

IMPROVCHAT: AN AI-ENABLED DIALOGUE ASSISTANT
CHATBOT FOR ENGLISH LANGUAGE LEARNERS (ELL)

by

Yawen Guo

April 2018

A thesis exhibition and defense presented to OCAD University in partial fulfillment of the requirements for the degree Master of Design (MDes) in Digital Futures. 49 McCaul, 11th - 19th April 2018, Toronto, Ontario, Canada, April, 2018

Yawen Guo: *ImprovChat: An AI-enabled Dialogue Assistant Chatbot for English Language Learners (ELL)* (2018)

© This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>.

Creative Commons Attribution-NonCommercial 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution-NonCommercial 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

You are free to:

- Share — copy and redistribute the material in any medium or format
- Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- NonCommercial — You may not use the material for commercial purposes.
- No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

- You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.
- No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I authorize OCAD University to lend this thesis to other institutions or individuals for scholarly research. I understand that my thesis may be made electronically available to the public. I further authorize OCAD University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for scholarly research.

Signature _____

ABSTRACT

ImprovChat: An AI-enabled Dialogue Assistant Chatbot for English Language Learners (ELL)

Master's degree of Design 2018

Yawen Guo

Digital Futures

OCAD University

This thesis asks, how can an AI-enabled Dialogue Assistant in the form of a chatbot suggest new and unexpected forms of response to English Language Learners? The thesis draws upon theories of improvisation and play, alongside computer science to facilitate text-based conversation in English for non-native speakers. ImprovChat is a digital solution for communication issues that I encounter every day. My experience with improvisational theatre also informs the ImprovChat project. For my thesis research-creation, I developed the web application by implementing API powered by machine learning to generate AI phrases. The phrase generation model is pre-trained using the Twitter feeds created by people who speak English as a first language. ImprovChat dialogue assistant will provide phrase suggestions; these could potentially be used in online chatting scenarios where the participants have two different native languages. It can also be used to inspire unexpected and playful forms of response.

Keywords: AI, machine learning, dialogue assistant, conversation, improvisation, play, English Language Learning, ELL, chatbot

ACKNOWLEDGEMENTS

Thanks so much to my supervisor Judith Doyle and co-supervisor Adam Tindale. Their inspiring guidance and invaluable constructive criticism has helped me throughout the time it has taken to research and write this thesis. I have been so happy to work with you. Thanks also to Jessica Wyman and Phoebe Wong for providing writing strategies for the text, and Rebecca Martin for editorial assistance.

Thanks for my lovely cohort who created a friendly, diverse research environment together. I learned a lot from you and sincerely wish you all will have a bright future.

I also sincerely appreciate the help from Christoph Rühlemann and Matthew Brook O'Donnell who generously shared their linguistic research data for my project as the training corpus for the machine learning model.

Thanks for Feng Liu's design suggestions for the user interface and technical advice from Yang Lu. I also thank the six "anonymous" participants for my user testing whose feedback will inform the future directions of my research.

And my parents, love you both!

CONTENTS

1	INTRODUCTION	1
1.1	Research Questions	4
2	RELATED WORKS	7
2.1	A Playful Attitude and Learning	7
2.2	Improv and Conversation Competencies	8
2.3	Improv and Second Language Acquisition	9
2.4	Artificial Intelligence - Human Intelligence Collaboration	11
2.5	Dialogue System and Chatbot Facilitated Language Learning	12
2.6	Related Works	14
3	METHODOLOGY	17
3.1	Practice-Based Research	17
3.2	Autoethnography	17
3.3	Research through Design	18
3.4	Methods: Software Development	19
3.4.1	Prototype One	19
3.4.2	Prototype Two	24
3.4.3	Corpus	24
3.4.4	Chatterbot Library	25
3.4.5	Backend: Django and Twilio	27
3.4.6	User Interface	27
3.5	Methods: User Testing	29
3.5.1	User Interaction Feedback	29
4	FINDINGS	33
4.1	Artificial Intelligence	35
4.2	Artificial Stupidity	36
4.3	Research	39
5	CONCLUSION	41
5.1	Future Directions	41
6	WORKS CITED	43
A	REB APPROVAL	49
B	REB CONSENT FORM	51
C	LETTER OF INVITATION	53
D	LETTER OF FEEDBACK	55
E	QUESTIONNAIRE FOR ELL	57
F	QUESTIONNAIRE FOR NES	59
G	USER FLOW OF PROTOTYPE ONE	61
H	DIALOG SYSTEM EXPERIMENTS	65

LIST OF TABLES

Table 3.1	Sub-dataset of the results in “Five Things”	23
-----------	---	----

LIST OF FIGURES

Figure 1.1	I feel I am robotic	2
Figure 1.2	a-because-b	3
Figure 1.3	The chatbot helped me to improvise a funny dialogue	4
Figure 2.1	a-then-b relationship	9
Figure 2.2	SLA Filter Hypothesis	10
Figure 2.3	Dialogue system diagram	13
Figure 3.1	Chatterbot Process Flow Diagram (“About Chatterbot”)	26
Figure 3.2	Architecture of the web application	27
Figure 3.3	Sketches of the chatbot icons	28
Figure 3.4	Sketches of the user interface and user interaction . .	28
Figure 3.5	Showcase of the UI design	28
Figure 4.1	What should I say now?	33
Figure 4.2	I don’t realize that I can say that	34
Figure 4.3	A traditional way of learning new phrases	34
Figure 4.4	My AI dialogue assistant makes chatting much more enjoyable	35
Figure 4.5	AI dialogue assistant join in the SLA	36
Figure 4.6	Different-able AI	38

1

INTRODUCTION

ImprovChat fuses ideas of improvisational theatre with chat, including chatting and chatbots. Improvisational theatre, or improv, is a form of acting where performers act spontaneously by accepting random suggestions from the audience, without prepared scripts. Behind the scenes, improv teaches beginners “to play games with good nature, and to fail gracefully” (Johnstone, “Improv” 22). With years of training and improvisational experiences, Keith Johnstone, the founder of ‘Theatresports’ concludes that improv can help people “develop storytelling skills (these are more important than most people realize), improve interpersonal skills and encourage a life-long study of human interaction and improve ‘functioning’ in all areas” (23).

ImprovChat aligns with other applied improvisation¹, as it is designed to apply techniques from improvisational theatre for non-performing conversational settings. It is not a surprise that improv methods have been applied to many other fields such as consulting, teaching, generating ideas and therapeutic settings (“Applied Improvisation”). For example, the Second City Comedy Company provides workshops called ‘Improv for Anxiety’. Like other improv training classes, these workshops create a safe and nurturing environment for the participants and help students improve their communication and social skills (“Improv For Anxiety”) by reading verbal and non-verbal cues, assessing emotions, storytelling, understanding other perspectives, and ultimately building confidence and connections. In their book, *Improvisation*, John Hodgson and Ernest Richards explain how improv can influence the participants in a section entitled, “Living and Responding.” They write,

Because (improvisation) places people in a human situation involving other people, it calls for fairly quick thinking and at times for different levels of thought at one and the same time. Decisions have to be made by the individual in the situation, but because it is an experimental situation, he can learn by his errors or adjust to the utilization of his mistakes. (23)

Theories of second language acquisition suggest that various forms of inputs from conversational English are necessary. For instance, a learner can acquire English by picking up conversational phrases from movie dialogues, TV shows, and interview programs. A friend of mine talked about her immigrant parents, originally from China two decades ago, and how they learned to perform better in their English conversations. Her father would write down the catchphrases and slang from TV shows and practice many times by himself until he fully understood and was confident enough to use them. He would use those phrases deliberately and adequately in front of his family, sometimes leading to hilarious situations. Her mother seemed to have more issues with her accent, so she tended to have a clear phonetic

¹ Applied Improvisation employs the principles of improvisational theatre in non-theatre settings: life, work, relationships, group dynamics, and the like (“Applied Improvisation - Anima Learning”).

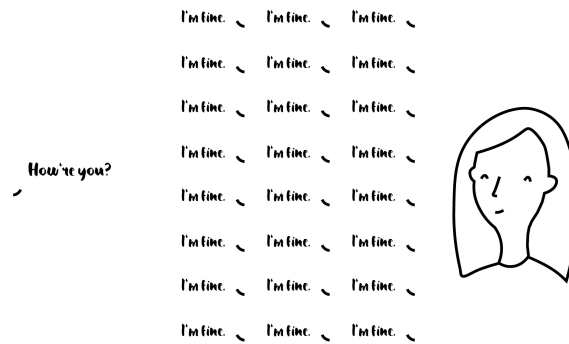


Figure 1.1: I feel I am robotic

awareness. She liked to associate words that sound similar and use them deliberately in one sentence, thus to make amusing effects. These experiences resonate with many ELL learners who are adequate enough speakers to communicate information, but want to add one more layer of emotional impact or personality on top of the plain words that they communicate.

Another block for speaking English is the social pressure. Researchers have found that adults are often “embarrassed by their lack of mastery of the language and they may develop a sense of inadequacy after experiences of frustration in trying to say exactly what they mean” (Lightbown and Spada 67-73). This embarrassment can turn into social anxiety for people who must communicate in another language, for example, immigrants living in a country with a different dominant language. A chatbot could provide a supplement that would relieve this sense of stress and inhibition.

I have been learning for ten years, and I still find speaking English to be a very challenging task. I have thought of myself as very robotic (Figure 1.1) when having conversations in English. The reason lies in that I only have one answer to a certain question and my means of expression tends to be simply listing my knowledge. An important goal for me at this stage is to “handle even a slightly challenging unofficial conversation and, orally or in writing, to tell about everyday things with a little detail” (26) as well as to add a bit of my sense of humor. Conversational harmony is vital to me because I want to have better interactions with friends who can only speak English. Though making jokes or amusing friends is not the primary goal for communication, I feel more accepted by the dominant cultural group² when I can express myself clearly, in more diverse ways, and also share in humor.

I joined my first improv workshops in 2013 when I wanted to get more practice speaking English. Through practice, I found that it is possible to draw parallels between the conversational settings for an ELL learner and the Improv game settings for a new improviser. Communication is essential to improv. After the first step of building confidence, the next level for the players is “to listen as well as talk and to respond to the needs of others” (Hodgson and Ernest 66). Communication skills including “keep(ing) the others talking in turn in an evenly balanced conversation” (66) can be practiced and coached through improv. One common warm-up game is when players are asked to develop the conversation from a given suggestion, us-

² Moyer’s study shows that acquiring more confidence in using the second communication make immigrants feel more like they were part of the L2 culture (Moyer).

I like man/woman as I like **a** because **b**

suggestion

player's response

Speaker1: I like a woman as I like a table because they both have four limbs.

Speaker2: I like a man as I like a table because both of them are stable.

Figure 1.2: a-because-b

ing a fixed phrase, for example, 'Yes, and...'³. This game is easy and fun to play and can be applied to daily conversations once a player learns the technique of accepting what is said and doing something with it. It is a way to support the other speaker as well as to encourage ourselves to speak.

Another challenging group game I played was to develop the sentence using the frame "I like a woman/man as I like 'a'⁴, because 'b'". This game builds the player's metaphorical skill. In our game, the given suggestion is 'cigar'. I contributed one sentence: "I like a man as I like a cigar, because both are my hobby!" My responses surprised myself for two reasons: Firstly, the metaphor comes so easily. Secondly, I dared to talk about a private topic on stage without the burden of self-consciousness! Improviser Katie Novotny shared a similar story, and she said, "I was shocked those words came out of my mouth.... I was the high-status hero. I played to the height of my intelligence. I did it. After that class, I realized that I shouldn't only be playing to the height of my intelligence when I'm on stage, but in every aspect of my life" (Carrane 27). My experience on the improv stage echoed with Katie and motivated me to connect improv with daily life.

Improv games break down the art of language into techniques, using funny and achievable exercises. Players can focus on building one skill at a time. By playing the phrase games, we begin to recognize specific patterns. Players stretch their imaginations, thus gaining the ability to use more creative expressions. Gerald Graff and the coauthors explained the benefit of phrase exercises:

After all, even the most creative forms of expression depend on established patterns and structures. . . . Even the most avant-garde, cutting-edge artists (like improvisational jazz musicians) need to master the basic form that their work improvises on, departs from and goes beyond, or else their work will come across as uneducated child's play. Ultimately, then, creativity and originality lie not in the avoidance of established forms but the imaginative use of them. (Graff et al. 10)

I started to practice my speech and communication skills by taking improv classes at the Second City Comedy School in the spring, 2017. The art of speaking and using phrases creatively filled in my mind again. I improvised in a dialogue with a friend and how did the AI technology caught my eyes (Figure 1.3) and inspired me to use the chatbot for my communication issues. To further explore using the chatbot to demonstrate new and different creative expressions, I continued to test chatbot generated phrases

³ "Yes, And..." is game but also the most common method of advancing scenes in an improv skit: accept everything said and done, and do something with it ("Improv Encyclopedia - Yes And"). It is one of the foundational concepts for improv that "scenes spontaneously generate themselves if both actors offer and accept alternately (Johnstone, 'Improv' 99)."

⁴ "a" is the suggestion from the audiences, "b" is the reason the player need to come up with. Figure 1.2 shows us two examples when the audiences give the suggestion "table".

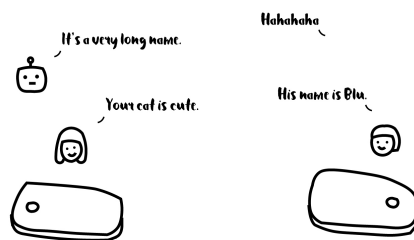


Figure 1.3: The chatbot helped me to improvise a funny dialogue

in various online chatting scenarios with people. The bot would prompt me with new phrases, for example, “so many memes, so little time.” Some phrases make me laugh, for they show a sense of humor. Though it was the first time I saw these unexpected responses, I immediately remembered all of them. I assume that I can remember those phrases very efficiently because I physically typed them out and therefore mentally engaged in using them. In doing so, the phrase and the experience become related, so that I can more easily recall them. On most occasions, the chatbot text made no sense and could be offensive, so I could not rely on it completely. That being said, this way of interacting with native English speakers successfully taught me new, advanced phrases that I would never find in a book. Attaining new knowledge of English has never been as much fun for me with this approach. To put it simply, one moment of improvisation in my life inspired me to make an AI dialogue assistant⁵ to build my capacity for creative expressions in English. I was eager to explore the phenomenon.

The computer has played an ever-increasing role in the field of language learning with the maturity of Natural Language Processing (NLP) technology. As stated by Nerbonne: “NLP focuses on how computers can best process language, analyze, sort, and search it. It seems natural that NLP should be applied to the task of helping people learn language” (Nerbonne 678). Integrating the insights from NLP and corpus linguistic makes the computer assisted language learning possible and desirable (Antoniadis et al. 89). In my research and prototype, I hypothesize that the chatbot⁶ can facilitate conversation, filling a gap between inadequate English knowledge and the conversational urge experienced by the second language learners.

1.1 RESEARCH QUESTIONS

- How can improv games facilitate practice-based language learning? How can I transform the useful parts of improv games into a digital form?

⁵ Trouvain stated that in the ideal computer assisted language learning settings, in his research specifically the computer assisted pronunciation training, would adopt the best practices which is teacher assisted, book assisted, audio-assisted and dialogue partner-assisted (Trouvain et al 1).

⁶ The chatbot, or say dialogue systems been built from the data-driven approaches is feasible (Serban et al. 1).

- What is the best approach for creating a dialogue assistant embedded with intelligence from the theory and techniques of improvised games? How can the AI dialogue assistant make the content creative and rich? Can an AI textual dialogue assistant help generate humor through its unpredictable responses?
- How can an AI dialogue assistant (enabled to execute embedded AI phrases) facilitate chat in English for non-native speakers in their daily communication? Will using an AI dialogue assistant enable better acquisition of English (for example, by prompting multiple options of response to a chat)?

2 | RELATED WORKS

As a component of this thesis, I reviewed contemporary literature on a number of relevant subjects. This chapter outlines these subjects and identifies vital concepts that have influenced my design work. The first section starts to look at my prototype within the context of contemporary theories on and explanations of play. The second and third sections address the central concept of improvisation performance and a few related sub-themes such as applied improvisation, second language acquisition, conversational competence and speech therapy. The fourth and fifth sections focus on the computational vision for the development of the prototype which covers the artificial intelligence, chatbot and dialogue system. The last section reviews the related projects that deployed machine learning models in order to synthesis from different approaches. The above fundamental pillars outline the theoretical framework of the research presented in this paper. The iterative design methodology applied to this framework enables the design of the AI dialogue assistant that serves as a bridge between my desire to express my personality and my ability to express myself creatively in English.

2.1 A PLAYFUL ATTITUDE AND LEARNING

In their book, *Improvisation*, Hodgson and Richards write that, "Every improvisation is to some extent a play" (103). Improv pioneers acknowledge that play is a way to "make learning a beloved activity" ("Geniuses"). **Keith Johnston** shared a story in his book *"Impro for Storytellers"*: Benjamin Constant was aged four when his tutor suggested that they invent a language. They went around the estate, naming everything, and working out a grammar. They even invented special signs to describe the sounds. Ben was aged six before he discovered that he'd learned Greek (Introduction). Spolin's approach was directly influenced by Neva Leona Boyd. Boyd's innovation in education "explored the constructive potential of play, using games with inner city and immigrant children to help them adjust to society (Scott 119)." Spolin emphasizes that "games develop personal techniques and skills necessary for the game itself, through playing. Skills are developed at the very moment a person is having all the fun and excitement playing a game has to offer — this is the exact time one is truly open to receiving them" (Spolin 5).

Given the above, "play" is a very important topic to investigate in my research. "Play is the act of creatively engaging with the world, with technologies, contexts, and objects, from games to toys and playgrounds, exploring them through ludic interaction" (Sicart 17). It is designed as mediated by things created to facilitate the emergence of play (7) and "what we want is the attitude of play without the activity of play" (21). Play involves competition, and there is a risk of failure. However, embracing the failure is also part of the play and is especially encouraged in improv. Improvisational

theatre provides a non-judgmental playground for people to practice their creative expression. Theatresports founder Keith Johnstone reflects on his coaching: "I tell improvisers to follow the rules and see what happens, and not to feel in any way responsible for the material that emerges" (Johnstone, "Improv" 111). During the play sessions, participants expressed themselves. In a word, improv can bring playful attitude to learning experience.

The idea of ImprovChat comes from an improvised activity. ImprovChat is designed to facilitate the emergence of funny conversations. Also, the desire to bring the fun part into the long journey of gaining ability in English motivates me to create the ImprovChat. With the help of the dialogue assistant designed to know creative expressions in English, I explore what I can say and how I can say things in the second language domain. It also reflects an introvert's childhood experience in lack of partners whom I can play dialogue with. When I encountered difficulties expressing myself through conversations in daily life, the creative part of me which desires to interact with people in a playful way (mainly by adding humour) was also blocked. When I took improv workshops, this desire awakened, and the first shock I got was to consider, "why is it so hard for me to be happy?" After the literature review, I ask the question "why is it so hard for me to play improv games?" Surprisingly, as an adult, I have to take another look at "play" and possibly I need to learn how to play. Surprising as this may be, this idea of adults gaining the ability of play is not new to researchers coming from the play theory background. Bateson and Martin argue that "when play is overt, it may be seen more often in children than in adults, but adult humans are perfectly capable of playing and, sometimes benefit from it" (Bateson and Martin 5). They suggest that "humans and organizations can exploit playfulness as a tool for fostering creativity and innovation" (5).

Since "almost any space can become a playground" (Sicart 7), it is proper to use the playground as the metaphor referring to the social environment where the conversations happen. Online chatting scenarios are my playgrounds for testing phrases and dialogues. With the facilitation of the customized dialogue assistant, I can enter the playground confidently, even though I am not fully capable of using the tool-English. Knowing that I have the backup from an English knowledge powered dialogue assistant, I can simply become immersed in the conversation and, "play for the sake of playing" (26).

2.2 IMPROV AND CONVERSATION COMPETENCIES

In his book, *Improvising Improvisation*, Gary Peters defines the origin of improvisation as "the desire to begin something for the sake of the beginning and not the end" (7). Pioneers of Improv call it spontaneity. Adding to this, Viola Spolin writes that, "the intuitive can only respond in immediacy — right now. It comes bearing its gifts in the moment of spontaneity, the moment when we are free to relate and act, involving ourselves in the moving, changing world around us" (1).

From a linguistic perspective, the "Yes, and..." game employs what William Labov and Joshua Waletzky describe as an "a-then-b"⁷ (30). This relationship is seen by Labov and Waletzky as the "defining characteristic of narrative" (22). By this they mean that the players in "Yes, and..." games are required to use at least two temporally junctured narrative clauses in a sen-

⁷ "a-then-b" relationship can be illustrated as Figure 2.1 (Rühlemann and O'Donnell 318).

a then b

S1: yeah no, what happen was the alarm went off in their block or something and they all had to go into a room. I don't know (KPY-N1)

Figure 2.1: a-then-b relationship

tence (Rühlemann and O'Donnell 317-318). Thus, the conversational narrative will be formed in the "Yes, and..." game naturally. The practice of conversational narrative will also help the ELL learner in their daily conversation. My speech therapist even suggested that I "always ask for a story, not just a simple answer to a question." So the question, "what did you do today?" is slightly better than, "how are you?" because the former gives the speaker a chance to tell a story about herself/himself. Thus, the conversation will flow on its own. Since the speaker's brain is engaged in searching for story details, the speaker will be released from the stress of "what to say next" without any cues. Hodgson and Ernest write that "[w]hat is required during improvisation is thinking within a situation or thinking about a situation after it has been experienced (23)." Improvisation is the act of real-time, dynamic problem solving (Magerko et al. 1). In the essay "'Improvisation'. A 'Method Of Philosophical Consultation,'" Neri Pollastri has pointed out that there are two important elements to the improvisational method: the practice of skills and the reference to a corpus of tradition (3). In other words, while "previously learned material plays a central role in improvisation," (Sterritt 166) this alone is not enough—the practitioner has to, "'embody' those knows and competencies"(Pollastri, 3). The ultimate goal then is to supplant both instinct and knowledge with habit. The advertisement "Kids Secretly Control a Speed Dating Session" on YouTube (Mini-Wheats Canada) shows us a good example of how improvisation can be applied to a conversation to create an amusing effect. In this designed scenario, a woman actively takes unusual suggestions from kids hidden in a near-by room, and uses her imagination to respond to the kids' prompts to facilitate a series of strange activities during the speed dating sessions. The application of improvisation to the situation made her interactions with the other daters very amusing and memorable.

2.3 IMPROV AND SECOND LANGUAGE ACQUISITION

Second language learning occurs in a natural environment, normally in a country where the language is spoken. Second Language Acquisition⁸ (SLA) emphasizes the degree to which automatic processes may play a role in the more natural situations where a language from the immediate environment is adopted (Nerbonne 678). Acquisition is responsible for our fluency in second languages, while learning serves only as a monitor, or editor: that is, we use our conscious knowledge of rules only to make corrections, either be-

⁸ Second Language acquisition theory distinguishes subconscious language acquisition from second language learning. Acquisition is hypothesized to be subconscious in the sense that while it is happening we are not usually aware of it; our focus is elsewhere, on the message that is being communicated. The results of acquisition are also subconscious — we cannot always describe our acquired knowledge, but, rather, have a "feel" for correctness in a language we have acquired. Conscious learning is "knowing the rules," or explicit knowledge. In every speech, it is "grammar." (Kreshan, "Principle" 14-15)

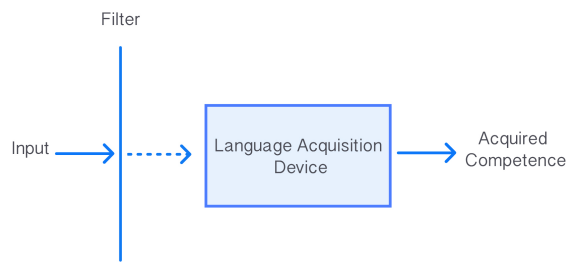


Figure 2.2: SLA Filter Hypothesis

fore or after we produce our sentence in the second language (Krashen, "The Din in the Head" 43). SLA reveals why having conversations has helped me to acquire knowledge of English. As indicated by this section's heading, in this section I will explore how improv and technology can enhance the process of language acquisition. Ability in a second language can be developed. As Terry Winograd and Fernando Flores write in their book *Understanding Computers and Cognition*,

Communicative competence means the capacity to express one's intentions and take responsibilities in the networks of commitments that utterances and their interpretations bring to the world.... People's conscious knowledge of their participation in the network of commitment can be reinforced and developed, improving their capacity to act in the domain of language. (162)

Adding to this, it is essential to know how one might gain this ability. Linguistics professor Stephan Krashen developed a theory about second language acquisition, first published in 1982, that has been confirmed by subsequent research. Krashen points out that compared to "language learning," there is another very different way to gain ability in another language, a method that he terms "language acquisition." According to Krashen, comprehensible input and the strength of the filter⁹ are the true causes of second language acquisition ("Principles" 31). Getting comprehensible inputs from various sources is good for language acquisition (Krashen, "hypothesis" 299).

During a chat with an English instructor who actively applied a communicative approach (rather than memorization) to pedagogy, she shared a very interesting insight from her teaching experience. My instructor said that she could tell that some students like to copy the conversational model been taught to them, while other students are able to improvise dialogue. One important sign of improvising dialogue is the speaker's spontaneity. This caught my attention immediately because "spontaneity" is the key el-

⁹ Filter hypothesis, represented in Figure 2.2, claims that the effect of affect is "outside" the language acquisition device proper. It still maintains that input is the primary causative variable in second language acquisition, affective variables acting to impede or facilitate the delivery of input to the language acquisition device. Most affective variables relate to success in second language acquisition can be placed into three main attitudinal factors (motivation, self-confidence and anxiety). Krashen has found that these factors tend to show stronger relationships to second language achievement when communicative-type tests are used (Krashen, "Principles" 30).

ement for the improv theatre that I discussed it in the previous section. I asked her "what does spontaneity mean in your class?" She replied that, "to be spontaneous means using what you know for a new situation." For example, a learner would normally know basic phrases like "It is good/-great/bad....", "I want", "I use", "how is it going" and "what you know" and these phrases would be a good start for them to join a class discussion. If a learner doesn't know how to say that the meat dish is too heavy and oily for him, he can simply express: "I'm full" to avoid being invited to eat more. A learner can be as creative as she needs to be to create her own words. For example, if she can't remember the word "refrigerator," she can refer to it as an "ice box" instead. The new learner shouldn't feel shy or ashamed that she/he doesn't have the perfect vocabulary and expressions, because playing with limited knowledge also means using the language in new ways — new ways that highly relate to the learner's own experience¹⁰. My instructor encourages her students to make mistakes because people can only help to correct you when you try to speak.

So, given the above, I propose that improv games can be useful for getting knowledge of conversational competence and SLA, "because much of the work involves us in talking, we gain confidence in a flow of intelligent speech....The use of language varies in different situations, and in the variety of the subjects of improvisation, there is a chance to realize how vocabulary, word order, and images vary from circumstance to circumstance" (Hodgson and Ernest 26). As a developer and designer, I believe there must be something that I can do to connect improv, English acquiring and the communication.

I remember that in one improvised performance session, I felt the urge to praise my partner – my "husband" in that play. Somehow, my brain was just a blank. Suddenly I saw my instructor doing a gesture to me which hinted that I should recall what I learned from the "Give me five things" game. I immediately knew to count five good things that my "husband" did. I looked at my partner and said: "Yes, you are such a good husband, because you washed all the dishes in the morning, you sent our kids to school on time and...." Recalling this experience, I realized how helpful and how smart my instructor was. The idea of having a digital assistant which could take his role and facilitate my daily conversation has grown in my mind since then.

2.4 ARTIFICIAL INTELLIGENCE – HUMAN INTELLIGENCE COLLABORATION

Artificial Intelligence, or AI, was first defined as the part of the computer sciences concerned with designing systems that exhibit the characteristics associated with human intelligence—understanding language, learning, reasoning, solving problems and so on (Barr and Feigenbaum 1: 8-11). This definition of AI was first proposed in the 1950s. From its definition, we know that AI has the ability to understand language and learning.

Winograd and Flores, in their book *Understanding Computers and Cognition*, formulate new theoretical foundations for the design of computer technology and bring up a new way of understanding both human and arti-

¹⁰ For example, my instructor said that she had several different accents. These accents came from her early experience of living in different cities.

ficial intelligence. They argue that “computers are not only designed in language but are themselves equipped for language. They will not just reflect our understanding of language, but will at the same time create new possibilities for the speaking and listening that we do — for creating ourselves in language” (162). Winograd and Flores propose an alternative conception of artificial intelligence in which we recognize that all computational systems can truly do is reflect the intelligence of the people who designed them. Thus, the true potential for computers lies not in the mimicry or duplication of human intelligence but in the facilitation of human-to-human communication (qtd. J. Tanenbaum and K. Tanenbaum 254).

AI development needs open source code and an open tool kit for developers, programmers and researchers to bring the knowledge into application. For example, the best programming languages for AI include Python, Java, and C++ which are all open source. To better serve researchers and users, some companies build open tool kits by combining programming language, basic code, products and so on. For example, Google launched Tensorflow¹¹ in 2015 and it is widely employed by users at different skill levels; Keda Xunfei, a multi-language conversation center which aids people in traveling around world by translating the information in another language to the user’s first language also launched state-of-art AI technology services in 2010. These include open sourced software development kits¹² (SDK) that have made it possible for me — a developer who does not have sophisticated AI background — to learn and build an AI project.

2.5 DIALOGUE SYSTEM AND CHATBOT FACILITATED LANGUAGE LEARNING

Experiments using dialogue systems formed through a computational agent — very often a chatbot as a tool to learn and practice a second language — have been conducted across the world. These kinds of experiments have resulted in varying degrees of failure, however according to Bayan Abu Shawar and Eric Atwell, despite the failures of experiments using a chatbot dialogue system, “users found it an interesting tool to practice the language and enjoyed chatting” (38.) This being the case, researchers “concluded that even with its key — word based matching technique, a chatbot could be used as a tool for unknown languages” (38). In some educational settings, dialogue systems have cognitive advantages for learning and memory (Bowden et al. 14). Students learned better when they interacted with an agent¹³ that utilized a personalized dialogue (Moreno et al. 6). The research implies further that a pedagogical agent should focus on extending the capabilities of speech recognition and natural language systems to facilitate instructional conversations (Moreno et al. 6).

11 Tensorflow is an open-sourced machine learning framework launched by Google.

12 Software development kit (SDK) provides a set of tools, libraries, relevant documentation, code samples, processes, and guides that allow developers to create software applications on a specific platform (Sandoval).

13 In computer science, a software agent is a computer program that acts for a user or other program in a relationship of agency. Software agents interacting with people (e.g. chatbot) may possess human-like qualities such as natural language understanding and speech, personality or embody humanoid form (“Software Agent”).

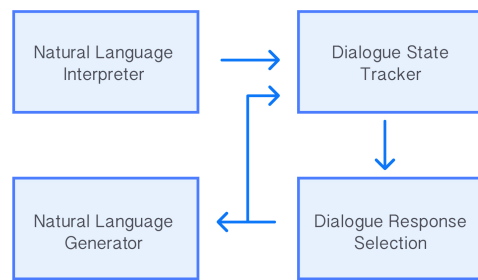


Figure 2.3: Dialogue system diagram

Dialogue systems¹⁴, also known as interactive conversational agents, virtual agents or sometimes chatterbots, are useful in a wide array of applications ranging from technical support services to language learning tools and entertainment (qtd. in Serban et al. 1). Many researchers have pursued building dialogue systems over the decades. Research on non-goal driven dialogue systems dates back to the mid-60s. It began, perhaps, with Weizenbaum’s famous program ELIZA, a system based only on simple text parsing rules that managed to convincingly mimic a Rogerian psychotherapist by persistently rephrasing statements or asking questions (Weizenbaum 36-45). Much of the recent progress is due to the availability of larger public datasets, increased computing power, and new machine learning models, such as neural network architectures” (Serban et al. 2). Data-driven approaches, which apply state-of-the-art machine learning technologies, contrast to systems where each component is hand-crafted by engineers (3). The standard architecture for text-based (written) dialogue systems, shown in Figure 2.3, incorporates a Language Interpreter, State Tracker, Response Generator, Natural Language Generator (Serban et al. 3). However, there is a so called end-to-end dialogue system architectures which does not conform to the architecture shown in Figure 2.3. In its purest form, these models take as input a dialogue in text form and output a response (or a distribution of responses). End-to-end dialogue systems can be divided into two categories: those that select deterministically from a fixed set of possible responses, and those that attempt to generate responses by keeping a posterior distribution over possible utterances (Serban et al. 6).

System architecture forms the bones while the training corpus for the system is the meat of my research. To facilitate further research on building data-driven dialogue systems, researchers conducted a broad survey of available dialogue corpora (Serban et al. 2). Programmers need to pick the proper corpus based on the design purpose. For example, the programmer may select the corpus called “Cornell Movie Dialogue Corpus”¹⁵ to train the dialog system for Pyggy - the artificial improviser (Mathewson and Mirowski 66). Google researchers use the corpus collected from the IT helpdesk troubleshooting service chatting log to make the chatbot capable of the best communication on the topic of “IT problems” (Vinyals and Le 3). In

¹⁴ Over the course of my research, I have relied mainly on Iulian Vlad Serban and his co-researchers’ investigations and complete overview of dialogue systems (Serban et al. 2) to acquire knowledge on the subject.

¹⁵ This corpus contains a large metadata-rich collection of fictional conversations extracted from raw movie scripts (Danescu-Niculescu-Mizil and Lee 76-87).

the machine learning field, engineers always say, “garbage in, garbage out,” which means a bad quality or improper training data will never reach an expected result. So, the corpus is very essential to the final result. Developing open-domain conversational dialogue systems is difficult, since the huge variety of user utterances makes it harder to build knowledge resources for generating appropriate system responses (Sugiyama et al. 334). Microblogs such as Twitter containing a wide range of topics are generally a good resource for making the open domain conversational system in recent years (Serban et al 30).

I give more detailed programming information of my dialogue system and the corpus related to my project in the Chapter 3.

2.6 RELATED WORKS

Smart Reply, available in Inbox by Gmail and Allo came out in March, 2017. This application saves users’ time by suggesting quick responses to emails (Strope and Kurzweil). Researchers who worked on this project found an effective path by comparing a simple hierarchy of vector representations of multiple features corresponding to longer time-spans (Strope and Kurzweil). Shortly after the release of the app, users started to post their user experience on their online blogs. One user carried out a 24-hour experiment on using the Google smart reply. She listed all of her email content and compared the smart reply suggestions to her own opinions of the suggested content. She described the experience as “weird” and finally decided to rely on her own typing instead of using the Smart Reply app (Santino). My friend Judith confided that she thought the smart replies were more polite than her own responses. I use Smart Reply sometimes, but usually follow up with my own words. Smart Reply itself does not convey the whole of my emotional attachment to any given correspondence. Though Smart Reply’s function shares similarities to my ImprovChat, the starting point of Smart Reply is totally different from the intended use of my chatbot. Smart Reply is designed for people who already use English fluently in their daily communication, while ImprovChat is designed for people who are lack language competency in English. So, Google researchers might assume that the three phrases, “How about tomorrow?”; “Wanna get together tomorrow?”; and, “I suggest we meet tomorrow,” are not very different options for users (Corrado). These three phrases, however, can show ELL learner three ways to express the same intent rather than sticking to one standard answer, and therefore they are useful for such a user. That’s the niche where ImprovChat fits.

In addition to the above, I also find inspiration from machine learning projects that deal with text generation, as well as AI assistants and collaborators. For instance, Robin Sloan, whose background is as a journalist and an author, has created an extensible text editor powered by machine learning technology. Sloan’s extensible text editor is responsive, and its inline autocomplete function is powered by a recurrent neural network trained on a corpus of old sci-fi stories (Sloan). Another example, Dango, is a floating assistant that runs on your phone and predicts emoji, stickers, and GIFs based on what you and your friends are writing in any App. For example, it suggests emojis of a crown and a bumble bee for the text “Beyoncé” because she is Queen Bee” (Snelgrove). This AI assistant helps people to weave imagery with text in their online chatting. The data-driven approach to emoji

prediction demonstrates how people around the world use visual language to communicate. Inspiration also comes to me from “Pyggy-Artificial improvisation model,” the first artificial improviser created by Kory W. Mathewson (Mathewson and Mirowski 1). Mathewson developed this chatbot that he can perform with during improvised performances and he and his collaborator Piotr Mirowski propose to use improvisation as a way to pass Turing Test.

In a conclusion, machine learning technology shows a promising ability for making tools to facilitate human to human communication and computing literature. My context review also indicates the fact that there are several software development kits that only require a short learning curve for development, thus the developer without a computer science background can also access and use the complicated machine learning algorithm. Following the context review, I conducted a series of self-studies based on the top rated online courses and the part time training courses offered by other experts. The whole learning experience made me believe that I could make a prototype. I also got a good sense of the AI industrial geography. I’m grateful for the good resources available in this booming technological city, Toronto.

3 | METHODOLOGY

ImprovChat engages both embodied practice-based research supported by autoethnographic journaling and research through design methodology structured by iterative design. These components weave together to form the foundations of my research.

I built my first prototype to be a digital version of the improv games called Give me Five Things. This prototype two uses the findings from prototype one in order to extend the ImprovChat's helper function. The findings at this stage guided me to recall the embodied practices that I undertook earlier in my research and to look back at my autoethnographic journals. I have started to see these embodied practice-based experiences as evidence that support my understanding of artificial intelligence and even artificial stupidity. I imagine the chatbot asking the human, "do you understand my humour?" or even welcoming the human to see the world through machine vision. I will talk more about these findings in the Chapter 4. This chapter will focus on the methods I deployed to build the prototypes and the data collection.

3.1 PRACTICE-BASED RESEARCH

Practice-based research implies an original investigation undertaken in order to gain new knowledge partly by means of practice and partly through the outcomes of that practice (Candy 3). Claims of originality and contribution to knowledge may be demonstrated through creative outcomes which might include artefacts such as images, music, designs, models, digital media and more (Candy 3). The research programme must consist of a continual reflection upon that practice and on the results that continually inform that practice (4). The embodied dimension of my own study includes participation in improv workshops where I took classes to get first-hand immersive experience with improv. My research is also informed by readings in improv theory and techniques. Finally, the development of the

ImprovChat prototypes engages iterative design including practice-based prototyping and programming.

3.2 AUTOETHNOGRAPHY

In *Composing Ethnography: Alternative Forms of Qualitative Writing*, autoethnography is envisaged by editors Carolyn Ellis and Arthur P. Bochner as a writing form of "creative nonfiction to take certain expressive liberties associated with the arts, but (also) to feel the ethical pull of converting data into experiences readers can use" (28). Ethnographer Ruth Behar asserts that "what happens within the observer" must be made known (6). In *The Oxford Research Encyclopedia of Education*, Susan Gannon points out that

one current in autoethnography is the experience of otherness, which suggests that autoethnographic form may be unique to each author's particular lived experience and influences (14). Tami Spry comments on her research experience that "for me, autoethnographic texts express more fully the interactional textures occurring between self, other, and contexts in ethnographic research" (708). In *Conversing Life: An Autoethnographic Construction* Christopher Hoelson, and Rod Burton catalogued selected moments before they "jointly allocated and organised them into meaningful clusters or categories to reflect the unitary nature and effects of the selected moments" (100). Following Hoelson and Burton, at the stage of exploring embodied-practice through improv, I wrote down important moments using journals, and by engaging in self-study through chat sessions, organized these moments into categories in order to understand the nature of these moments.

3.3 RESEARCH THROUGH DESIGN

The creative process is iterative. This means the ideas are developed, tested and refined a number of times, with weak ideas dropped in the process ("The Design Process: What Is the Double Diamond?"). The question of "Does it work?" is not approached as a calculation before construction, but as an iterative process of prototype testing and refinement. David Kelley, founder of the design firm IDEO, has stated that, "enlightened trial and error outperforms the planning of flawless intellect" (qtd. in Winograd 1257). Research through design (RtD) as one type of design research¹⁶ activity "was the process of iteratively designing artifacts as a creative way of investigating what a potential future might be" (Zimmerman et al. 313). It is notable that "new artifacts (products, environments, services, and systems) where the artifact is itself is a type of implicit, theoretical contribution" (313). The power of these artifacts has been described by Zimmerman in terms of how they codify the designers' understanding of the current state, including the relationships between the various phenomena at play therein, and the description of the preferred state as an outcome of the artifact's construction (313). In other words, while the main concern of RtD is to inform a research question, it also must be concerned with the end product of the design (Godin and Zahedi 1). Horst Rittel and Melvin M. Webber write that, "[t]he two most critical aspects of this approach are that it allows researchers to engage with wicked problems"¹⁷ (155-169) and "to become active constructors of possible futures (Forlizzi et al. 2894)." Some researchers think of RtD as a way of broadening the scope and focus of designers, of challenging current perceptions on the role and form of technology (Zimmerman et al. 310).

During the programming process, I scanned all possible software tool kits that support rapid prototyping and ran their starter kits before deploying a more complicated model. Multiple online tutorials and examples from other developers made short develop/test cycles possible. I demonstrated one prototype in the final presentation, but there are two more prototyping experiments in their second prototyping stage. Though these two experiments have not reached a polished form, they contributed to and deepened

¹⁶ Christopher Frayling first introduced the idea of three different forms of interactions between research and design / art: research for design, research into design and research through design (Frayling 5).

¹⁷ "Wicked problems" refers to problems beyond usability (for example, societal problems such as sustainability) that cannot be easily solved (Zimmerman et al. 310).

my knowledge of the research topic. This being the case, I have listed these two experiments in the appendix section.

3.4 METHODS: SOFTWARE DEVELOPMENT

3.4.1 Prototype One

The nature of improvised theatre relies on spontaneous generative conversational abilities. Improvised theatre training relies on teaching actors games that force them to perform fast-paced word associations (e.g., ‘electric. . . car. . . company’) or sentence without over-thinking any of their decisions (Mathewson and Mirowski 66). The first iteration of my project sought to find game rules that would be suitable to be transformed into the dialogue system or that would inform the dialogue system’s built-in improv formula.

The first prototype was developed using JavaScript, the Web Speech API for speech recognition, the Rita computational literature library for natural language processing and the Node.js backend framework for web applications. Using these programming tools, ImprovChat is intended for chat-based gameplay based on improvisation and association¹⁸ games. The process of development included programming followed by my own testing and journaling about these tests.

I picked two entry-level games from my practice with the Second City Improv class — “Five Things” and “Words that Sounds Similar.” Both of these games play with association and encourage the spontaneity (Merlin, “Word Association”; “5 Things”). In the game “5 Things”, one player initiates by pointing at another and asking him or her to say five things as fast as he or she can base on a category of his or her choosing. For example, one player points to someone and asks her/him: “Give me five songs you like.” As the player names off each one, the rest of the class counts along, cheering when five things have been said. The objective of this game is to revel in wordplay, to free your mind and commit to the moment (“5 Things - Canadian Improv Games”).

“Word sounds similar” is another fun word association game. This game asks players to vocalize their first impulse without hesitation. The objective is to connect players to their impulses and to free players from the perceived necessity to be creative (“Word ball - Canadian Improv Games”). As a part of this research project, I made digital versions of these two games (see Appendix G).

Over the course of playing these games for several hours, I documented how many new English words I was able to add to my vocabulary. In doing so, I am able to track how I am personally gaining— or not gaining—knowledge of English based on chat-based interactions. And, more importantly, I understood how does the game help me getting new knowledge of English. To measure the experiment results under the theoretical framework, I will answer the following two questions:

- How much do I engage in this game? How does the SLA theory reveal itself in each play session?
- What new thing do I learn from playing the game? What do I know that I didn’t expect to know?

¹⁸ “Association is one of the subcomponents of creativity that contribute to our intuitive understanding of what creativity is” (Jordanous 286).

- Is there any self-questioning happening in the play process?

I played my digital version of “Five things” from January 10th to January 13th, 2018 for three hours in total and I used screen capture software to record the whole process (See for instance ??). I had played this game in person, in Improv A before. The digital version was not as much fun as when I played in the Improv workshop for two reasons: firstly, because I selected the batch of data, I already knew the questions I would get from the chatbot, so even though the questions are randomly prompted, the game is not challenging for me. Secondly, there is no cheering from peers at the end of the game, which makes the experience less exciting. Despite these two flaws, testing the games gave me some new insights into what I want the dialogue system to achieve.

During the first hour of game play, I realized that the speech recognition system was extremely sensitive in recognizing the first input (the first thing I said). It could recognize almost all the words correctly — the system has a high confidence in recognition. The same word, for example, “goji berry,” would be recognized 100 percent of the time if it was the first input, while it might not be recognized at all if it was the third or fourth input. I would naturally repeat the same word until it could be recognized. Sometimes, the program simply failed to respond to me after too many repeats. Realizing this issue, I would say the complicated words first or second to avoid pausing of the system. As a participant, I tended to adapt my behavior to the system as to pass the game. Adding to this, the “Answer” button (which provides right answers) was alluring. I always wanted to click on the button to see the right answers, and in doing so I would expect to see new words. If I encountered a new word, I would first guess the pronunciation. If the computer could not recognize it, I would check an online dictionary for the definition, listen to the pronunciation, and speak the word to the system again. Thus, the happiest moments came from the computer recognizing a new word that I was able to say.

When it came to the second hour of interaction, I began to feel unwilling to answer questions which asked for objects under one category, such as: “Give me five fruits/vegetables/cars/herbs.” However, other questions such as: “Give me five household objects/pizza toppings/industries” are still interesting to me because these questions would usually have a broader set of correct answers available. I could simply look around my room and name the surroundings. I would picture a pizza in my mind to visualize the pizza toppings; visual association works well for this kind of questions. Adding to this, I very much enjoyed questions such as: “Give me five art-movements/music genres/types of beer/rooms.” The answers follow a pattern which an extension is added to a word that I already knew. For example, I know the words “folk music,” “rock,” “pop,” and it was interesting to learn how to extend these words into genres such as “freak folk”, “indie rock”, “Latin jazz”, “jazzy hip-hop,” and so on. The same pattern would apply for art movements — for instance, “post-minimalism,” “Russian futurism,” “photo-realism,” “video game art,” and so on. It was also enjoyable to learn that “stout,” “ale,” and “lager” are all types of beer — I hadn’t know this before. Extended knowledge of beer could include: “barrel-aged”, “table- served”, “dark Belgian-styled” and so on. This is what I had hoped to find as rich expressions. It means diversity, a new way to describe a familiarized object¹⁹. I also paid more attention to my pronunciations since the computer might be inclined to understand “ale” as “L”.

¹⁹ This always show up in a new adjective-noun combination, for example, “barrel-aged beer”.

Adding to this, I naturally made my own cheerful sound — “Bing ” — at the end of “Five things.” This is the reward for having successful interaction with the computer. “Bing ” means, “Hey bot, we made it together!” I did not see the process of testing as the process of solving obstacles created by the system, but instead thought of the process as us — the computer and I — collaborating.

In the third hour of interaction, I skipped all questions that asked for specific nouns such as “fruits/vegetables/cars.” I could now come up with newspapers such as “Atlantic news,” “New York Today,” and “Toronto Star,” for the question “Give me five newspapers.” In the past, I struggled to play word association games because I did not have the raw materials — a strong vocabulary. The helper function can list all the words related to publication so I would start to recognize a pattern in the names of various newspapers. I would assemble “journal,” “news,” “daily,” and a location to make up a newspaper, for example “Newfoundland Today.” If I can remember the pattern of the newspapers’ names, I can give more creative response to my partner in future improv games, for example “the journal of my broken nose.”

The same thing happens in giving answers to dogs, cat, and bread. It is hard to memorize all of the various dog names such as “Dalmatian” and “Papillon”. But it is easy to combine the words like “semi-long hair,” “smooth,” “twirl,” with locations like “British,” “Scottish,” “China,” and breeds like “terrier,” “spaniel,” and “hound” to make up a new type of dog—for example: “Asian semi-long hair terrier.” This discovery coincided with the instructions from the Improv, game coach. He encouraged us to think out of the box that instead of saying “Manx,” we could say “Tom’s cat,” “black and white spotted cat” or “ a drinking milk cat.”

Besides acquiring new vocabulary, being exposed to rich expressions, and recognizing new patterns, computer humour is another thrilling aspect of the program that I discovered in the testing phase. For instance, when the game asked for types of beer, the helper function suggested “out-of-category beer” as one answer. This made me laugh. If someone says the same thing, I will think of the person as witty or humorous. I will personally use this in a future conversation with another person. In this way, the computer taught me a “byte of humour.” As the developer, I knew what the mistake was. I could fix it by eliminating that phrase from the data, but I decided not to. All I thought was, ‘this is funny’ and I hope to encounter more “bytes of humour” during my interaction with the prototype.

In summary, the testing session unveils my English acquisition process in the “Give me Five Things” game based on the SLA theory. It was tedious to learn too many new nouns all at once. Adding to this, it was not efficient to memorize a noun if I did not have any direct experience related to it. But I never get bored with picking up a new adjective + noun combination or a new way to describe things. Knowing more adjective + noun combination phrases has allowed me to develop what I have called “rich expression.” In learning these combinations, I recalled memories of reading books by the writer Jan Morris. She always has brilliant descriptions for ordinary things. For example, she writes about Sydney: “The whole matter of Australia, history, character, reputation, attitude, finds its best epitome in this particular corner of the great land mass, where Sydney stands beside its fjord-like harbor....It is equipped with all the statutory metropolitan tokens – city marathon, revolving restaurant, supine veiled figure by Henry Moore, breakfast TV and Bahai temple. (Morris 4)” I understand the main

message in this short paragraph but also notice the new words “epitome,” “fjord,” and “supine” that she uses to describe the city. By checking the meaning of the three words, I get three new vocabulary “finds its best epitome,” “fjord-like harbor” and “supine veiled figure” for my describing of things. I am amazed by her magical short sentences and the way of using adjective creatively such as “this particular corner of the great land mass” and “equipped with all the statutory metropolitan tokens”. The “comprehensive inputs hypothesis” in SLA theory explains the reason of my preferring of picking new adjective + noun combination during the game session and from the reading.

The findings pointed me to a clearer direction for the next prototype. In prototype two, I wanted to explore a different corpus that is enhanced with “rich expressions.” Prototype two is akin to a more advanced “Five Things” game where the chatbot is smart enough to detect the intention in the texts, helping me to generate new textual content as a response. The three hours testing for prototype one also showed a distinct feature of this tool. It worked well as a surrogate conversation partner in my own process of learning English. I used it several times until I fully mastered the skill. Building on the first prototype, Prototype two eliminates the speech recognition, since this function does not work well in the continuous interaction.

In summary, prototype one shows three things that an intelligent machine can help me through three ways for my English acquiring:

- Unpack more adjective + noun phrases to help me get better at describing things. This allows me to see, feel and know the rich expressions.
- Show me multiple proper and related answers to a question.
- A reward system and the ability to deliver some machine humour make the interaction and language acquiring process more interesting.

Question (Give me five things)	Answer 1	Answer 2	Answer 3	Answer 4	Answer 5
Fruits	pummelo cranberry boysenberry green Kingfisher playing chakalaka Canada goose	grapefruit Meandering dragon fruit Castro American robin swallow	durian raspberry chili pepper scarlet tanager Kingfisher American flamingo	purple mangosteem grapefruit grapefruit home to grab all quill	gem ball Canary melon kiwi fruit N/a Taylor Swift darning
Art movements	Video game art Russian futurism Russian post futurism	Russian futurism Video game art minimalism	Performance art minimalism super flat	N/a Renaissance humanism	N/a N/a formalism
Newspapers	New Times York The News Journal Chicago Sun Times	alien newspaper Atlantic City Press star gazer	South florida sentinel newspaper Las Vegas Review Journal Northwest Indiana	The Sun The Washington Post USA Today	N/a The Blade The Detroit news
Beer categories	pumpkin beer table beer foggia style strong l American style as lager	table beer foggia style strong l	American in Paris Flooring styles American light lager	Chocolate flavoured beer Australian Style Pale Ale Amber beer	Scotch ale Lego N/a

Table 3.1: Sub-dataset of the results in "Five Things"

3.4.2 Prototype Two

The first prototype achieved two improv games by using the specific functions from the Rita library. It carries only two single functions and works well for a game which has only one simple rule. Prototype one uses the mini-corpus database called “Corpora²⁰” to generate the text content. This prototype utilizes a collection of static corpora (plural of “corpus”) that are potentially useful in the creation of weird internet stuff. However, it does not achieve the higher level of performance required for prototyping a dialogue system which is capable of generating a dialogue on a specific topic or even an open domain dialogue²¹ — for instance, small talk. In their blog introducing Smart Reply, Google researchers have pointed out that, “the content of language is deeply hierarchical, reflected in the structure of language itself, going from letters to words to phrases to sentences to paragraphs to sections to chapters to books to authors to libraries, etc.” (Strope and Kurzweil). In prototype two, I want to move from categorized words to phrases or even short sentences. This will require that I search for suitable conversational corpus or literature corpus fused with a sense of humour and narrative characterization.

At the same time, I realize that my research is not to make a “killer application”²². ImprovChat is neither an amusing chatbot nor a perfect messaging chatbot. At the sketching stage of the research, the concept of a dialogue assistant is often misunderstood as a chatbot. Peers always think I am making a chatbot that can chat with the user and test her/his English fluency—thus, the chatbot will learn more about the user and provide further English materials. Contrary to building a chatbot which interacts with the user spontaneously, the dialogue assistant will prompt dialogue suggestions only when the user asks for them. This means ImprovChat is not a full dialogue agent. The dialogue system will only give a single response to the user when the user clicks on the bot image (just like clicking the “Answer” button will fire the helper function).

I built the phrase generation system and the phrase management system with the Chatterbot python library and customized corpus. After the success of testing, I built the back-end with Django python framework, sqlite3 database, and integrated the front-end user interface with HTML, CSS and Javascript.

3.4.3 Corpus

ImprovChat was built with a conversation support system extracted from a built-in improv formula. It deals with natural dialogues which means that conversations are unconstrained and unscripted — for example, between interlocutors who are not instructed to carry out a particular task, to follow a series of instructions, or to act out a scripted dialogue (Serban et al. 9).

Corpus contributes meaningful value to two dimensions in ImprovChat. Firstly, it matters to the user because it decides the content of the comprehensive inputs for the language acquisition device. The potential of corpora for language pedagogy has been widely acknowledged since Tim Johns’ pi-

²⁰ Corpora <https://github.com/dariusk/corpora>

²¹ Open domain conversational dialogue systems should be able to respond to any user utterance on any topic” (Sugiyama et al. 335).

²² Most AI researchers have concluded that the generation of natural language narratives that are both domain independent and Turing test compliant is a “killer application”: it defines the outer limit of computational creativity (Lönneker et al.).

oneering suggestions for the use of corpora²³ in language learning, also called "Data-driven-learning"²⁴(Braun 307).

Corpus²⁵ also includes a dataset transformed by machine learning algorithms. The transformed data is used to generate the responses of the dialogue assistant. My exploration of the corpus covers works of literature, microblog dialogues, online Q&A forums, and movie scripts until it narrows to a corpus called Narrative Corpus hereafter NC²⁶ introduced by Christoph Rühlemann and Matthew Brook O'Donnell, alongside with the corpus tailored from a series of Twitter feeds and Reddit subsets.

3.4.4 Chatterbot Library

ChatterBot is a Python library designed to make it easy to create software that can engage in conversation. ChatterBot uses a selection of machine learning algorithms to produce different types of responses (see Figure 3.1). The machine-learning nature of ChatterBot allows an agent instance to improve its own knowledge of possible responses as it interacts with humans and other sources of informative data. An untrained iteration of ChatterBot starts off with no knowledge of how to communicate. Each time a user enters a statement, the library saves the text that they entered and the text that the statement was in response to. As ChatterBot receives more input, the number of responses that it can generate and the accuracy of each response in relation to the input statement increases ("About Chatterbot").

The program selects the closest matching response by searching for the closest matching known statement that matches the input, it then chooses a response from the selection of known responses to that statement. In my dialogue system, I use "best match logic adapter" for the phrases generation and set the response setting method to "random select phrases". In some circumstances, the system cannot find a best match phrase²⁷, the system will return a scripted response, such as "Sorry, I don't understand you."

23 In linguistics, a corpus (plural corpora) or text corpus is a large and structured set of texts. They are used to do statistical analysis and hypothesis testing, checking occurrences or validating linguistic rules within a specific language territory ("Text Corpus").

24 "Data-driven learning" (DDL) describes how language learners could become language detectives to explore language data themselves (Boulton 23).

25 In machine learning field, corpus is a collection of machine-readable texts that have been produced in a natural communicative setting (Pustjovsky and Stubbs).

26 The NC is a specialized corpus containing narratives extracted from the demographically-sampled sub-corpus of the British National Corpus (BNC) (Rühlemann and O'Donnell 315).

27 Calculated by the algorithm and return as the confidence. If the confidence is lower than the setting threshold, the system will turn to use the "low confidence logic adapter" to generate the response.

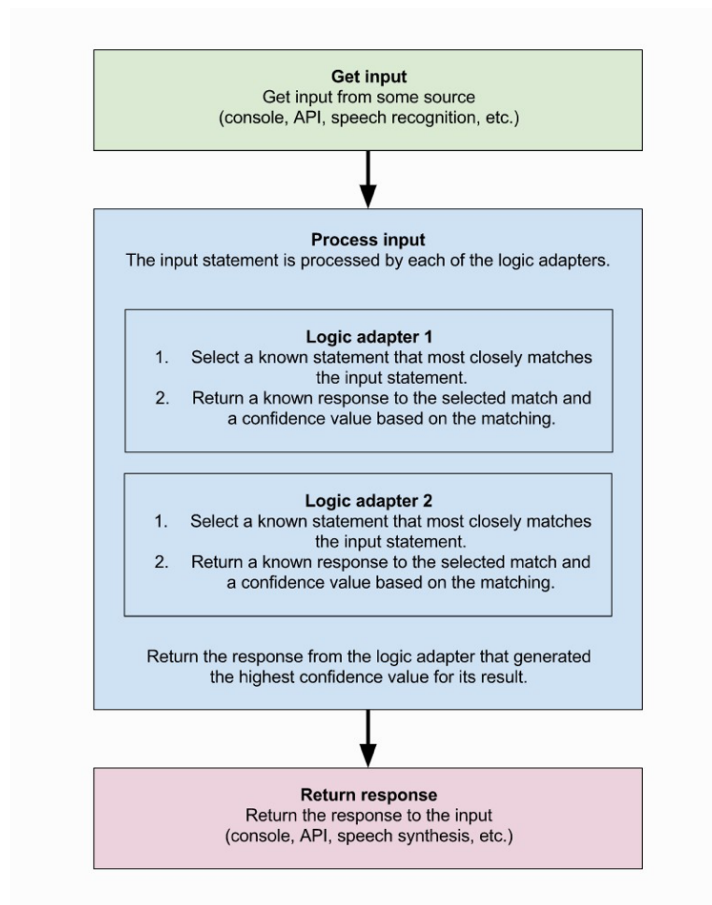


Figure 3.1: Chatterbot Process Flow Diagram ("About Chatterbot ")

3.4.5 Backend: Django and Twilio

The application was finally wrapped up with Django web framework²⁸. Figure 4.1 shows the architecture of the whole application.

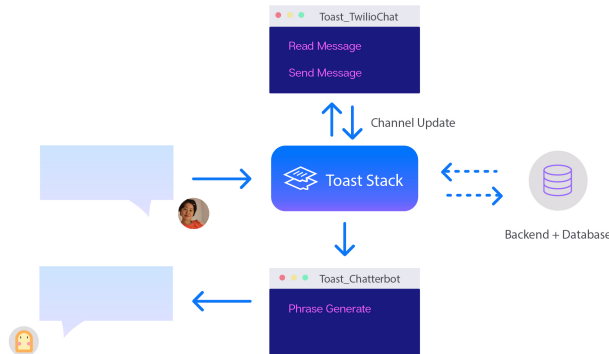


Figure 3.2: Architecture of the web application

3.4.6 User Interface

In his book *Play Matters*, Miguel Sicart writes that, “playful technologies are designed for appropriation, created to encourage playfulness” (31). This being the case, I wanted to sketch and think of a delightful design for my *ImprovChat* (see Figure 3.3). I tried to achieve this by making an adorable chatbot face and using a vibrant color. There is one shape in particular that kindles my mind. It looks like a piece of “Toast,” and reminds me of the image of toast jumping out of a toaster when it is having been warmed. This motion can be transformed to the animation to make the interaction more interesting. Toast is a common breakfast food in English speaking countries, but it is not popular in my home country at all. So, it stands for “a thing that is basic in English but not so familiar to me”. Toast also refers to an English alternate meaning. To toast also means “cheers” — it refers to the joyful emotion in the interaction.

²⁸ Django is a free and open-source web framework, written in Python, which follows the model-view-template (MVT) architectural pattern.

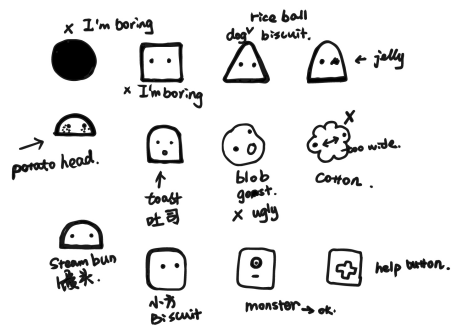


Figure 3.3: Sketches of the chatbot icons

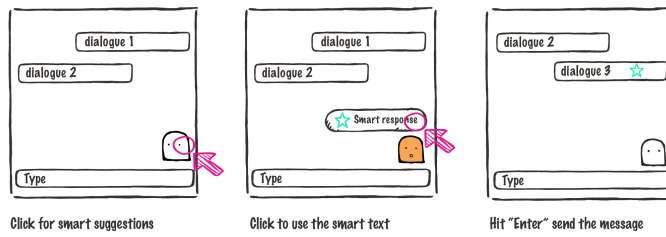


Figure 3.4: Sketches of the user interface and user interaction

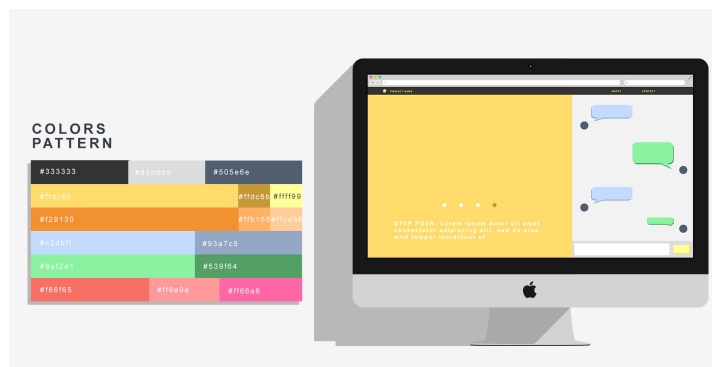


Figure 3.5: Showcase of the UI design

3.5 METHODS: USER TESTING

The research process includes user testing to gather documentation of interactions with the prototypes that will assist in display, archiving, and promotion when developing the two iterations of the prototype. The goal is to observe reactions, listen to critical feedback about the current prototype, gain insights into what could be changed or improved, and to include that knowledge in future iterations of the prototypes. There has been one round of testing in total. Three teams of two participants each were tested. Each test session ran 30 minutes and was conducted between March 15–16, 2018. Each 30-minute session included a text chat session (20 minutes) and a questionnaire (10 minutes) for a total of 30 minutes in duration.

The participants included three international undergraduate students from OCAD University who have English as a second language (English Language Learners, or ELL) and three undergraduate or graduate students who are Native English-language Speaker (NES). Each ELL was paired with an NES in an online text chat interface. ELL participants used the application (ImprovChat) that I am developing as an aid during the chat. The ImprovChat application could display one to three phrases that might be used as responses. The user could choose to select one of the phrases provided by the ImprovChat, or not. What was being tested is whether the application provides useful and pleasurable alternative responses from the viewpoint of the English Language Learner. Pairs of participants used laptops for the 20-minute text chat sessions. The Native English-language Speaker (NES) started the chat, opening with a question such as, "how is your day going?" After the 20-minute chat interaction, each participant filled out a questionnaire requiring approximately 10-minutes. The NES was asked questions that evaluated whether the chat achieved meaningful English usage. The ELL evaluated the user experience of the ImprovChat application. Each participant was provided with a box in which to add personal comments. The complete questionnaires have been attached to this document as Appendix D and E. The research output includes documentation of the chat in text and the tabulated results of the questionnaire.

3.5.1 User Interaction Feedback

All the information gathered during the testing phase was collected from the questionnaires. Some of the ELL users were not confident about writing about their user experience, so they used their first language instead, and the researcher carefully translated their feedback to English afterward.

Native English speakers did not use the chatbot functions at all; they were only in charge of chatting and providing feedback regarding the chatting experience with the ELL users. After the interaction, they got to look at the user interface from the ELL users' side to know how the chatbot was working before writing the questionnaires.

Native English Speakers

NES User 1:

User 1 remarked, "The chat seems humorous sometimes. I remembered I once received an emoji from the chatting partner who is so cute (this emoji comes from the ELL user, not from the chatbot). It is hard for me to tell if the chat has a particular tone. I wished I could get some notifications/hints

when the other side was typing. 2. It is user-friendly, if it can be used on some portable devices (phones, tablets, et.) 3. Thinking about the future application, it could be used as an educational tool. You could design this app for some audiences with special topic/subject/context to improve this language learning experience.”

NES User 2:

User two remarked, “I think the chat was quite playful. But I was not able to tell which responses came from the chatting partner and which ones came from the chatbot. The pot question was very funny (this comes from the chatbot suggestion). I think this is a chat bot with a lot of potential for users such as international students, or new English speakers. What I hope will be added is a now “typing” function which would allow me to know when to expect an answer. Again, from my point of view, I was not able to tell which ones were answered by the users and which ones were answered by the bot, so it was hard to me to understand which was the friend’s personality. I questioned myself, wondering was that him or it?? It also seemed a bit awkward between my friend and me because we already know each other too well, and it might be hard for him to answer any serious questions. Thanks for working so hard on this chatbot! I believe after some design interaction it will be much more useful for the world.”

NES User 3:

User three remarked, “There is one point that my chatting partner said: “I have been known to take shortcuts now and then” (this phrase comes from the chatbot suggestion). I feel this is so interesting because the tone of his response is quite different from what he would normally say. I would like to know my friends’ opinions of using it, and I hope that I can try it too.”

English Language Learners

Three ELL users ranked the quality of the phrases provided by the chatbot as poor, acceptable and good. They did not use the suggested phrases very often during the chat exchange. On the whole, this indicates that the performance of the chatbot is not quite stable. But at the same time, their feedback about the specific user experience was inspiring, and all of them have given me useful suggestions for future development.

ELL User 1: User one responded “I found that the chatbot only had several fixed suggestions for me. I was hoping to see more different things. Thus I feel it is not super smart. I hoped that it could have a real human mind. For example, that it could understand which topic I am talking about, and give me highly related suggestions. The overall experience is fun except that sometimes I don’t understand the prompts at all, and sometimes I feel it is weird. Well, some vocabulary is above my English level, so I hoped this application could give me transcriptions of some new words.

There is one turn that caught my difficulties in understanding. I asked my chatting partner: “Which country music do you like?” She answered: “I forget the name, sorry for the bad memory.” And I clicked the chatbot after her reply. The chatbot suggested, “if we were together we would be lovely.” Obviously, the bot doesn’t understand that what we were talking about was all about music- it jumped to a new topic about showing affection to another person. I found it weird, but it also lightened my mind. Probably because I’ve never responded to anyone before like this. But on the other hand, if I used its unrelated suggestions as my responses, I worry that my chatting partner will see me as impolite or abrupt. Most of the phrases

look fresh to me since it was my first time to meet these expressions, and this triggers my interests. I felt the bot can do a good job in broadening and discovering conversation topics. So, it can turn the awkwardness to less awkwardness. Another bias I found immediately is the mistranslation of polysemy in the speech. The chatbot does not seem to fully understand the context of the statements. I would like to use it in a situation that I have to communicate with people using different languages and when the chatbot is good enough to give me 2 – 3 good responses. Last, but not the least, some online customer service provider who will chat with you to know your customer experience or inquiries, they will sometimes just tell you "they don't know" or are "unable to help," but this chatbot could keep searching for me. I appreciate its effort. For the future iterations, I hope the artificial intelligence could enter a new advanced level. It should be able to tell the ongoing topic in the conversation and at the same time broaden the topics for my conversation."

ELL User 2:

There is one turn in our conversation that seems funny to me because the chatbot tried to suggest something about a person, and it seemed to me that the chatbot knows whom we are talking about and his suggestion coincidentally described that person. The chatbot seems like it knows part of my life. The chatbot seems to have a personality, and I will say that half of the time, it is light-minded and the other half of the time, it is serious. The chatbot mainly can help me break an awkwardness in the conversation with my chatting partner. It gives me some funny ideas to continue the conversation, with a better response. Though I might not use the suggested phrased directly, its prompts me something, so I can think about it. I can also catch some new words and new phrases from its suggestions, especially some phrase containing specific interpretation for different situations. For the future, I might use this application if it is popular among my friends. I did not get the point of the chatbot at the first 10 minutes so I hope the designer can make an introduction page to show me how to use the application. The chatbot reaction is a little bit slow; this can be improved for sure.

ELL User 3:

It is hard for me to tell the chat session is humorous because I am not much of a chatting person. I found the chatbot does not quite answer the question well. It seems adorable but also a little bit inflexible. It fits more for the young people's way of chatting, or chatting when it is not about serious things. For the future iteration, it can be developed as an add-on application embedded in other chatting applications. The chatbot could have more animated actions to indicate to the user whether it is processing or if there is an error. Maybe it could also suggest emojis, gifs, and pictures to me. And since I am not into chatting, I hope the chatbot can suggest phrases to end the topic or the whole conversation. It could be fun to suggest some black humour thing to me to make the partner not want to chat with me anymore.

4 | FINDINGS

In this chapter, I will recall the difficulties I had with conversations in English and insist on seeking effective solutions to acquire²⁹ English knowledge for chatting. As I tried to study myself, this “myself” became a third person. So, I feel that it is easier to explain this process by using a persona. Her name is Rose. Rose represents me but could also refer to other English learners who might have similar issues as me and want to try an innovative way to build conversation competence in a second language.

Rose is an adult; her first language is Mandarin and she received English training in China for over five years. She has primary English vocabularies, and she can conduct self-learning by reading books, watching movies, and checking the dictionary. She did not have abundant experience talking with native English speakers. She found it is difficult to express freely and creatively in social situations. This issue followed her when she moved to Canada where English is required for communication every day.

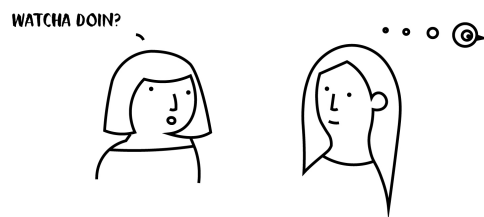


Figure 4.1: What should I say now?

During the first year, Rose always found herself stuck in the chatting, because she doesn't understand what the other people mean. For example (see Figure 4.1), her classmate asks: “Watcha doing?” or “What’s up?” She immediately gets confused because she doesn't know what that means. The course book only showed her “How are you?” as a greeting. She remembers to say “I’m fine” well but often forgets to ask back as a form of respect. Thus, the conversation makes both her classmate and herself feel embarrassed. In short, the stuck moment happens because Rose has less exposure to general conversational settings.

As she starts to acquire more of the language, Rose improves in the second year to the extent that she can initiate a chat, though she still struggles to understand the cultural context. Rose understands what the other people

²⁹ I say “acquire” is because I did not “learn” under a structured course outline, whether it is participating in a formal English class at school or following chapters in a course book. I mainly get new English knowledge by engaging myself in the conversations, and simply by listening to other people talking in life or on TV.

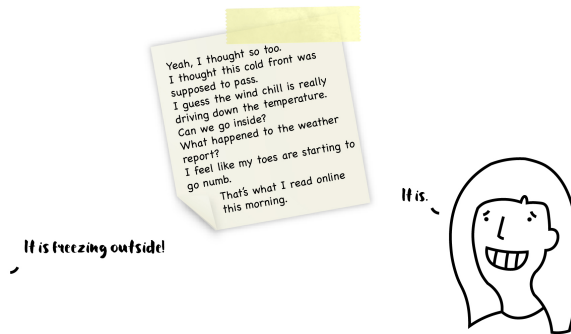


Figure 4.2: I don't realize that I can say that

are saying, and she knows how to mono-respond. For example (see Figure 4.2), a friend says: "It is freezing outside!" Rose naturally responds: "It is." But she feels stuck again because this is the only response she knows, compared to a fluent English speaker who might have many different responses, such as "I guess the wind chill is really driving down the temperature," "Could we go inside?" or "I feel like my toes are starting to go numb." The conversation ends quickly because she doesn't know how to extend this "talking about weather" context. She starts to pay attention to a new ability to generate responses containing richer content and different expressions as an opposed to mono-response.

If Rose gets stuck in a conversation, she would google the meaning, refer to a dictionary, check phrases books, ask her smart friend Allen after the interaction or stay unknowledgeable (bravo!) until she gets stuck again in another conversation (see Figure 4.3). At this stage, understanding the message is no longer the hardest part for her. Instead, remembering the phrases and using them properly become the core issue. *Case1:* If Rose



Figure 4.3: A traditional way of learning new phrases

remembers to use phrase A, then she has successfully acquired this phrase A by using it in her conversation. But, one risk is that she might stick to this successful phrase as in the "I'm fine" example.

Case2: In another situation, she forgets phrase B that she just learned from the book, so this phrase B might be forgotten forever or it might have sunk into her subconscious, waiting to be used at another time.

Case3: Many times, Rose can't remember the phrase C that she just learned from the book, but at the same time, she is aware that she did learn a useful phrase somewhere before. Either by reviewing her phrase book or by googling, she can refresh her memory of this phrase C.

All three cases reveal one problem in the traditional way of learning new phrases and using them in conversation: the learning process and the practice process (as in, applying the knowledge in conversations) do not always happen at the same time. Or one might say, the input stage and the output stage are separated. This could cause inefficiency in gaining conversational competence in a second language.

4.1 ARTIFICIAL INTELLIGENCE

This section will demonstrate to the reader the ways in which the AI dialogue assistant will join the conversation and provide support and why this is an innovative solution (see Figure 4.4). Speaking is not the cause of better acquisition, but a result of it (Krashen, 'Din' 43). Without enough comprehensible inputs being poured into the acquirer's brain, she/he will encounter difficulty initiating conversation, maintaining the conversation and ending the conversation. Based on the overview of the SLA theory in the Chapter 2, we know how language acquisition devices (LADs) works. Adding to this, through the two rounds of user testing for the prototype one and prototype two, it is clear that ImprovChat can provide a number of comprehensive inputs across broader conversational topics in the form of authentic phrases, short sentences and funny messages when the user needs its help. Since the AI dialogue assistant can contribute to a conversation as a support for human intelligence, it can decrease a users' stress in awkward conversations. This lessening of the user's stress will also lead to users' better apprehension of English. There are many occasions in which the user may experience the benefits of collaborating with the chatbot's artificial intelligence. The dialogue system was made through a data-driven approach and, given that machines are good at storing and sorting large amounts of data, humans can be released from duplicating this data-driven study. Therefore, the AI dialogue assistant will challenge Kreshan's statement that "real language production happens only after the acquirer has built up competence via input" (43) in that this approach allows the user

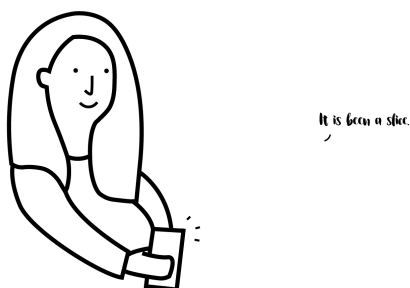


Figure 4.4: My AI dialogue assistant makes chatting much more enjoyable

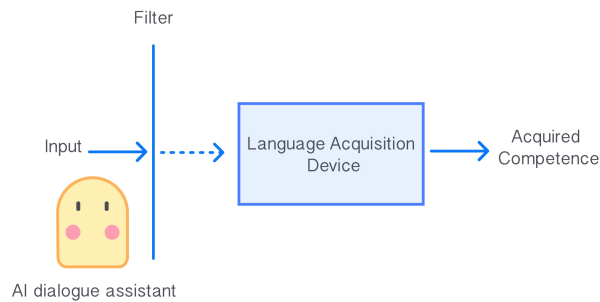


Figure 4.5: AI dialogue assistant join in the SLA

to experience a different way to produce dialogue before developing the expression skills of higher-level creative expressions or advanced English knowledge (see Figure 4.5).

The time delay in getting a response on the chatbot is very important for carrying on the conversation. Those who have not mastered a language may experience the need for more time to translate or interpret before finally coming up with a proper response. Thus, they might miss the right moment to respond to other people in a conversation. ImprovChat shortens the time required for the user to come up with the right words and reduces the self-doubt they may experience with their choice. ImprovChat also releases the user from the pressure of, “I don’t know what to say” (also seen as a block in the speaker’s head) and encourages the user to think, “what I can say” with help from the AI dialogue assistant.

Another advantage is that ImprovChat can help the user to solve a problem right at the moment of the human-human communication. A very interesting hypothesis is that we acquire language best when the pressure is completely off, when anxiety is at zero, and when the acquirer’s focus is entirely on communication (Krashen, “Din” 43). As Hodgson and Ernest write, “[w]hen we grasp of a situation with our whole being, it means that conscious effort at memory work is unnecessary. What is learned is placed upon the deepest knowing and away from mere repetition for the sake of rote learning” (24). ImprovChat does not force the user to accept the knowledge compulsively but through the user actively seeking new knowledge. This can be used to explain the phenomenon I experienced before (mentioned in Chapter 1) where I could remember the phrases or the short sentences suggested by a chatbot right after I used them in my online chatting.

All in all, ImprovChat will not replace human intelligence, but will inspire human intelligence.

4.2 ARTIFICIAL STUPIDITY

Brain Eno has talked about his interest in using AI to compose a new piece of music and gave a very brilliant speech about how artists interpret technology in a unique way: “Computers make some very weird mistakes, and a lot of those mistakes are very interesting. One of the things artists are interested in for technology is the things that they do that they’re not supposed to do. The dominant texture of any era is really captured in the shortcomings of

those technologies (qtd. Beer).” In short, he calls this “artificial stupidity” and that is where the section title comes from. This section will talk about a new scope of my design research which “challenges current perceptions on the role and form of technology” (Zimmerman et al. 310). My new perception comes from looking at the portion of AI technology which I call it chatbot humour. I have talked about the ‘byte of humour’ in the section entitled “Prototype One,” and confirmed that this humour is due to a mistake in the dataset. In the section entitled “Prototype Two,” I deliberately selected the corpus with humorous characterization, such as the data retrieved from “Kids Write Jokes” Tweeter feeds. With the chatbot’s weak algorithm, there are a lot of mismatches that happen during the phrase generation process. This being the case, I will discuss why I view machine weirdness as part of the chatbot’s own personality rather than fixing these so-called mistakes.

The computer manipulates syntax to represent the language, which is also called “thought vector.” But it does not track the semantic meaning behind the thought vectors. That’s why human testers in the user testing can tell the chatbot does not have real consciousness of the meaning of the chatting. ImprovChat hints at new options for the user—including, sometimes witty ways to jump across certain obstacles in the conversation. This facet of ImprovChat matches what Colin and Baedeker suggested when they said to ‘leapfrog over the expected response.’³⁰ This is why the chatbot outputs look meaningful to us in some occasions regardless of the weirdness. My research echoes the views of Paul Prudence. I believe that the machine fallibility will lead creative AI to the kind of emergent weirdness essential for improvisation³¹.

When a kid says “when I was young,” you laugh because you think that a kid doesn’t have a long enough living history to announce “when she/he was young.” In the same way, a programmed chatbot can produce the kind of interesting thoughts that would come from an adult’s experience, and the contrast itself already make things funny. There might be a way to analyze the grammar of the humorous texts. But we also need to consider other important ingredients in a conversation such as the speaker’s state of mind and whether she/he has a playful attitude. Humour is a type of improvisation which happens at the moment, at a location, and between people. In improv, humour does not mean making jokes. The funny parts of improv always come from the speaker’s spontaneous responses, but not from trying too hard to amuse people. To make amusing the audiences a goal in improv would put so much pressure on the performers that it could freeze the players’ creativity. Humour is more likely to happen when people feel relaxed. Sometimes, simple-mindedness and stubbornness can make for humour too. I once played a “Jump in the Circle and Sing” game at improv, where the rules are that you jump into a circle and sing a song that you associate with the last song sung by the last player in the circle. During this game, the pace of singing went so fast at first that it felt thrilling to me. Gradually however, the atmosphere became embarrassing because the players almost ran out of songs to sing. Then, suddenly, someone jumped in and started to sing “Happy birthday to you.” Everyone laughed! I wish I was the one who jumped in but I did not! I realized that I wanted to sing songs that made me look smart, but this kind of thinking does not fit well with the spirit of improv games. Whenever you don’t know what to sing,

³⁰ Leapfrog over the expected response is recommended by speech experts as a way to transform small talks into smart conversations (Colin and Baedeker).

³¹ Prudence said: “... machine fallibility will lead creative AI to the kind of emergent weirdness essential for composition” (Prudence 23).



Figure 4.6: Different-able AI

it is enough to just sing "HAPPY BIRTHDAY." I would argue that this is the same principle as when the chatbot says, "I still don't have any ideas," when it only has one answer. From this perspective, I am confident to say that the chatbot is good at improvisation because it does not "think hard" to generate funny things — rather, it just says whatever it knows. Even still, this kind outspokenness can still not be called spontaneity because a chatbot cannot adapt to a conversational topic organically at the current stage that it is in. This is why human intelligence is important in this machine – human intelligence collaboration and plays the role of monitoring the artificial intelligence outputs.

A Tweet about AI caught my eyes recently (see Figure 4.6) I could feel the pulse of crying when I saw it. My response had nothing to do with the creator of this Tweet, but only with the empathy that I have developed for Artificial Stupidity during my research. Why should an AI be teased for not running elegantly? Why couldn't AI have its own running postures? Some people might run like this for their whole lives and be viewed as a weirdos. Some people create humour through self-mockery. This can't happen if people do not fully accept themselves, especially their awkwardness, or human fallibility. In one episode of Friends, Phoebe runs awkwardly — like the AI in Figure 4.6 — which causes her friend Rachel to be embarrassed. Interestingly however, at the end of the episode, Rachel has changed her mind and wants to mimic Phoebe. She carefully looks around to make sure no one is watching her and then suddenly releases her silliness to run just like her friend. At first, the viewer might think Phoebe is weird, but at the end, we are led to think that Rachel is funny. She is funny for her childish behavior, her innocent expression, and most of all for her acceptance of her friend's own unique weirdness.

Alan Turing wrote that “If a machine is expected to be infallible, it cannot also be intelligent.” Prudence reworded Alan Turing’s maxim to “if a machine is expected to be infallible, it cannot also be creative” (Prudence 23). If improv does not make efforts to cultivate a non-judgmental culture and encourage players to make mistakes, there will not be a flow of the improvisational creativity. During the research, I always laugh at the machine weirdness. When I am laughing, I am not consciously judging it as a mistake. My drawing class teacher always said that we need to make over 3000 mistakes before internalizing the knowledge. Fostering creativity takes the same path. Creativity needs to be written in the human mind’s subconscious first by repetition. Only then can creativity be released from a subconscious state. Through state-of-the-art machine learning technology, machines have started the process of building “sub-conscious” layers through repeatedly reading data. Lack of real intentionality is one of the criticisms leveled against artificial intelligence and machine creativity in general (Prudence 23). The machine will only know how to express creativity on its own when we humans know how to build a “conscious” layer (i.e. the ability to transform and mediate) for machines. When the chatbot knows how to control its outputs with its conscious layer, it could behave like it has real intelligence. But before that, I won’t consider its output as stupidity, but rather think of the outputs as coming from the schmozzle of the machines “subconscious” data. At the current stage, ImprovChat is unorganized and naïve, but demonstrates originality and creativity as well.

4.3 RESEARCH

Research through design is still in its emerging stages. Researchers argue that there is a need for serious development of RtD into a proper research methodology that can produce relevant and rigorous theory (Zimmerman et al. 316). I experienced what participants in the surveys of RtD mentioned in that, “RtD and designers conducting research in general suffer from a romantic view of design” and “there is a lack of examples and critical analysis of these examples for this kind of research” (316). The romantic view of design in ImprovChat shows that as I spent a lot of time looking for the latest AI technology, it turns out that I am not able to use it because of the invisible depth of the knowledge. However, I still have benefited a great deal from applying RtD in my research, because “RtD is the closest to the actual design practice, recasting the design aspect of creation as research” (Findeli) and “allows design researchers to ignore commercial concerns in order to focus on new understanding of technology” (Zimmerman et al. 310). I’ve been involved in the ImprovChat project for the last 8 months, and most of the findings come from the prototyping directly. Being a programmer and a designer, I am able to research the technology by shaping it for design. I can figure out the reasons for the machine fallibility right away, thus deepening my understanding of the pros and cons of the technology.

5 | CONCLUSION

The scholar Sherry Turkle foresees a significant change in modern conversations when “we are tempted to talk not only through machines but to them, with them.” (16-17). The ImprovChat AI dialog assistance is my response to social communication issues inspired by my lack of English language fluency, alongside my experience with improvisation games. The iterative design process helped me to discover the underlying rules of improv that can be useful in acquiring individual expression skills for daily conversation. The process experimented with building a dialog system that uses AI technology for the individual acquiring language through communication and for communication. ImprovChat provides an opportunity for second language learners to start a conversation using her/his digital assistant. ImprovChat is a glance toward the future, where people will have a higher acceptance of AI; it tests the form in which AI will exist in our daily lives. A great deal of communication in the future will include AI tools adapted to individual needs, assisting individuals in seamlessly chatting in a second language, and ImprovChat is proud to be at the cutting edge of this progression. As a researcher, I see the dangerous point where the machines might overstep the voice and the character of the speaker since machine intelligence does not necessarily connect to the user’s intelligence. Turkle has warned about the “the flight from conversation”³², and the deficiency of machines that are unable to take care of humans during the human-machine interaction era. I assert that the AI dialog assistant as a crowd-sourced³³ computational agency, will unpack a multitude of expressions for second language learners but will not control the communication directly. This device is more of augmentation, inspiration, and a call and response (Sloan).

5.1 FUTURE DIRECTIONS

External knowledge sources such as online encyclopedias and English to the user’s first language translators can be introduced as extensive features for future iteration. As the only user of ImprovChat at this stage, I realize that the interpretation of the AI phrases and concordance require a level of analytical skill. This factor might become one obstacle for broader users. As Krashen notes, “knowledge of components of conversational competence is one thing. Developing conversational competence in students is another (Krashen, “Principles” 70).” ImprovChat is only at its beginning stage and there is more to do to truly encourage the user to participate in the conversations continually. On the other hand, the nature of language acquisition also indicates that ImprovChat will be a surrogate tool for the user. It will only serve as a bridge, a transportation for the user to reach a more confident

³² And we use conversation with others to learn to converse with ourselves. So our flight from conversation can mean diminished chances to learn skills of self-reflection (Turkle).

³³ Sherry Turkle used “crowd-sourced therapist” to describe the chatting robot in her book “The Second Self”.

level of second language usage for their communication. If we look at the process where a person gains her/his first language, we will find that there is usually a teaching figure involved who serves as dialogue assistant. This figure could be the mom, dad and whoever played the role of delivering the fundamental knowledge of that language. At the initial stage, when the child cannot use the language for their conversations independently, this figure will intervene in the conversational settings with social tolerance and help the child to analyze, phrase and rephrase. But still, the child might not fully understand or remember the phrase. The teaching figures will repeat this process with patience until the novel user is fully capable of speaking that language correctly by herself or himself. If the AI technology is powerful enough, such a figure — a robust dialogue assistant — would be developed for the second language learners. The comparison between first language acquisition and second language acquisition will provide more insights on this topic in the future.

Advanced machine learning algorithms, or more complicated models, will be much more effective than the ImprovChat prototype. An improved ImprovChat would require a large training dataset and highly efficient computer to carry out the calculations needed for success. A team of computer science PhDs might be able to push the design to a champion version. However, the technical insufficiencies do not devalue the prototype. As Winograd has stated, “rather than basing AI competence on a logical representation of the setting and the agent’s knowledge, there is an interplay between general adaptive mechanisms and world experience, which leads over time to intelligent behaviors” (Winograd 1258). Prudence even challenges the current research environment in AI: “computer scientists working on [AI generated music] right now, don’t seem to have the aesthetic sensibilities to bend their algorithms far enough into alien dimensions” (Prudence 22-25). This statement indicates the necessity of interdisciplinary cooperation between computer scientists and artists, and calls for the participation of people with art, music and theatre backgrounds to contribute to artificial intelligence.

6 | WORKS CITED

- "5 Things – Canadian Improv Games" **Improv.Ca**, <http://improv.ca/5-things/>.
- "The Design Process: What Is the Double Diamond?" **Design Council**, 2005, <https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond>. Accessed 23 Jan 2018.
- "Writing with The Machine." **Robin Sloan**, 2017, <https://www.robinsloan.com/notes/writing-with-the-machine/>.
- "About Chatterbot — Chatterbot 0.8.4 Documentation". **Chatterbot.Readthedocs**. Io, 2018, <https://chatterbot.readthedocs.io/en/stable/index.html>.
- "Software Agent". **En.Wikipedia.Org**, 2018, https://en.wikipedia.org/wiki/Software_agent.
- "Text Corpus". **En.Wikipedia.Org**, 2018, https://en.wikipedia.org/wiki/Text_corpus.
- Abu Shawar, Bayan, and Eric Atwell. "Chatbots: Are They Really Useful?" *The Journal for Language Technology And Computational Linguistics*, **LDV-Forum 2007**, 2007, pp. 29-49, <http://www.gldv.org>.
- Antoniadis, George et al. "Integrated Digital Language Learning." 2009, p. 89.
- "Applied Improvisation." **En.Wikipedia.Org**, https://en.wikipedia.org/wiki/Applied_improvisation.
- "Applied Improvisation - Anima Learning." **Anima Learning**, <https://animalearning.com/applied-improvisation/>.
- Barr, Avron, and Edward A. Feigenbaum. **The Handbook of Artificial Intelligence**. Heiris-tech Press, 1981, pp. 8-11.
- Bateson, Patrick, and Paul Martin. **Play, Playfulness, Creativity and Innovation**. Cambridge University Press, 2015, pp. 103-109.
- Beer, Jeff. "Brian Eno Talks About Using Artificial Intelligence To Create Music And Art". **Fast Company**, 2016, <https://www.fastcompany.com/3061088/brian-eno-talks-about-using-artificial-intelligence-to-create-music-and-art>.
- Behar, Ruth. "The Vulnerable Observer: Anthropology That Breaks Your Heart." 1997.
- Boulton, Alex. "What Data For Data-Driven Learning?". 2011, p. 23.
- Bowden, Kevin K. et al. "M2D: Monolog to Dialog Generation For Conversational Story Telling". **9Th International Conference on Interactive Digital Storytelling**, 2016, p. 14, <http://www.springer.com/us/book/9783319482781>. Accessed 9 Feb 2018.
- Candy, Linda. "Practice Based Research: A Guide". 2006, p. 3.
- Carrane, Jimmy. **The Inner Game Of Improv: 5 Steps To Getting Bigger In Your Improv Career**. 1st ed., 2015, p. 27.
- Catherine, Santino. "I Used Gmail's Automated Responses For 24 Hours Straight And It Got Weird". **Levo League**, 2017, <https://www.levo.com/posts/i-used-gmail-s-automated-responses-for-24-hours-straight-and-it-got-weird>.
- Cho, Kyunghyun et al. "Learning Phrases Representation Using RNN Encoder-Decoder For Statistical Machine Translation". 2013.
- Colin, Chris, and Rob Baedeker. "How To Turn Small Talk Into Smart Conversation". **Ideas.Ted.Com**, 2018, <https://ideas.ted.com/how-to-turn-small-talk-into-smart-conversation/>.
- Corrado, Greg. "Computer, Respond To This Email.". **Google Research Blog**, 2015, <https://research.googleblog.com/2015/11/computer-respond-to-this-email.html?m=1>. Accessed 10 Feb 2018.
- Dalinina, Ruslana. "Word Embeddings: An NLP Crash Course". **Datascience.Com**, 2017, <https://www.datascience.com/resources/notebooks/word-embeddings-in-python>.
- Danescu-Niculescu-Mizil, Cristian, and Lillian Lee. "Chameleons In Imagined Conversations: A New Approach To Understanding Coordination Of Linguistic Style In Dialogs". **ACL Workshop On Cognitive Modeling In Computational Linguistics**, 2011, pp. 76-87.

- Findeli, Alain. "La Recherche-Projet: Une Méthode Pour La Recherche En Design". 2004.
- Forlizzi, Jodi et al. "From Design Research To Theory: Evidence Of A Maturing Field". 2009, p. 2894.
- Frayling, Christopher. "Research In Art And Design. Royal College Of Art Research Papers Se-ries". Vol 1, no. 1, 1993.
- Gannon, Susanne. "Autoethnography". textbfOxford Research Encyclopedia of Education, 2017. Oxford University Press, doi:10.1093/acrefore/9780190264093.013.71.
- Godin, Danny, and Mithra Zahedi. "Aspects Of Research Through Design: A Literature Review."
- Graff, Gerald et al. "They Say / I Say". 3rd ed., Your Coach In A Box, 2015, p. 10.
- Hodgson, John, and Ernest Richards. **Improvisation**. Methuen Drama, 1991, p. 26.
- Hoelson, Christopher N., and Rod Burton. "Conversing Life: An Autoethnographic Construction". textbfThe Qualitative Report, 2012.
- Howe, Daniel C. "Rita: Creativity Support For Computational Literature". p. 206. "Improv Encyclopedia - Yes And". **Improvencyclopedia.Org**, http://improvencyclopedia.org/glossary//Yes_And.html.
- "Iwordball – Canadian Improv Games". **Improv.Ca**, <http://improv.ca/wordball/>.
- Johnson-Laird, P. N. "How Jazz Musicians Improvise". **Music Perception**, vol 19, no. 3, 2002, pp. 415-442. **University Of California Press**, doi:10.1525/mp.2002.19.3.415.
- Johnstone, Keith.**Improvisation And The Theatre**. 1st ed., Richard Clay (The Chaucer Press), Ltd., 1981, p. 111.
- Johnstone, Keith. **Impro For Storytellers**. Faber, 1999, p. Introduction.
- Jordanous, Anna. "Defining Creativity: Finding Keywords For Creativity Using Corpus Linguistics Techniques". Accessed 9 Feb 2018.
- Krashen, Stephen. "The Pleasure Hypothesis". **Georgetown University Press**, **Round Table On Languages And Linguistics**, 1994, pp. 299-322.
- Krashen, Stephen. **Principles And Practice In Second Language Acquisition**. Prentice Hall International, 1995.
- Krashen, Stephen D. "The Din In The Head, Input, And The Language Acquisition Device". **Foreign Language Annals**, vol 16, no. 1, 1983, pp. 41-44.
- Labov, William, and Joshua Waletzky. "Narrative Analysis: Oral Versions Of Personal Experi-ence". Seattle: University Of Washington Press, **Proceedings Of The 1966 Annual Spring Meeting Of The American Ethnological Society**, 1966, pp. 12-44.
- Lê, Phong. "Learning Vector Representations For Sentences: The Recursive Deep Learning Approach". 2016, <https://dare.uva.nl/search?identifier=8d8a427c-141b-481b-b093-3167dfbe13b7>.
- Lenci, Alessandro. "Distributional Models Of Word Meaning". **Annual Review Of Linguistics**, vol 4, no. 1, 2018, pp. 151-171. **Annual Reviews**, doi:10.1146/annurev-linguistics-030514-125254.
- "Leonard Newmark - Teflpedia". **Teflpedia.com**, http://teflpedia.com/Leonard_Newmark.
- Li, Jiwei et al. "A Diversity-Promoting Objective Function For Neural Conversation Models". **Eprint Arxiv:1510.03055**, 2015. **ARXIV**, doi:arXiv:1510.03055.
- Li, Jiwei et al. "A Persona-Based Neural Conversation Model". **Corr**, abs/1603.06155, 2016. **Arxiv**, doi:arXiv:1603.06155.
- Lightbown, Patsy M, and Nina Spada. **How Languages Are Learned**. Oxford Univ. Press, 2013.
- Lonneker, Birte et al. "Story Generators: Models And Approaches For The Generation Of Literary Artefacts". 2005.

- Magerko, Brian et al. "An Empirical Study Of Cognition And Theatrical Improvisation". **2009 ACM SIGCHI Conference On Creativity And Cognition**, 2009.
- Mathewson, Kory W., and Piotr Mirowski. "Improvised Theatre Alongside Artificial Intelligences". Association For The Advancement Of Artificial Intelligence, **The Thirteenth AAAI Conference On Artificial Intelligence And Interactive Digital Entertainment**, 2017, pp. 66-72.
- Merlin, Shana. **Improv Warm-Ups : Improv Warm-Ups: Word Association In Pairs**. 2008, <https://www.youtube.com/watch?v=rnwEd17kedQ>.
- Merlin, Shana. **Improv Warm-Ups : Improv Warm-Ups: 5 Things**. 2008, <https://www.youtube.com/watch?v=gvoVm5Fq2oI>.
- Mini-Wheats Canada. **Kids Secretly Control A Speed Dating Session**. 2017, <https://www.youtube.com/watch?v=EugvWqdhue8>.
- Moreno, Roxana et al. "Life-Like Pedagogical Agents In Constructivist Multimedia Environments: Cognitive Consequences Of Their Interaction". Association For The Advancement Of Computing In Education (AACE), **ED-MEDIA 2000—World Conference On Educational Multimedia, Hypermedia & Telecommunications**, 2000, pp. 776-781, <https://www.learntechlib.org/p/16160/>. Accessed 9 Feb 2018.
- Morris, Jan. *Journeys*. 1st ed., Oxford University Press, 1984, p. 4.
- Moyer, Alene. **Age, Accent And Experience In Second Language Aquisition**. Multilingual Matters, 2004.
- Murray, Garold et al. **Identity, Motivation And Autonomy In Language Learning**. Multilingual Matters, 2011, pp. 57-71.
- Nerbonne, John. "Natural Language Processing in Computer-assisted Language Learning ". 2003, p. 678.
- Newmark, Leonard. "A Minimal Language Teaching Program". **The Psychology Of Second Language Learning**, vol 32, 1971, pp. 11-18.
- Paakki, Henna. "Difficulties In Speaking English And Perceptions Of Accents: A Comparative Study Of Finnish And Japanese Adult Learners Of English". 2013, pp. 26-103.
- Peters, Gary. **Improvising Improvisation**. The University Of Chicago Press, 2017.
- Pollastri, Neri. "Improvisation. A 'Method' Of Philosophical Consultation". 2006, http://www.ub.edu/practicafilosofica/arxiu/N_Pollastri.pdf.
- Prudence, Paul. "Machine Fallibility & Emergent Weirdness - When The Ghost In The Machine Sings Of Its Broken Ontology". **Neural Magazine**, 2017, pp. 22-25.
- Reporter, Weekly. "The Father Who Breeds Geniuses". **The M&G Online**, 2017, <https://mg.co.za/article/1989-12-01-00-the-father-who-breeds-geniuses>.
- Rittel, Horst W. J., and Melvin M. Webber. "Dilemmas in A General Theory Of Planning". **Policy Sciences**, vol 4, no. 2, 1973, pp. 155-169. **Springer Nature**, doi:10.1007/bfo1405730.
- Rühlemann, Christoph, and Matthew Brook O'Donnell. "Introducing A Corpus Of Conversational Stories. Construction And Annotation Of The Narrative Corpus". **Corpus Linguistics And Linguistic Theory**, vol 8, no. 2, 2012. **Walter De Gruyter GmbH**, doi:10.1515/cllt-2012-0015.
- Sandoval, Kristopher. "What Is The Difference Between An API And An SDK?". **Nordic Apis**, 2016, <https://nordicapis.com/what-is-the-difference-between-an-api-and-an-sdk/>.

- Santino, Catherine. "I Used Gmail's Automated Responses For 24 Hours Straight And It Got Weird". *Levo League*, 2017, <https://www.levo.com/posts/i-used-gmail-s-automated-responses-for-24-hours-straight-and-it-got-weird>. Accessed 10 Feb 2018.
- Scott, Jeffrey. "Improvisation In The Theatre: An Intersection Between History, Practice, And Chaos Theory (MFA Thesis)". 2014.
- Serban, Iulian Vlad et al. "A Survey of Available Corpora For Building Data-Driven Dialogue System". 2015, p. 51. *ARXIV*, doi:arXiv:1512.05742.
- Serban, Iulian Vlad et al. "A Survey of Available Corpora For Building Data-Driven Dialogue Systems". *Eprint Arxiv:1512.05742*, 2017. *ARXIV*, <https://arxiv.org/abs/1512.05742>.
- Shang, Lifeng et al. "Neural Responding Machine For Short-Text Conversation". *Corr*, abs/1503.02364, 2015. *ARXIV*, <http://dblp.org/rec/bib/journals/corr/ShangLL15>.
- Sicart, Miguel. *Play Matters*. MIT PRESS, 2017, p. 3.
- Snelgrove, Xavier. "Dango - Your Emoji Assistant". *Getdango.com*, 2017, <https://getdango.com/emoji-and-deep-learning/>.
- Spolin, Viola. *Improvisation For The Theater*. Northwestern University Press, 1999, p.4.
- Spry, Tami. "Performing Autoethnography: An Embodied Methodological Praxis". *Qualitative Inquiry*, vol 7, no. 6, 2001, pp. 706-732. *SAGE Publications*.
- Sterritt, David. "Revision, Prevision, And The Aura Of Improvisatory Art". *The Journal Of Aesthetics And Art Criticism*, vol 58, no. 2, 2000, pp. 163-172.
- Strope, Brian, and Ray Kurzweil. "Efficient Smart Reply, Now For Gmail". *Research Blog*, 2017, <https://research.googleblog.com/2017/05/efficient-smart-reply-now-for-gmail.html>.
- Sugiyama, Hiroaki et al. "Open-Domain Utterance Generation For Conversational Dialogue Systems Usingweb-Scale Dependency Structures". *SIG-DIAL 2013 Conference*, 2013, Accessed 9 Feb 2018.
- Tanenbaum, Joshua, and Karen Tanenbaum. "Improvisation And Performance As Models For Interacting With Stories."
- Trouvain, Jurgen et al. "On the Assessment of Computer Assisted Pronunciation Training Tools". 2017.
- Turkle, Sherry. *The Second Self*. MIT Press, 2005, p. 294.
- Turkle, Sherry. *Reclaiming Conversation*. 1st ed., Penguin Books, 2015, pp. 16-17.
- Turkle, Sherry. "Opinion — The Flight From Conversation". *Nytimes.Com*, 2018, <https://www.nytimes.com/2012/04/22/opinion/sunday/the-flight-from-conversation.html>. Accessed 9 May 2018.
- Weizenbaum, Joseph. "ELIZA—A Computer Program For The Study Of Natural Language Communication Between Man And Machine". *Communications Of The ACM*, vol 9, no. 1, 1966, pp. 36-45. *Association For Computing Machinery (ACM)*.
- Williams, Garnett P. *Chaos Theory Tamed*. Joseph Henry Press, 2001.
- Winograd, Terry. "Shifting Viewpoints: Artificial Intelligence and Human - Computer Interaction". *Artificial Intelligence*, vol 170, no. 18, 2006, pp. 1256-1258.
- Winograd, Terry, and Fernando Flores. *Understanding Computers and Cognition*. Norwood, 1986, p. 162.
- Zimmerman, John et al. "From Design Research To Theory: Evidence Of A Maturing Field". 2009.

A | REB APPROVAL

Dear Ms. Judith Doyle, Miss Yawen Guo,

The Research Ethics Board has reviewed your application titled "ImprovChat - A machine learning powered dialog assistant with knowledge of creative expression". Your application has been approved. You may begin the proposed research. This REB approval, dated January 22, 2018, is valid for one year less a day: January 21, 2019. Your REB number is: 2018-09.

Throughout the duration of this REB approval, all requests for modifications, renewals and serious adverse event reports are submitted via the Research Portal.

Any changes to the research that deviate from the approved application must be reported to the REB using the amendment form available on the Research Portal. REB approval must be issued before the changes can be implemented.

To continue your proposed research beyond January 21, 2019, you must submit a Renewal Form before January 14, 2019. REB approval must be issued before research is continued.

If your research ends on or before January 21, 2019, please submit a Final Report Form to close out REB approval monitoring efforts.

If you have any questions about the REB review & approval process, please contact the Christine Crisol Pineda, Manager, REB secretariat.

If you encounter any issues when working in the Research Portal, please contact our system administrator.

Sincerely,

Nancy Snow Chair, Research Ethics Board

B | REB CONSENT FORM

Date: November 1, 2017
Project Title: ImprovChat

INVITATION You are invited to participate in a study that involves research. The purpose of this study is to gather feedback and documentation for the iterative design of an artificial intelligent agent, or chat bot, to facilitate the user to practice the improvised performance games on the digital application, suggest the user more authentic English expressions for the use in different conversational scenarios and help the user to build the creative expression skills.

WHAT'S INVOLVED As a participant, you will be asked to interact with this prototype individually and will offer your feedback which will dictate the development of the prototype. The Researcher will note down the read-outs of the web application. The user test and interview should take approximately 20 minutes. You will take approximately 30 minutes in each user testing and 1 hour in total of your time. Video, voice recording, and still images will be taken during this session for the purpose of archival project documentation, and project promotion.

POTENTIAL BENEFITS AND RISKS Possible benefits of participation include the opportunity to provide feedback, get a direct honorarium of 20 dollars, and to have an interesting interaction with a custom device, while also contributing to taking a critical stance towards connected devices. There also may be risks associated with participation. There may be risks associated with participation, such as general frustration associated with understanding and interaction of the prototype.

CONFIDENTIALITY Due to the nature of user testing and documentation involving image and voice, confidentiality can not be totally provided, however participants will have the option to be identified in all project data by one of the following: 1. Their full name, 2. By the first name only, 3. Anonymous - by alias or participant number (e.g. participant 1, 2, 3...). Data (voice / image / video) collected during this study will be stored on a password protected thumb drive held by the research. I plan to keep the data (voice / image / video) indefinitely. In my future endeavors, I might undertake Ph.D. studies and I predict that my research in my Master's studies will be most valuable then. Access to this data will be restricted to Yawen Guo, and Judith Doyle.

VOLUNTARY PARTICIPATION Participation in this study is voluntary. If you wish, you may decline to answer any questions or participate in any component of the study. Further, you may decide to withdraw from this study at any time, or to request withdrawal of your data prior to data analysis in February 28, 2018, and you may do so without any penalty or loss of benefits to which you are entitled.

PUBLICATION OF RESULTS Project documentation, archival and promotional material will be used at conferences, be used on the internet in the context of a promotional video, and portfolio images, and be used at a stu-

dent thesis defense at OCADU in April 2018. Feedback about this study will be available after June 2018, and you can contact Yawen Guo at 3150089@student.ocadu.ca.

CONTACT INFORMATION AND ETHICS CLEARANCE If you have any questions about this study or require further information, please contact the Principal Investigator (Yawen Guo) or the Faculty Supervisor (where applicable) (Judith Doyle) using the contact information provided above. This study has been reviewed and received ethics clearance through the Research Ethics Board at OCAD University [1879]. If you have any comments or concerns, please contact the Research Ethics Office through email.

CONSENT FORM I agree to participate in this study described above. I have made this decision based on the information I have read in the Information-Consent Letter. I have had the opportunity to receive any additional details I wanted about the study and understand that I may ask questions in the future. I understand that I may withdraw this consent at any time.

Name: _____

Signature: _____

Date: _____

Thank you for your assistance in this project. Please keep a copy of this form for your records.

C | LETTER OF INVITATION

Date: November 1, 2017
Project Title: ImprovChat

I, Yawen Guo, Graduate researcher, from the Department of Graduate School, OCAD University, invite you to participate in a research project entitled ImprovChat.

The purpose of this research project is to develop an web application called ImprovChat. The prototype avail AI technology to build an intelligent agent, or chat bot, to facilitate the user to practice the improvised performance games on the digital application, suggest the user more authentic English expressions for the use in different conversational scenarios and help the user to build the creative expression skills. Should you choose to participate, you will be asked to interact with this prototypes individually and will offer your feedback which will dictate the development of the prototype.

The expected duration is 30 minutes in each user testing session. The first user testing will be set up at the beginning of December, 2017 and the second user testing will be set up in January, 2018.

TTC travel expenses will be covered either by the Researcher or the participants, and healthy snacks will be provided at each session.

If you have any pertinent questions about your rights as a research participant, please contact the OCAD University Research Ethics Officer.

If you have any questions, please feel free to contact me (see below for contact information).

Thank you,

[Insert Principal Investigator's Signature]

This study has been reviewed and received ethics clearance through OCAD University's Research Ethics Board.

D | LETTER OF FEEDBACK

Date: November 1, 2017

Dear (Insert Name of Participant),

Thank you so much for your participation in testing The ImprovChat. Your feedback was quite helpful, and over the next few months I will be taking your suggestions on improvements under consideration.

Thank you as well for letting me record your interaction with the device. This will be very helpful as archival project footage.

I will keep you informed of the upcoming thesis show dates. I plan on sharing this documentation at conferences, presentations, and online as project promotion and archival images. If you have concerns, or questions, or if you would like to withdraw from this study, please con-tact me at.

This study has been reviewed and received ethics clearance through the Ontario College of Art And Design University (file #).

Best,

Yawen Guo

E

QUESTIONNAIRE FOR ELL

- Overall, did you understand the English clearly during this experiment? Please choose one of the following:
1) very well 2) well 3) acceptable 4) poor 5) very poor
- Did you decide to use any of the phrases provided by the ImprovChat during the chat exchange? Please choose one of the following:
1) often 2) not much 3) not at all
- If you chose 1 or 2, please rank the quality of the ImprovChat phrases using one of the following:
1) great 2) good 3) acceptable 4) poor 5) very poor
- Did the chat seem humorous at any point? Please describe below:
- Did the chatbot seem to have a personality? If so, please describe it in a few words:
- Would you use the ImprovChat application in the future?
1) Yes 2) Maybe 3) No
- Please, add any further comments you would like to make about the ImprovChat interaction.

F

QUESTIONNAIRE FOR NES

- Overall, how would you rank the clarity of English language during this experiment? Please choose one of the following:
1) very clear 2) clear 3) acceptable 4) poor 5) very poor
- Could you tell when phrases provided by the ImprovChat were used during the chat exchange? Please choose one of the following:
1) often 2) not much 3) not at all
- If you chose 1) or 2), please rank the quality of the ImprovChat phrases using one of the following:
1) great 2) good 3) acceptable 4) poor 5) very poor
- Did the chat seem humorous at any point? Please describe below:
- Did the chat communication seem to have a particular tone (for example, humorous, playful, confusing, scattered...)? If so, please describe it in a few words:
- Would you recommend the use of the ImprovChat application in the future?
1) Yes 2) Maybe 3) No
- Please, add any further comments you would like to make about the ImprovChat interaction.

G | USER FLOW OF PROTOTYPE ONE

Press the button 'Start to Play'

Wait for the Speech Recognition start to work

Listen to the bot to ask the question

User gives one answer

User click the 'Answer' button to get suggestion

User use the suggestion as one answer

User interact with the bot until they make five answers

Speech Recognition stop automatically

Case 1 - Five things you will get a request for giving five things for a random suggestion.
[View the custom dictionary](#)



Answer

localhost:8080/game2



ANSWER

East Grey

Mango Indica

East Grey

green tea

Mango Indica

East Grey

green tea

mango indica

Cloud

East Grey

green tea

mango indica

black dragon

cloud

localhost:8080/game2

Press the button 'Start to Play'

Game 2: give a word that sounds similar



You said: ...
Bot said: *devised*

Wait for the Speech Recognition start to work

localhost:8080/game2

Listen to the bot to speak the word

You said: ...
Bot said: *devised*

User speak a word that sounds alike

You said: *divide*
Bot said: *divine*

Wait for the bot to response

You said: *Vine*
Bot said: *vie*

User interact with the bot until they stop with a long silence

You said: *wide*
Bot said: *wine*

Speech Recognition stop automatically

localhost:8080/game2



RECAST.AI (AI CHATBOT)

Recast.ai is a collaborative platform to build, train, deploy and monitor intelligent bots for developers. I employ Recast.ai to extract user input information. The user input information will be parsed as a JSON file which includes intents and entities. It responds and adapts quickly so I can keep up-dating it. It consists of a mix of rule based NLP method and abundant examples so a developer can learn it in a short time.

END-TO-END DIALOGUE SYSTEMS

End-to-end dialogue systems can be divided into two categories: those that select deterministically from a fixed set of possible responses, and those that attempt to generate response by keeping a posterior distribution over possible utterance (Serban et al. 6). As Serban has written, "end-to-end dialogue system architectures based on neural networks have shown promising results on several dialogue tasks. In their purest form, these models take as input a dialogue in text form and output a response (or a distribution over response)" (6). Google researchers proposed the recurrent neural network conversational model by applying the sequence to sequence framework. They "experiment with conversations obtained from a noisy dataset of movie subtitles, and find that the model can hold a natural conversation and sometimes perform simple forms of common sense reasoning" (Vinyals and Le 1). Another group of researchers proposed a persona-based neural conversation model. "The Speaker Model integrates a speaker-level vector representation into the target part of the SEQ2SEQ model" (Li et al. 2).

Through the literature review and discussion with machine learning engineers, I know that the algorithm RNN is one of the main stream for phrase-based statistical machine learning. RNN Encoder-Decoder naturally generates a continuous-space representation of a phrase. "It is clear that the RNN Encoder-Decoder captures both semantic and syntactic structures of the phrases." Researchers have indicated that their model is capable of capturing linguistic regularities in the phrase pairs and proposing well-formed target phrases (Cho et al.). Researchers from Noah's Ark Lab propose employing a neural encoder-decoder for solving the response generation problem in open-domain conversation modeling, named Neural Responding Machine (NRM). "NRM essentially estimates the likelihood of a response given a post. The estimated probability should be complex enough to represent all the suitable responses" (Shang et al. 2). Due to my own limitations in knowledge, I am not able to adapt their algorithm into my prototyping, but NRM shows me that my design can be achieved when I attain enough technological supports.

INFORMATION RETRIEVAL AND RANKING-BASED SYSTEMS

Information retrieval and ranking-based systems are systems that search through a database of dialogues, and pick responses with the most similar context. In this case, the mapping function projects the dialogue history into a Euclidean space (e.g. using TF-IDF bag-of-words representations). The response is then found by projecting all potential responses into the same Euclidean space, and the response closest to the desirable response region is selected. This task only requires knowing the sequence of utterances, which can be extracted automatically from transcribed conversations (Serban et al. 6). Word embedding is a set of feature engineering techniques that map sparse word vectors into continuous space based on the surrounding context. Word meanings can be well described with numerical vectors, often reflecting co-occurrence frequencies with other words in corpora (Lê 49). This vector representation provides useful properties for comparing words or phrases. For example, if "salt" and "seasoning" appear within the same context, the model will indicate that "salt" is conceptually closer to "seasoning," than, say, "chair." (Dalinina). A rich and diverse literature has emerged on distributional semantics, with many successes in tasks like similarity judgements, word sense disambiguation and information retrieval (Lenci). I can use the word embedding model to build my multidimensional association dictionary.