdila či nikoliv. Toto tvrzení by se dalo vyvrátit či potvrdit za předpokladu, že bych měla v analyzovaném vzorku hovory, ve kterých virtuální asistent komunikuje nekonzistentně. Co se týče osobnosti bota, ta byla v analyzovaných hovorech, zejména u virtuální zdravotní sestry, provázána se schopností komunikovat lidsky.

#### 4.2.2.4 Přizpůsobit se příjemci

Potřeba této dovednosti se potvrdila.

V hovorech při zadávání platby k úhradě i při vymáhání pohledávek bylo několik situací, kdy se člověk doptal voicebota na detailní informace. Většina však prošla bez nutnosti tyto informace znát. Dá se tedy říci, že se bot přizpůsobil cílové skupině příjemců. Virtuální zdravotní sestra zase pracovala s kontextem, díky kterému se mohla přizpůsobit pacientům, kteří volali se život ohrožujícím problémem. Stejně tak bot z oblasti věřitelů přizpůsoval své promluvy dle kontextu stavu dluhu u daného klienta. Na těchto příkladech můžeme vidět propojení dvou dovedností, kdy práce s kontextem přímo navazuje na schopnost přizpůsobit se příjemci.

---

<sup>103</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 33.

#### 4.2.2.5 Schopnost porozumět a naslouchat

Potřeba této dovednosti se potvrdila.

U hovorů ve scénářích pro zadání platby i vymáhání se potvrdila potřeba velmi přesného naslouchání, protože domluva na splacení dluhu, stejně jako zadání platby k úhradě jsou závazné činnosti. Z toho důvodu si všechna sdělená data voicebot s člověkem explicitně ověřoval. U virtuální sestry nebyl takto velký důraz na přesnost dat. I v případě, že se voicebotovi nepodařilo zachytit jméno či číslo pojištěnce, předal vždy danou zprávu do ordinace. Pacienti zde měli možnost diktovat poměrně dlouhou promluvu o svém zdravotním stavu či žádosti o předepsání určitých léků. Ačkoliv byla většina prepisů v naprostém pořádku, v některých situacích bot neporozuměl některému slovu správně, což však neovlivnilo celkový kontext zprávy. Zde je možné poukázat na logické propojení schopnosti porozumět se schopností pracovat s chybou.

#### 4.2.2.6 Schopnost udržet kontext

Potřeba této dovednosti se potvrdila.

S potřebou znát kontext klienta jsem se setkala u všech hovorů. Konkrétně to byla práce se situačním kontextem (urgentní hovor v život ohrožujícím stavu automaticky přepojen do ordinace či nabídka individuálních splátek při informaci, že člověk nemůže požadovaný dluh zaplatit z nedostatku financí), s kontextem dat nasbíraných v průběhu hovoru (rekapitulace jen těch údajů k platbě, které si člověk přál zadat), tak kontextem dat o klientovi (oslovení jménem, informace o výši a stavu dluhu). V případě, že by bot v analyzovaných hovorech s kontextem nepracoval, pozbývaly by smyslu. Proto tuto dovednost považují za potřebnou.

#### 4.2.2.7 Pracovat s chybou a neporozuměním

Potřeba této dovednosti se potvrdila.

Z analýzy se ukázalo, že čím větší je důraz na správné porozumění, tím lépe musí umět voicebot pracovat s možnou chybou a neporozuměním. Dalším poznatkem, který z analýzy hovorů vyplynul, je to, že ne všechny chyby a neporozumění jsou způsobeny chybou

technologie. V některých případech se objevila chyba konverzačního designu, kdy některé klíčové slovo nebylo přidáno do správného intentu a voicebot proto reagoval chybně. To souvisí s potřebou hovory nejen důkladně testovat, ale také dodatečně upravovat dle analýzy proběhlých hovorů a mít u voicebota schopnost asistovaného učení.

#### 4.2.2.8 Minimalizovat

Potřeba této dovednosti se částečně potvrdila.

U hovorů z bankovního sektoru byla minimalizace upozaděna za vzájemné porozumění. Voicebot například četl jak kód banky, tak její název, aby si byl jistý, že platbu posílá na správný účet. Stejně tak v některých promluvách nebyla stručná ani virtuální sestra, která při žádost o diktování čísla pojištěnce rovnou sdělila i poměrně zdlouhavou informaci, co číslo pojištěnce je a kde jej lze najít. V komunikaci ve zdravotnictví je dobrou praxí poskytnout pacientům co největší prostor, aby mohli dobře popsat svůj problém.<sup>104</sup> Proto i při diktování důvodu hovoru virtuální sestra pacienty nikde nepřerušovala a pokud to bylo potřeba, měli možnost promluvu dvakrát doplnit o další řeč.

Naopak velmi stručné byly hovory ve scénáři k vymáhání pohledávek, kdy byl bot ve svých promluvách velmi stručný a stejně tak směřoval ke stručnosti i odpovědi člověka díky použití otázek, na které se dalo většinou odpovědět ano či ne. Zde minimalizace souvisí také s osobností bota, ve dvou dalších scénářích je pak provázaná s dovedností přizpůsobit se příjemci.

#### 4.2.2.9 Komunikovat lidsky

Nelze určit, zdali se potřeba dovednosti potvrdila či vyvrátila.

Styl komunikace velmi souvisí s konkrétní personou. Ani v jednom z analyzovaných scénářů nebyl prostor pro small talk. V analyzovaných hovorech jsem se s otázkou člověka na bota “Jak se máš?” setkala pouze v jednom případě, a to u scénáře pro zadání platby. Zároveň v této analýze nemám dostatek dat pro to, abych mohla porovnat, zdali je zdvořilost, ohleduplnost, možnost small talku a humoru potřeba či nikoliv.

---

<sup>104</sup> BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. Studies in Conversational UX Design [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3, str. 42.

#### 4.2.2.10 Vzbudit respekt

Potřeba této dovednosti se potvrdila.

Vzbudit respekt je dovednost, kterou jsem do seznamu zařadila dodatečně po analýze hovorů. V malém množství z nich jsem se totiž setkala s neochotou s voicebotem vést konverzaci, což jsem vyhodnotila jako důsledek absence dovednosti vzbudit u lidí respekt.

**24.06.2022 8:50:18:** Robot voice - "Pro české občany je číslo pojištění stejné jako rodné číslo bez lomítka, najdete ho na kartičce pojištění. Cizinci mají své číslo pojištění také na kartičce. Až budete připraveni diktovat, vyslovte prosím můžu diktovat."

**24.06.2022 8:50:35:** Human voice - "Ježišmarja co to je tohle"

**24.06.2022 8:50:38:** Robot voice - "Až budete mít připravené číslo pojištění k diktování, prosím, vyslovte můžu diktovat."

**24.06.2022 8:50:47:** Human voice - "Ježíš já už jsem tam u pana doktora byla já se chci jenom objednat"

**24.06.2022 8:50:51:** Robot voice - "Číslo pojištění tedy dohledáme později."

*obr. 32 – detail hovoru (scénář 2)*

V kontextu této analýzy chápu vzbuzení respektu jako přesvědčení lidí o tom, že voicebot zvládne vyřídit jejich požadavek stejně dobře, jako lidský operátor, případně že komunikace s robotem převáží jiné nevýhody, například čekat několik minut na obsazené lince do ordinace. Tato dovednost je provázaná jak s nastavením očekávání, tak také s dovedností vybudovat důvěru. Zcela jistě souvisí i se schopností vést konverzaci. Voicebot, který umí vybudovat respekt by měl být asertivní, mít tzv. tah na branku a udržet pozornost klienta tak, aby konverzace došla do správného konce.

#### **4.2.3 Cíl práce: Vytvoření modelu pravidel**

Hlavním cílem této diplomové práce je definovat model pravidel konverzačního designu, která jsou zásadní pro to, aby byla komunikace lidí s roboty co nejvíce přirozená. Nejprve je však potřeba vymezit, co v tomto kontextu model pravidel znamená a k čemu slouží. Model zde chápu jako typ či formu, kterou je možné aplikovat na tvorbu voicebotů v



určité kategorii. V případě této diplomové práce a v kontextu analyzovaných dat je to kategorie virtuálních asistentů komunikujících po telefonu. Pravidla v tomto případě reprezentují poznatky z provedené analýzy. Na tento model pravidel se můžeme dívat také jako na seznam funkčních požadavků, které jsou kladeny na virtuální asistenty, aby byli schopni konverzovat s lidmi prostřednictvím telefonních hovorů.

#### 4.2.3.1 Matice pravidel

Při ověření poznatků na empirických datech se potřeba jednotlivých dovedností buď potvrdila, nebo částečně potvrdila. U dvou dovedností nešla potvrdit, ani vyvrátit. Ani u jedné z dovedností však její potřebu empirická data nevyvrátila. Zároveň bylo zjištěno, že jsou na sebe jednotlivé dovednosti navázané. Model pravidel jsem proto vytvořila jako matici, kdy je v horním řádku tabulky uvedeno všech deset dovedností voicebota a v levém sloupci jsou pak uvedeny konkrétní pravidla či požadavky. Ty jsou rozděleny na dvě sekce – na technické požadavky a na konverzační pravidla. V tabulce je dále vyznačeno, se kterou dovedností se daný požadavek či pravidlo protíná, tedy které schopnosti ovlivňuje.

Tato tabulka může nyní sloužit jako pracovní nástroj pro konverzační designery. Lze ji totiž využít na získání informací o tom, co daná funkce (například lemmatizace) ve skutečnosti ovlivní. Zároveň ji lze použít opačným způsobem, kdy můžeme jednoduše vyfiltrovat, které požadavky je potřeba splnit, aby byl voicebot schopný určité dovednosti (například vést konverzaci).

#### 4.2.3.1.1 Technologické požadavky

TECHNOLOGICKÉ POŽADAVKY																					
		Schopnost vést konverzaci	Vybudovat důvěru	Být konzistentní	Přizpůsobit se příjemci	Schopnost porozumět a naslouchat	Pracovat s chybou a neporozuměním	Minimalizovat	Komunikovat lidsky	Vzbudit respekt											
↑ více důležité	Správně přepsat řeč uživatele				x	x															
	Eliminovat šum				x	x															
	Klasifikovat text	x			x	x															
	Lematizovat	x			x	x															
	Mluvit syntetickým hlasem	x																			
	Skloňovat	x	x		x																
	Mít možnost asistovaného učení				x	x		x													
	Zamezit zacyklení při opakovaném neporozumění	x			x	x	x	x	x												
	Ukládat získaná data v průběhu do proměnných	x	x		x	x	x														
	Pracovat s dostupnými informacemi o klientovi		x		x	x	x			x											x
↓ méně důležité	Ověřovat dostupná data s databází		x		x	x			x	x											x
	Ukládat přepisy hovorů ke zpětné analýze				x	x			x												
	Pracovat s jistotou porozumění					x			x	x											
	Znáť expertizu klienta a frekvenci komunikace		x		x						x										x
	Reagovat bez pauz mezi přechody promluv	x			x					x											
	Vyhodnotit tempo člověka a přizpůsobit se mu					x					x			x		x					x
	Umět přepojit na lidského operátora					x				x	x										x
	Zpomalit při čtení složitých prvků (adresa, email)					x															
	Umět přehrát zvuky v pozadí		x				x														
	Umožnit skákání do řeči									x		x	x	x							

tab. 1 – technologické požadavky

Při tvorbě seznamu technologických požadavků jsem vycházela především z teoretické části práce a také z principů fungování hlasových asistentů, které popisují v kapitole Jak komunikují voiceboti. Při pohledu na tabulku lze konstatovat, že většina technologických požadavků je provázána na schopnost porozumět a naslouchat, na práci s chybou a neporozuměním a také na přizpůsobení se příjemci. U prvních dvou dovedností to celkem logicky vychází právě z hlavní funkce, kterou technologické pozadí hlasových asistentů má, a to je přepsat řeč člověka. Hledat v ní význam je pak společná činnost NLP a konverzačního designera. I proto uvádím požadavek na správné přepsání řeči uživatele na prvním místě. S tím souvisí požadavek na eliminaci okolního hluku. Následují požadavky na práci s textem, tedy jeho klasifikace, lematizace a skloňování (například jmen), před které jsem umístila schopnost mluvit syntetickým hlasem. Následuje výčet dalších požadavků, které jsou seřazeny podle důležitosti. Zajímavým poznatkem je to, že schopnost komunikovat lidsky zde protínají pouze dva požadavky, a to umět se přizpůsobit tempu řeči člověka a umožnit skákání do řeči.

#### 4.2.3.1.2 Konverzační pravidla

		KONVERZAČNÍ POŽADAVKY									
		↑					↓				
víc důležit é	↑	Používat konverzační značky	x	x	x						
		Klást uzavřené i otevřené otázky, instrukce	x		x						
mén ě důležit é	↓	Dát najevo střídání mluvčích	x		x						
		Aktivně naslouchat	x		x						
		Reagovat na neporozumění			x	x					
		Nabídnout řešení při opakovaném neporozumění			x	x					
		Reagovat na ticho			x	x					
		Dát možnost opravit jakýkoliv vstup			x	x	x				
		Upřednostnit opravu před pokračováním ve scénáři			x	x	x				
		Potvrzovat, že klientově řeči rozumí	x	x		x	x				
		Dát možnost doptat se na to, čemu člověk nerozumí		x	x					x	
		Napovědět člověku, co může odpovědět	x		x	x	x				
		Komunikovat jednoznačně		x	x			x		x	
		Umět vyhodnotit situaci klienta		x		x	x			x	
		Přizpůsobit se expertize člověka				x	x				
		Předvídat důvod hovoru		x		x	x	x		x	
		Vyjadřovat se stručně a věcně, být efektivní			x			x			
		Být univerzální, ale konkrétní			x			x			
		Mít osobnost		x	x				x	x	
		Informovat o nahrávání hovoru a zabezpečení dat		x	x						
		Mít konkrétní účel	x	x	x					x	
		Správně nastavit očekávání	x	x						x	
		Být transparentní a mluvit pravdivě		x				x		x	
		Dát vybrat maximálně ze tří položek				x	x	x	x		
		Informovat, že člověk hovoří s robotem				x	x	x		x	
		Položky z listu vyjmenovávat jednotně			x	x		x	x		
		Zařadit to nejdůležitější na konec promluvy	x			x		x	x		
		Informovat že může nastat pauza				x		x			
		Informovat za jakou společnost hovoří		x						x	
		Nevyžadovat diktování celého údaje		x		x	x	x			
		Místo obecných instrukcí použít konkrétní s příkladem				x	x	x			
		Reagovat kdykoliv na žádost o počkání				x			x	x	
		Reagovat kdykoliv na žádost o přepojení				x				x	
		Být asertivní								x	
		Oslovovat klienta jménem		x		x	x			x	
		Vhodně vykat či tykat		x	x	x			x	x	
		Umět vhodně ukončit konverzaci	x								
		Být zdvořilý		x		x			x	x	
		Reagovat na small talk v průběhu celé konverzace		x		x			x	x	
		Pracovat s humorem		x		x			x	x	

tab. 2 – konverzační pravidla

Konverzačních pravidel je podstatně více. Příkládám to tomu, že je tato diplomová práce zaměřena především na analýzu konverzace. Pokud bychom se blíže podívali na technologické pozadí hlasových asistentů, jistě bychom našli další požadavky, které by bylo možné přidat do první tabulky. Nejvíce pravidel (celkem 22) se protíná v dovednosti

přizpůsobit se příjemci. To příznačně navazuje na skutečnosti, že se voicebotům říká hlasoví nebo virtuální asistenti. Už z názvu vyplývá, že se jedná o robota, který nám má asistovat a tedy se co nejvíce přizpůsobit příjemci.

Další dvě dovednosti, které v tabulce měly nejvíce záznamů, jsou schopnost vybudovat důvěru (17) a vzbudit respekt (18). To interpretuji dvěma způsoby. Jednak spolu tyto dvě dovednosti úzce souvisí, takže se dané pravidlo pravděpodobně dotýkalo obou z nich. Z analýzy hovorů i z poznatků z literatury zároveň vnímám skutečnost, že pro některé lidi jsou hlasoví asistenti stále neznámí, mnozí se s nimi při analyzovaných hovorech mohli setkat vůbec poprvé (například starší lidé u virtuální zdravotní sestry). Myslím si tedy, že je nyní potřeba budovat důvěru v komunikaci s roboty, ale v blízké budoucnosti tato potřeba bude stále menší. To ostatně souvisí i s citátem, který uvádím v úvodu této diplomové práce, tedy, že se dobré technologie postupně začleňují do našeho každodenního života, až je bereme za automatické. Vzhledem k rychlosti rozvoje hlasových technologií v posledních letech si troufám tvrdit, že bude komunikace s nimi stále častější a proto se sníží i potřeba budovat respekt a důvěru.

### **4.3 Náměty pro další výzkum**

Návazným výzkumem může být kvalitativní či kvantitativní analýza hodnocení schopností virtuálního asistenta přímo lidmi, kteří s ním hovořili. V kvalitativní analýze by bylo možné zaměřit se na hloubkové rozhovory s uživateli, kteří prošli konkrétním hovorem. Bylo by možné vytvořit vzorek dle sentimentu hovoru, aby byly zastoupeni uživatelé, kteří jsou s robotem ochotni mluvit, stejně tak, jako ti, kteří mají vůči komunikaci s virtuální asistentem výhrady, či ji odmítají úplně. V kvantitativní analýze na toto téma by bylo možné zjistit dané informace prostřednictvím standardizovaného šetření zákaznické spokojenosti Net Promoter System (NPS), kdy lidé hodnotí svou zkušenost s konkrétní službou. Toto šetření by mohl provést stejný virtuální asistent, se kterým lidé hovořili, jako součást daného hovoru. V takovém případě by se na konci hovoru mohl zeptat, jak s komunikací s ním byli lidé spokojeni, zdali měli pocit vybudování důvěry, jestli dobře vedl konverzaci a dal jim možnost oprav, zdali měli pocit, že jim dobře naslouchá a podobně.

Námětem pro další výzkum je analýza konverzací provedených přes jiný kanál, než je telefonní hovor. V hovorech je totiž omezen přenos dat, takže dochází k automatickému

snižování kvality přeneseného zvuku. Je tedy možné, že při konverzaci s virtuálním asistentem, který je součástí chytrého telefonu, domácnosti či auta, je menší riziko chyb při nepřesném přepisu. Vzhledem k situačnímu kontextu zde bude pravděpodobně potřeba po jiném souboru schopností a dovedností. Je možné, že k hlasovým asistentům typu Alexa či Google Home si lidé mohou vybudovat důvěrnější vztah. Komunikace také pravděpodobně bude mít i jiný, než pouze transakční účel, nebude potřeba, aby virtuální asistent byl schopen vzbudit v uživateli respekt. A nebo se zde bude pravděpodobněji objevovat požadavek na humor a small talk.

## 5 Závěr

V této diplomové práci jsem analyzovala konverzace lidí s hlasovými asistenty na telefonu, abych je porovnávala se současným stavem vědění, tedy poznatky z odborné literatury a také z provedených studií. Tyto poznatky by se neobešly bez teoretického ukotvení o tom, co je to konverzace, jak se na ni můžeme dívat, jak ji můžeme analyzovat a také, jak fungují hlasoví asistenti. Zaměřila jsem se také na to, jak lze analyzovat konverzace voicebotů a jak to udělat, abychom my lidé byli schopni s virtuálními asistenty komunikovat.

Z této analýzy jsem pak vytvořila model pravidel, která jsem převedla do matice technologických požadavků a konverzačních pravidel. Dohromady to tvoří funkční požadavky na voicebota, aby byl schopen konverzace. A to jak po technické stránce, tak také po té konverzační. Cílem bylo dát dohromady soubor takových doporučení, aby byla konverzace co nejvíce podobná té mezilidské. Tento model pravidel je nyní možné využít při reálné tvorbě virtuálního asistenta.

## 6 Seznam použité literatury

BENNETT, Gregory A. Conversational Style: Beyond the Nuts and Bolts of Conversation. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 161-180 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_8

BICKMORE, Timothy, Ha TRINH, Reza ASADI a Stefan OLAFSSON. Safety First: Conversational Agents for Health Care. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 33-57 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_3

CLARK, H. H. a S. E. BRENNAN. Grounding in communication. *Perspectives on socially shared cognition*. American Psychological Association, 1991, 127–149. Dostupné z: doi:10.1037/10096-006

CLARK, Leigh, Nadia PANTIDI, Orla COONEY, et al. What Makes a Good Conversation?. *Conference on Human Factors in Computing Systems - Proceedings* [online]. Ithaca: ACM, 2019, 1-12 [cit. 2022-07-05]. ISBN 1450359701. Dostupné z: doi:10.1145/3290605.3300705

DASGUPTA, Ritwik. *Voice User Interface Design Moving from GUI to Mixed Modal Interaction*. Berkeley, CA: Apress, 2018, 1 online resource (xi, 108 pages). ISBN 1-4842-4125-8. Dostupné z: doi:10.1007/978-1-4842-4125-7

DEIBEL, Diana a Rebecca EVANHOE. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, 2021. ISBN 1-933820-26-8.

DUBBERLY, Hugh a Paul PANGARO. What is conversation? How can we design for effective conversation. *Interactions Magazine*. 2009, **16.4**, 22-28.

ENFIELD, N. J. *How we talk: the inner workings of conversation*. New York: Basic Books, 2017, ix, 257 stran : ilustrace ; 22 cm. ISBN 978-0-465-05994-2.

GRICE, H. P. (H. Paul). *Studies in the way of words*. Cambridge (Mass.): Harvard University Press, 1991, vi, 394 s. ISBN 0-674-85271-0.

HARRIS, Randy Allen. *Voice interaction design crafting the new conversational speech systems*. San Francisco, CA: Morgan Kaufmann Publishers, 2005, 1 online resource (619 p.). ISBN 1-281-01658-6.

MARGE, Matthew, Carol ESPY-WILSON, Nigel G WARD, et al. Spoken language interaction with robots: Recommendations for future research. *Computer speech & language* [online].

LONDON: Elsevier, 2022, **71**, 101255 [cit. 2022-07-06]. ISSN 0885-2308. Dostupné z: doi:10.1016/j.csl.2021.101255

MOORE, Robert J. a Raphael. ARAR. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. San Rafael: Morgan & Claypool Publishers, 2019, 1 online resource (318 pages). ISBN 1-4503-6304-0.

MOORE, Robert J., Margaret H. SZYMANSKI, Raphael. ARAR a Guang-Jie. REN. *Studies in Conversational UX Design*. Cham: Springer International Publishing, 2018, 1 online resource (204 pages). ISBN 9783319955797. Dostupné z: doi:10.1007/978-3-319-95579-7

NIJHOLT, Anton, Oliviero STOCK, Alan DIX a John MORKES. Humor modeling in the interface. *CHI '03 Extended Abstracts on human factors in computing systems* [online]. ACM, 2003, 1050-1051 [cit. 2022-07-12]. ISBN 9781581136371. Dostupné z: doi:10.1145/765891.766143

PANGARO, Paul. Questions for conversation theory or conversation theory in one hour. *Kybernetes*. 2017, **46,9**, 1578-1587.

PEARL, Cathy. *Designing voice user interfaces: principles of conversational experiences*. O'Reilly Media, 2016. ISBN 978-1-491-95541-3.



PIERACCINI, Roberto. *The voice in the machine: building computers that understand speech*. Cambridge, Massachusetts: MIT Press, 2012, 1 online resource (355 p.). ISBN 1-280-49915-X.

RADZIWILL, Nicole a Morgan BENTON. Evaluating Quality of Chatbots and Intelligent Conversational Agents. *Software quality professional* [online]. Milwaukee: American Society for Quality, 2017, **19**(3), 25 [cit. 2022-07-25]. ISSN 1522-0540.

RUAN, Sherry, Jacob WOBROCK, Kenny LIU, Andrew NG a James LANDAY. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. *Proceedings of ACM on interactive, mobile, wearable and ubiquitous technologies* [online]. Ithaca: ACM, 2018, **1**(4), 1-23 [cit. 2022-07-25]. ISSN 2474-9567. Dostupné z: doi:10.1145/3161187

SACKS, Harvey, Emmanuel SCHEGLOFF a Gail JEFFERSON. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Sotsiologicheskoe Obozrenie / Russian Sociological Review* [online]. 2015, **14**(1), 142-202 [cit. 2022-07-10]. ISSN 1728-192X. Dostupné z: doi:10.17323/1728-192X-2015-1-142-202

SARIKAYA, Ruhi. The Technology Behind Personal Digital Assistants: An overview of the system architecture and key components. *IEEE signal processing magazine* [online].

PISCATAWAY: IEEE, 2017, **34**(1), 67-81 [cit. 2022-07-25]. ISSN 1053-5888. Dostupné z: doi:10.1109/MSP.2016.2617341

SEARLE, John. Chinese room argument. *Scholarpedia journal* [online]. 2009, **4**(8), 3100 [cit. 2022-07-25]. ISSN 1941-6016. Dostupné z: doi:10.4249/scholarpedia.3100

SCHEGLOFF, Emanuel A. a Harvey SACKS. Opening up closings. *Semiotica*. 1973, **8**(4). Dostupné z: doi:10.1515/semi.1973.8.4.289

STILLMAN, Daniel. *Good talk: how to design conversations that matter*. Place of publication not identified: Management Impact, 2020, 1 online resource (265 pages). ISBN 94-6276-379-8.

STRAUSS, Anselm L., Stanislav JEŽEK a Juliet M. CORBIN. *Základy kvalitativního výzkumu: postupy a techniky metody zakotvené teorie*. Boskovice: Albert, 1999, 196 stran : tabulky. ISBN 80-85834-60-X.

SZYMANSKI, Margaret H. a Robert J. MOORE. Adapting to Customer Initiative: Insights from Human Service Encounters. *Studies in Conversational UX Design* [online]. Cham: Springer International Publishing, 2018, 19-32 [cit. 2022-07-25]. ISBN 3319955780. ISSN 1571-5035. Dostupné z: doi:10.1007/978-3-319-95579-7\_2

WEISER, M. The computer for the 21st Century. *IEEE pervasive computing* [online]. New York: IEEE, 2002, 1(1), 19-25 [cit. 2022-07-25]. ISSN 1536-1268. Dostupné z: doi:10.1109/MPRV.2002.993141

WOOFFITT, Robin a Norman FRASER. *Humans, Computers and Wizards: Human (Simulated) Computer Interaction*. CRC Press/Balkema, 1997. ISBN 9780415069489.

## 7 Seznam obrázků

obr. 0 – příklad analýzy konverzace	15
obr. 1 – detail konverzace s chatbotem Eliza	29
obr. 2 – detail aplikace	48
obr. 3 – detail hovoru (scénář 1)	52
obr. 4 – detail hovoru (scénář 1)	53
obr. 5 – detail hovoru (scénář 1)	53
obr. 6 – detail konverzačního stromu (scénář 1)	54
obr. 7 – detail hovoru (scénář 1)	54
obr. 8 – detail konverzačního stromu (scénář 1)	55
obr. 9 – detail skriptu (scénář 1)	56
obr. 10 – detail konverzačního stromu (scénář 1)	57
obr. 11 – detail hovoru (scénář 1)	58
obr. 12 – detail hovoru (scénář 1)	59
obr. 13 – detail hovoru (scénář 1)	60
obr. 14 – textace (scénář 1)	60
obr. 15 – detail hovoru (scénář 1)	61
obr. 16 – detail hovoru (scénář 1)	61
obr. 17 – detail hovoru (scénář 2)	63
obr. 18 – detail hovoru (scénář 2)	63
obr. 19 – detail hovoru (scénář 2)	63
obr. 20 – detail hovoru (scénář 2)	64
obr. 21 – detail hovoru (scénář 2)	65
obr. 22 – detail hovoru (scénář 2)	66
obr. 23 – detail hovoru (scénář 2)	66
obr. 24 – detail hovoru (scénář 2)	67
obr. 25 – detail hovoru (scénář 3)	68
obr. 27 – detail hovoru (scénář 3)	69
obr. 28 – detail hovoru (scénář 3)	69
obr. 29 – detail hovoru (scénář 3)	69
obr. 30 – detail hovoru (scénář 3)	70
obr. 31 – detail hovoru (scénář 3)	71
obr. 32 – detail hovoru (scénář 2)	79