

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

A Multiagent Architecture to Multilingual Interaction

EDUARDO FRANÇA

VITÓRIA

2017

EDUARDO FRANÇA

A Multiagent Architecture to Multilingual Interaction

Dissertação apresentada como requisito parcial à
obtenção do grau de Mestre em Informática.

Programa de Pós-Graduação em Informática.

Universidade Federal do Espírito Santo.

Orientador: Prof. Dr. Orivaldo de Lira Tavares.

VITÓRIA
2017

EDUARDO FRANÇA

A Multiagent Architecture to Multilingual Interaction

Dissertação apresentada ao Programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Aprovada em 11 de agosto de 2017.

Comissão Examinadora

Prof. Dr. Orivaldo de Lira Tavares
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Crediné Silva de Menezes
Universidade Federal do Rio Grande do Sul
Examinador interno

Prof. Dr. Alberto Nogueira Castro Junior
Universidade Federal do Amazonas
Examinador externo

Dedico essa dissertação à minha família, que de perto ou longe sempre tem me apoiado. E ao professor Orivaldo de Lira Tavares, sem o qual esse trabalho não teria sido concluído.

AGRADECIMENTOS

Agradeço ao meu orientador por ter me aceito de braços abertos no Laboratório de Informática na Educação, pela paciência com minha falta de foco, por brigar para que eu tivesse condições materiais de me manter no programa e, principalmente, pela persistência de me guiar até o término dessa dissertação, quando até eu duvidei de minha capacidade.

À minha família e amigos por entenderem minha ausência durante esse mestrado. Principalmente a minha mãe, Edileuza, por suportar meu temperamento imprevisível e comentários sem sentido, durante essa e muitas outras jornadas.

À minha companheira intergalática de reclamações sem fim com o estado das coisas, Larissa. Mesmo que tenhamos sido triturados por esse moinho, ainda hemos de reclamar muito dos absurdos quixotescos desse mundo de maravilhas.

Aos professores membros do LIEd — Crediné, Dedê e Orivaldo — que me proporcionaram o desequilíbrio necessário para construir novos conceitos e encarar os antigos com um novo olhar. Aos amigos do laboratório, Bruno, Tuler, Pinotti, Marques, Camila, Patrícia, pelos altos e baixos da vida acadêmica que compartilhamos semestre a semestre.

Aos demais professores, alunos e colaboradores do PPGI e do Departamento de Informática, numerosos demais para nomear, e com os quais convivi por todos esses meses de mestrado, como aluno, colega de trabalho, ou apenas mais um rosto conhecido correndo de um ponto ao outro do campus.

Por fim agradeço à CAPES e ao Ministério da Educação, por proporcionarem as condições materiais necessárias para que eu me dedicasse à essa pesquisa, e também pelo ensino da rede pública e do programa Universidade para Todos, sem os quais eu não teria chegado aqui. Educação é a chave.

“Todo conhecimento humano é incerto, inexato e parcial.”

Bertrand Russell

RESUMO

Esta dissertação apresenta um ambiente virtual especialmente construído com recursos para auxiliar na comunicação entre falantes de idiomas diferentes, voltado a possibilitar a interação em ambientes textuais entre pares ou grupos de pessoas, dentro de comunidades de aprendizagem que não possuam um idioma em comum. Nesse ambiente existe um agente digital (ou bot) que usa estratégias de comunicação para auxiliar os interlocutores na compreensão de mensagens em idiomas estrangeiros, podendo também definir termos técnicos, regionalismos e expressões, de acordo com dicionários e serviços usados. Os recursos do ambiente foram focados em interações síncronas, porém também podem ser usados em interações assíncronas com algumas adaptações. Além do suporte essas comunicação entre falantes de idiomas diferentes, as interações são armazenadas para pesquisa sobre o efeito dos agentes, estratégias e recursos sobre a comunicação.

Palavras-chave: mediação; agentes inteligentes; tradução de máquina; estratégias de comunicação;

ABSTRACT

This dissertation presents a virtual environment specially constructed with resources to aid in the communication between speakers of different languages, aimed at enabling interaction in textual environments between pairs or groups of people, within learning communities that do not have a common language. In this environment there is a digital agent (or bot) that uses communication strategies to assist the interlocutors in the understanding of messages in foreign languages, and may also define technical terms, regionalisms and expressions according to dictionaries and services used. The environment resources were focused on synchronous interactions, but can also be used in asynchronous interactions with some adaptations. In addition to supporting such communication between speakers of different languages, interactions are stored for research on the effects of agents, strategies, and resources on communication.

Keywords: mediation; Intelligent agents; Machine translation; Communication strategies;

LISTA DE FIGURAS

2.1 Text and Graphic mode Dialogs	29
2.2 Dr Sbaitso, a simple chatbot	29
3.1 Multilevel Architecture for a Multiagent Environment for Synchronous Interactions	41
3.2 The Pilot	44
3.3 SIGMO Talk & Understand	44
3.4 Clik	45
6.1 Chat with Translation - example of use	61
6.2 Multiagent Architecture to Support Multilingual Interaction	62
6.3 Agent State Model	68
6.4 Agent Action States	69
6.5 Answering Users' Commands	70
6.6 Clarification Attempt Steps	71
6.7 Consensus Collaborative Steps	72
6.8 Rephrasing for Clarification	73
7.1 Web Session with Two Participants	77
7.2 Access with a mobile client	78
7.3 Agent Event Log	79

LISTA DE QUADROS

Table 4.1 Important Dates in Machine Translation

47

LISTA DE SIGLAS

ALPAC - Automatic Language Processing Advisory Committee
API - Application Platform Interface
AVR - Automatic Voice Recognition
BLEU - Bilingual Evaluation Understudy
CBMT - Constraint Based Machine Translation
EBMT - Example Based Machine Translation
EF - Education First (International English School)
EFL - English as a Foreign Language
HICATS - Hitachi Computer Aided Translation System
HTTP - Hypertext Transfer Protocol
IRC - Internet Relay Chat
KBMT - Knowledge Based Machine Translation
LBMT - Lexical Based Machine Translation
LIEd - Laboratory Information Technology in Education
LVCSR - Large Vocabulary Continuous Speech Recognition
MT - Machine Translation
NMT - Neural Machine Translation
PBSMT - Phrase Based Statistical Machine Translation
RBMT - Rule Based Machine Translation
RNN - Recurrent Neural Network
SMT - Statistical Machine Translation
SL - Source Language
SP - Source Participant (who sent a message)
TISE - Taller Internacional de Software Educativo (Conference)
TL - Target Language
TP - Target Participant (message destinationary)
TTS - Text to Speech
XML – Extensible Markup Language

SUMÁRIO

1. Introduction.....	16
1.1 Rationale.....	18
1.2 Motivation.....	19
1.3 Objectives.....	19
1.3.1 General Objectives.....	19
1.3.2 Specific Objectives.....	20
1.4 Research Questions.....	21
1.5 Use Case Scenarios.....	22
1.5.1 Two people exchange.....	22
1.5.2 Two People Exchange With Human Helper.....	22
1.5.3 Mixed Language Group Lecture.....	23
2. Conceptualization.....	25
2.1 Translation.....	25
2.2 Machine Translation.....	28
2.3 Dialog Systems.....	29
2.4 Data Scraping.....	30
2.5 Text Corpus.....	30
2.6 Communication theory.....	30
3 Related Work.....	33
3.1 Academic Projects.....	33
3.1.1 Machine Translation Effects on Communication.....	33
3.1.2 Two is Better Than One: Improving Multilingual Collaboration by Giving Two Machine Translation Outputs.....	36
3.1.3 Exploring Communication Breakdown in Global Virtual Teams.....	37
3.1.4 What is a Better Translation? Reflections on Six Years of Running Evaluation Campaigns.....	38
3.1.5 Assessing the impact of real-time machine translation on multilingual meetings in global software projects.....	39
3.1.6 Ambiente Multiagente para Interações Síncronas.....	40

3.2 Commercial Projects.....	43
3.2.1 Not-so-Universal Translators.....	43
3.2.2 Online Machine Translation.....	45
4 Brief History of Machine Translation.....	46
4.1 Important Dates in MT History.....	47
4.2 State of the Art.....	49
4.3 Machine Translation Paradigms.....	50
4.3.1 Rule Based Machine Translation - RBMT.....	51
4.3.2 Knowledge Based Machine Translation - KBMT.....	51
4.3.3 Lexical Based Machine Translation - LBMT.....	51
4.3.4 Constraint Based Machine Translation - CBMT.....	52
4.3.5 Statistical Machine Translation - SMT/PBMT.....	52
4.3.6 Example Based Machine Translation - EBMT.....	52
4.3.7 Neural Machine Translation - NMT.....	52
5 Communication Strategies.....	54
5.1 Chunking.....	54
5.2 Goal Oriented Conversation.....	54
5.3 Explaining the Main Points.....	54
5.4 Talking Back.....	54
5.5 Avoiding Certain Words, Expressions, Constructions.....	55
5.6 Gesturing.....	55
5.7 Use of Images.....	55
5.8 Rephrasing.....	55
6 Architecture.....	57
6.1 Addressing Use Case Scenarios.....	57
6.1.1 Scenario C1.....	57
6.1.2 Scenario C2.....	57
6.1.3 Scenario C3.....	58
6.1.4 Other Scenarios.....	58
6.2 Agent Strategies.....	59

6.2.1 Machine Translation.....	59
6.2.2 Use of more than one MT output.....	59
6.2.3 Paraphrase / Paraphrasing.....	59
6.2.4 Dictionaries / Thesauri / Glossaries.....	59
6.2.5 Wikidata / DBpedia / OpenWordnet-EN / Princeton Wordnet.....	60
6.2.6 Collaborative Repair.....	60
6.2.7 Images.....	60
6.3 Multiagent Architecture to Support Multilingual Interaction.....	61
6.4 Agent State Model.....	68
 7 Prototype Implementation.....	 75
7.1 Examples of use of the prototype.....	76
 8 Final Considerations.....	 80
8.1 Return to research questions.....	80
Q1. Is it possible to promote communication between monolingual speakers of English and Portuguese with the use of technologies available on the web?.....	81
Q2. In what ways does this technology help (or hinder) communication?.....	81
Q3. What is the ideal scenario for the use of such technology?.....	81
Q4. Is it possible to build a multiagent architecture to support multilingual interaction?.....	82
8.2 Experiences acquired with the development of the environment.....	82
8.3 Future Research.....	83
 9 References.....	 85

1. Introduction

The main point of this dissertation is communication. Enabling it to happen between two people who do not have a language common. The language barrier is something that comes along with the diversity human culture communication itself. From the moment there was more than one language (or proto-language), we can imagine that there were also speakers who could not communicate because they did not know both languages or codes.

Today, in the twenty-first century, we have the consolidated certainty that our greatest asset is information. With the decentralization of knowledge repositories and the democratization of access created by the Internet, we see an arms race for the control of the media and its power to influence the paths of civilization.

And still, the communication barrier persists. Isolating communities, peoples, ideas. There are about 210 languages being spoken, only in Brazil, nowadays [OLIVEIRA 2008]. The complexity grows exponentially as we try to deal with regional variations, group and technical vocabulary.

English is the *de facto* language of the internet, of international trade and research. Not being able to speak English today is losing opportunities to learn, exchange information, do business, meet people, and have greater access to entertainment, news and all kinds of opportunities. The situation gets even harder in Research and Higher Education, where many researchers prioritize the publication of their articles in English [SCIENCE WATCH, 2009] [MOED, 2012]. In some areas, not knowing this particular language means lagging months behind 'state of the art' knowledge, especially in the Natural and Applied Sciences.

The past six decades have seen the creation of tools and methods that help translators, learners and monolingual speakers have access to information created in several languages. The interpretation, translation and post-editing market continues to develop, and so does the speed at which we create and publish content. Digital aid tools for translation, translation memories, *wordnets* and encyclopedias moved by the power of *crowdsourcing* are some of the tools that have been maturing since the beginning of the use of computers as a centerpiece in access to information.

Even with the large amount of tools available, the communication problem still prevents us from accessing opinions and content relevant to our interests. Most research engines still confine searches to the language in which they were made. And if the barrier is there when deal with multimedia content, we can assume that the situation worsens when dealing directly with people. How many interactions between

teachers, students, content creators and consumers did not fail to occur because of the language barrier.

This work proposes the creation of a system that mediates communication between speakers of Portuguese and English, and helps participants find artifacts generated in the message translation process, fix the errors, and finally, to have greater success in communicating with other people even having virtually zero prior knowledge of their language.

Even though most of the work here focus on Portuguese-English interaction, it is only so because of the prominence of the English language in the world stage today, since so much of the media we consume comes from sources that have English as their main language. Which is also true in the world of business and even more so in academic circles. That said, the only barrier to using the ideas here discussed to any other language pair is existence of enough tools for translation and reference between the intended language-pair.

The recording of this kind of interaction can be change of better understanding of how digital resources used interfere with communication, and use them to improve the system. New terms and phrases can be collected, serving as raw material to the creation of glossaries. Which in turn can be fed back into the system and be available as reference material.

1.1 Rationale

English language proficiency levels are still very low in Brazil, we are the 41st nation in EF's English Proficiency Index [EDUCATION FIRST, 2015]. Especially if compared with countries labeled as Developed World [EDUCATION FIRST, 2015].

If we take into account that the largest amount of content published on the Internet is in English [WEIJEN, 2012], we see that the opportunities are smaller for those who have no command of English.

Another reason is that we cannot expect everyone to spend a considerable amount of their time learning the language of another nation. There are people who would do that even if we had already solved the problem of cheap and reliable High Quality Machine Translation. But not all of us. And we have lives to live, places to be, people to meet people, thought-provoking ideas to explore. Ideas which Konstantin Andreev expresses in his English as a Lingua Franca short course [ANDREEV, 2011], dealing with questions such as the linguistic power imbalance every non-native English speaker suffers when being forced to write and speak in a language that is not their own, on a situation of disadvantage in comparison with native speakers, and also pouring into language learning time that would have been invaluable if spent in their specific fields of work or research.

This work aims to find methods by which Information and Communication Technologies can help mitigate the impact of this problem. And from the beginning we'd like to clarify that this technology does not intent to substitute learning a language or employing professional translators, but can be used when none of the previous, more complete solutions, are viable. And ideas here can also be applied in parallel with more traditional ones, as an aid for the communication.

1.2 Motivation

As a former EFL teacher, having worked as a translator/interpreter and being interested in languages, it was clear to me from the beginning this research would have to involve languages and communication somehow.

And, even though the tools have been around for several years, the use technology to enable communication between foreign participants has yet to become an integrated routine in our lives. Specially when compared to the ubiquity of smartphones and social networks.

It is challenging to imagine something that is yet to come. But with the advance of the access to the internet, and projects such as Google's Project Loon [KATIKALA, 2014] and Facebook [ZUCKERBERG, 2013], there will come a moment when everyone in the world will be a text/voice message away.

As global as English is, it is not a single language [CRYSTAL, 2012]. Neither is Portuguese. Nor any language whose speakers are spread over large areas. As realities vary from place to place, a Portuguese speaker from Manaus will not live in the same linguistic reality from a Portuguese speaker of São Paulo, Lisbon, Díli or Macau. And the language is a living animal, being colored by local creativity, enriched by technical language, neologisms, coining, calques and several others processes.

Enabling communication between speakers of two unrelated and non intelligible languages, making it easier to interact with people from other countries in several settings, such as text chat rooms and services. Also helping people understand regional or technical expressions with the use of glossaries, dictionaries and other reference tools.

Also creating a structure that can be easily modified to include more language services, resources and routines.

1.3 Objectives

In this chapter we define general and specific objectives that guided the development of this research.

1.3.1 General Objectives

Create a solution that serves as a flexible, extensible and reusable tool to enable, ease or aid communication in multilingual environments. Use free and open source tools and technologies in this process. Generate data, to be used as a way of evaluating and improving the overall solution.

1.3.2 Specific Objectives

Enable communication between speakers of different languages in several scenarios (which are further discussed and specified).

Use free and open sources technologies to implement the solution. Technologies such as python, XML and the Unix Dict dictionary server.

Compile interactions in a corpus, showing how digital agent intervention affects communication.

1.4 Research Questions

This dissertation project seeks answers to the following research questions.

Q1. Is it possible to promote communication between monolingual speakers of English and Portuguese with the use of technologies available on the web?

Q2. In what ways does this technology help (or hinder) communication?

Q3. What is the ideal scenario for the use of such technology?

Q4. Is it possible to build a multi-agent architecture to support multilingual interaction?

1.5 Use Case Scenarios

Starting from scenario in which no digital resources are used, how to replace each of the elements so that we depend less and less of human agents.

Can we replace the human interpreter for a digital / synthetic? What are the advantages, what are the disadvantages. Which technologies are available and how do they fit into this knowledge negotiation process?

Think about a human-machine hybrid architecture to support interaction in multilingual learning communities.

A digital architecture, and in this case completely digital, with no need for a human interpreter. However there may be times of maintenance in which the system works in supervised learning mode, watching and learning from how humans solve the problems.

1.5.1 Two people exchange

Scenario C1. Two people who do not have a common language need to communicate to exchange information. this dissertation presents the design and development of digital resources to enable the communication between these two people. It focuses primarily on speakers of Portuguese and English. The tools needed to make this exchange possible and the ways in which a digital agent can help it take place.

1.5.2 Two People Exchange With Human Helper

Scenario C2. Two people who do not have a common language need to communicate to exchange information, but in addition to software resources, they have the help of a third participant who knows the two languages they are using, serving validator and facilitator of communication. On the first scenario we must describe the technological devices that are needed to help the interaction. In this one we have to deal with a situation where there is an agent that helps a human who speaks both languages to mediate the interaction between other two participants who

do not understand each others' tongue.

1.5.3 Mixed Language Group Lecture

Scenario C3. A group of users attending a lecture or class in a virtual space, the session is being guided by a teacher who does not speak the language of all users, and those (with varying levels of knowledge in the language, but with little knowledge) work together to share, correct and validate the significance of the foreign participant's message. In this scenario the interaction is largely unidirectional (teacher-> students), but the checking and validation of the translation/understanding is collaborative.

In all cases, what the system provides is a set of tools to generate, test and if necessary correct problems in the communication until the message gets through to other side. Due to imperfections and defects that are common in all machine translation methods, encyclopedic sources help users check the translation, and by using an established communication protocol and repair workflow, better adapt it to the message they believe to be more faithful to what the other speaker meant.

Elements required in a digital architecture to support interaction: if it was a face-to-face dialogue, and both speakers are speech and hearing-able, we need a way to decode speech to text, and then, after the translation process, to re-encode sound.

In the scenario we choose to explore better, we assume that the parties are already in a synchronous textual communication environment. There are several other situation in which this scenario could apply, but for the sake of simplicity we are not including here face-to-face, audio or video communication.

Dealing with textual communication, we can have people with several levels of knowledge of their partners' language, the agent in charge of helping the communication, when you do not have information on any of the participants, it assumes that they understand nothing of the language spoken by others.

(It is necessary to create a chat state model, stating what languages were spoken so far, how many problems of understanding were accused by the participants, which support level of the agent gives to each participant, etc.)

A crucial tool that the agent uses is Machine Translation, and there several options of online and offline MT systems that can be used. However, based on the interaction needs, other systems can be used to clarify communication problems.

A list of problems and possible solutions:

- When the user indicates not understanding the previous message, the agent notifies the partner that there was a problem of understanding and proceeds to ask for the help of the participants to solve the problem.
- The first question asked is "what didn't the user understand". If a small word or passage, than it translates back to the original language and requests that the message sender to clarify the passage with more information.
- The user can request word definitions in their own language or in the language(s) of other participants.
- The user can also ask for usage examples of specific expressions in the language of their chat partner.
- In the case of objects/nouns that are easier to explain by visual means, the participant can choose an image and send as part of an explanation.
- In the latter case, the agent can use models (or templates) that participants can fill to have a greater chance getting the message through.
- After a dialogue, the agent can generate representations participants can use to review and better understand what was said.

The agent also watches the conversation and alerts the users when a single user is monopolizing the conversation and the others are not either confirming they are understanding or signaling their doubts. (A participant tagged as teacher, guide or session leader will be observed under different rules, since these need more time if the session is based on a more expositive structure.)

2. Conceptualization

This chapter presents the theoretical basis for the design of this dissertation. Translation is a multidisciplinary area, as are all areas of knowledge that deal with language. Even more so when language is the focal point through which people must negotiate their diverse identities and cultural backgrounds.

2.1 Translation

We can say that from the proverbial time of the fall of Babylon humanity has been looking for ways to communicate with other "tribes". Through language and habits in it condensed, confronting different ways of representing, codifying and giving boundaries to reality. Then rises, the herein defined, concept of translation.

In his treatise *Linguistic Theory of Translation* J. C. Catford [Catford 1965] defines translation as "an operation in languages:.. The process of replacing text in one language (source) for equivalent text in another language (target)" it is thus a unidirectional operation.

Roman Jakobson also makes clear that the translation process goes beyond "the text itself" when he says that translating is represented with a message more or less equivalent to what was said in another language:

"Most frequently, however, translation from one language into another substitutes messages in one language not for separate code-units but for entire messages in some other language. Such a translation is a reported speech; the translator decodes and transmits a message received from another source. Thus translation involves two equivalent messages in two different codes. [JAKOBSON, 1959]

Jakobson divides the translation process into three categories [JAKOBSON, 1959]:

- Intralingual or rewording, which is the interpretation of verbal signs by means of other signs of the same language;

- Interlingual, which is the interpretation of verbal signs by means of another language, and this is the process to which most people refer when talking about translation;
- Intersemiotic or Transmutation, which is the interpretation of verbal signals through nonverbal signals systems.

A more recent definition of the types of translation can be found in Brislin (2007), categorizing translation into four types:

- Pragmatic, focused on the accuracy of the information contained in the SL message.
- Aesthetic and poetic, in which there is an effort to carry effect, emotion and feelings of the text in SL to the TL. This addition to the message itself.
- Ethnographic, takes into account the cultural and social context in which certain expression choices are made in the SL and tries to find the expression that generate the closest effect to TL speakers.
- Linguistics, tries to find in TL equivalent meanings to the ones found in the morpheme and grammatical structures of SL.

However there is no final or unified translation theory [SOUZA, 1998], different schools face and define translation in different ways. There are even people who deal with the "impossibility of translation".

But the translator who lives up to the name does not end their analysis with the text, but brings to it (or inside it) a whole baggage of life experiences, interpretations and world views.

It gets even more complicated here, because social, political and historical contexts are not easily translatable, specially when you are only supposed to hint at them in the text, and assumes familiarity on the part of the reader. You must think about what the author of the text was thinking, or at least in the word choice that arose from their thought process and what feelings it would cause in the reader the message was aimed at. Such reader might live in the same country of the author, or at least "live" in

the same language, and share habits, know about similar facts. It is up to the translator to decide which information is important and how to imprint it on the translation in a way the reader could discern it. All the while trying to remain hidden, preventing his own taste or judgment to leave its traces on a text that is, essentially, not theirs.

Umberto Eco says that you can only really understand what translation is after translating and being translated (as an author) [ECO 2003]. In his book "Almost the Same Thing" he discusses the fundamental problem of translation: the non-equivalence of meaning between the words of two languages.

Eco defines languages as tools that create or define the boundaries of the real, dictating what exists and does not. They delimit reality with words like door, house, window, car, computer, translator, programador, aluno, orientador, Wissenschaft, beca, université and so on. And each of these systems has a logic of its own.

Each word is a node in a network of meanings and connotations that is unique to that speaking community. And even within communities that superficially share the same language, there are variations of use and meaning depending on age, formal education, profession, creed and several other factors.

2.2 Machine Translation

Machine Translation (MT), also known as Automatic Translation, is one of the oldest areas of computing, formally existing since the 1950s.

MT's efforts are in order to decompose the process of translation into steps that a computer can execute. The previous section makes it clear that it is not a trivial task, even for humans with years of experience in the field.

But having said that, much of what we produce on a daily basis as we communicate is predictable, messages are largely composed of fixed expressions, obeying basic usage patterns. Word usage frequency follows Zipf's law in all human languages, and many other systems [BLACK, 2009].

We have the ability to create and make up new expressions, words and meaning, yet most of what we say and write are repetitions, variations on a theme. Slight variations of sequences being reproduced since long ago.

In this chapter we talk a little history of machine translation. A lot of the text here comes the Hutchkins's work, a historian of this branch that intersects computing and linguistics.

“No lack of grammatical device in the language translated into makes impossible a literal translation of the entire conceptual information contained in the original.”
[JAKOBSON, 1959]

2.3 Dialog Systems

Any system that interacts with the user through dialogues can be called a Dialog System. Here defined as defined as any visual or textual component that gives the user information on program execution, and interacts with them through sound, text or images.

In computing, dialog system used to refer to dialog boxes such as the ones shown in Figure 2.1.

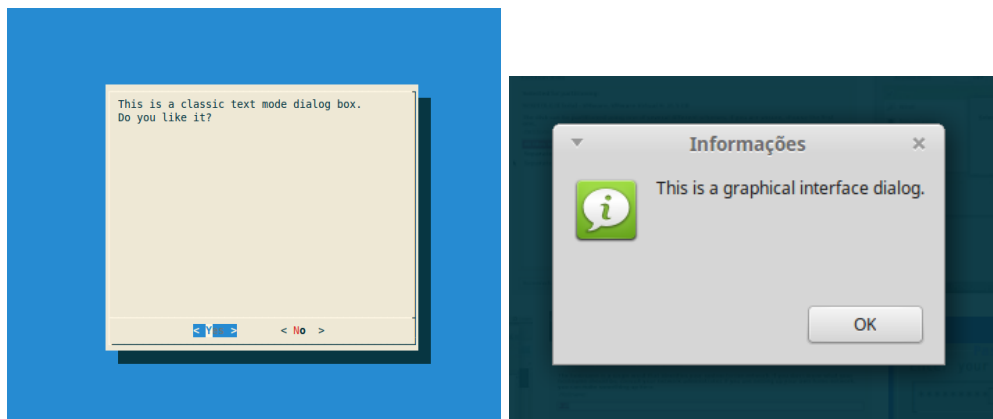


Figure 2.1 - Text and Graphic mode Dialogs [Fonte própria]

But with the popularization of chatbots “dialog system” also started to refer to systems that try to chat and have a “dialog”, a conversation with the user. Initially text based, such as Dr Sbaitso (Figure 2.2) and Eliza, these systems are the precursors to the personal assistants such as Apple’s Siri, Microsoft’s Cortana and Google Now.



Figure 2.2 - Dr Sbaitso, a simple chatbot [MATT, 2016]

2.4 Data Scraping

It is the act of extracting information from complex systems. In the case of this dissertation several online systems are used to create translations of users' messages. Other systems are used to obtain synonyms, antonyms, word definitions and classifications.

Scraping can also relate to the use of scripts and/or specialized software to gather and compile data from online services.

2.5 Text Corpus

Collections of text that may come from books, newspapers, movie subtitles, chat room interactions, email, social networks and several other sources. These usually represent how language is used by a certain group of people, at specified setting and certain moment.

A Corpus is the closest thing we have to taking a snapshot of language in moment in time. Greatly facilitating the process of experimenting with language hypothesis.

2.6 Communication theory

Is the field that theorizes and explains how communication is even possible. How we are able to translate our thoughts to sound waves, gestures, symbols, actions. Since our focus is led by technology and Information Theory, the way this work treats communication falls between what Craig [1999] describes as the Cybernetic and Sociocultural approaches to Communication Theory. The former deals with communication in terms of signals, channel noise, interference and information processing. The latter focuses on the social traits of communication, and how we reproduce learned behaviors and patterns that come from the society and environment we grew up in, at the same time influencing the whole network of interactions from which our reality emerges.

In addition to these two approaches, five other ways of dealing with communication are listed as the most significant in Communication Theory:

Rhetorical, which views communication as the practice or “art” of discourse, as striving to convince other and sway their judgment towards the ideas expressed by the speaker;

Semiotic, studies communication as an intersubjective mediation by signs. And as such, problems of communication are seen as problems in representation and transmission, when the parties involved do not have the same domain over the shared sign system;

Phenomenological, states that authentic communication can only be achieved through the direct and unmediated contact with others. It emphasizes that while communicating, what we feel in the “otherness” of our dialogue partners, paradoxically we feel that in ourselves. This tradition sees problems of communication as a manifestation of the need and impossibility of sustained authentic communication between those involved in the dialogue;

Sociopsychological, considers communication to be “a process by which individuals influence each other”. This process can occur directly, face-to-face, or can be mediated by any technological medium, and can also be one-to-one, one-to-many or many-to-many;

Critical, emphasises the always shifting nature of language, and how it is only meaningful when seen through the lenses of individuality through discourse, “Authentic communication occurs only in a process of *discursive reflection* that moves towards a transcendence that can never be fully and finally achieved — but the reflective process itself is progressively emancipatory.”

3 Related Work

In this section we explore some of the projects and research that aims to provide support for the same kind of problem situation, or that interface with the ideas of automatic translation and textual communication discussed in this dissertation.

3.1 Academic Projects

This section presents summaries of various academic projects on machine translation.

3.1.1 Machine Translation Effects on Communication

Yashiuoka and Bjørn (2011) touch in one of the central ideas this dissertation wants to explore: how the mediation of an MT reliant system interferes with the communication process.

They also add that, along with the inherited difficulties from translating from one language into another, there are also the problems that come along with textual communication, such as “typographical mistakes, breakdowns, accidental communication halts, inconsistencies and asymmetries of references pre and post translation” [YASHIUOKA, 2011]

Due that, but also to years of contact with imperfect and inadequate translation resulting from MT output, and to our very own bias against the belief that machines might one day outperform or even reach us in something so inherently human as “communication”, it is not a surprise the authors expected bad results:

Different from our expectation that quality of machine translation results determines communication process largely, our data indicates communication relies more on a dynamic process where participants establish common ground than on reproducibility and grammatical accuracy. (...) Instead, the largely for usage of project specific jargon contributes and maintaining common ground, despite establishing quality of translation

results.[YASHIUOKA, 2011]

The authors also note the importance of cultural differences in creating and maintaining this “common ground”:

Machine translation mediated communication embraces intercultural rooted in are mainly challenges, which computational communication, communication mediated artifacts and language technology. [YASHIUOKA, 2011]

Thus four possible outcomes might come from the use of this kind of system: (1) Participants can communicate even with the imperfections of MT; (2) Participants identify obvious translation mistakes and can still guess the meaning of the original message from the context of the communication and the parts MT got right. (3) Participants think MT mistranslated parts of the message and act accordingly, when in fact the translation was correct, taking for incorrect a message that in fact was right; (4): Participants take a mistranslation for a correct translation, and thus are misguided in the conversation, and continue to be so. In both case (3) and (4) the incorrect message will persist until other participants detect and point out the misunderstanding, if that ever happens.

They also point out how jargon (also community language or work language) is important in communication. And how having proficiency with the jargon in use can help a user maintain the common ground. The jargon used in this case was defined by the shared task, that asked participants to exchange information about places of touristic interest in a Japanese city. The participants spoke either English or Japanese as a first language, and had to communicate through the MT system provide by the Language Grid service.

An important finding is that while the MT system might be static, they way participants use language is not, as pointed out in:

(...) in order to get better translation, experienced users are known to adapt themselves to the system for example by simplifying and shortening their

messages to be processed preferably. Even less experienced students learned to prune sentences as well over time. Students mentioned in group notes how they adapted to the system. [YASHIUOKA, 2011]

Of course, it is important to highlight the problems that arise from this method of intercultural and inter-language communication:

From the analysis of intercultural collaboration process through machine translation mediated communication data, we identify four communication types. Interestingly, all teams mentioned difficulties of understanding translated messages in group notes and reflections on questionnaires, however, only one out of 15 teams could not complete travel plan. In other word, in spite that almost all teams faced language technological challenges caused by machine in task their they managed to translation, achieve collaboration with their informants. [YASHIUOKA, 2011]

So, even though most participants complained about the inaccuracies of machine translation, the majority were able to finish the assigned task.

Which is an evidence towards the premise that (a) MT might have evolved enough that meaning and sense can pass through to other side, at least to a minimal level, otherwise they would not have been able to complete the tasks; and (b) even though people are picky about what they consider "perfect" or "proper" renditions of their own languages, it is possible that several of the quirks, words, expressions, syntax and grammar we use, are more of a cultural habit than an actual requirement for communication to take place.

The authors also state that even though improvement in MT could clearly be beneficial, "seeking only for improvements of translation quality might not be an ultimate approach, at least for the time being" and that "beneficial resources and strategies, guessing and use supplement project jargon", pointing that engaging the human element might be just as important.

3.1.2 Two is Better Than One: Improving Multilingual Collaboration by Giving Two Machine Translation Outputs

Gao et al (2015) describes an experiment where English speakers were paired with Mandarin speakers and asked to perform an activity in which one of the participants had to give directions to the other, helping them to navigate a map and reach specific points of interest.

GAO et al (2015) assert that use MT to facilitate an exchange across languages brings not only benefits but also costs to the communication process. Even though people can use their own languages freely, know the system can handle it, the fact they cannot be sure of the specific wording the system is going to use can be detrimental keeping the common ground, specially when the point being discussed has to do with a specific thing and the MT output uses different translations for different sentences.

According to the authors, using two machined translated versions of the same message gives the participants more chances to work out what the original message meant:

Showing two translations rather than one makes the MT system more transparent, exposing seams in the underlying infrastructure. Showing the fact that there are alternatives doesn't just provide a resource for repair; it also foregrounds the fact that translation is happening in the first place and that it is imperfect. [GAO et al, 2015]

On a deeper reflection about not only MT based dialog systems but also about our world and lives are influenced by technology, the authors state these systems that decide what search results, ads, videos and songs we get see, such systems are constantly making assumptions and inferring about our behavior, and in this process they make correct as well as incorrect choices. Raising awareness and bringing this decision making to the foreground have positive effect in how we use and experience technology.

Systems that do filtering, translation, recommendation, and other tasks for us make choices. Those choices often have data attached to them that could help people decide whether the choices are appropriate. [GAO et al, 2015]

There's also a passing mention of other methods that can improve kind of mediated communication, such as using images together with that text and highlighting the most important part of the message.

They conclude by stating the benefits of using translation pairs outweighs the cognitive costs.

3.1.3 Exploring Communication Breakdown in Global Virtual Teams

In a research, DAIM et al (2012) study how cultural and linguistic differences affect global virtual teams (GVTs) communicating via technological means. And in this dynamic and diverse environment they explore how the extreme loss of context can make the dialog impossible, here referred to as "communication breakdown".

Dispersed work groups in which members are located in different countries face unique cultural differences that can affect the overall success of the group's performance. Language barriers in differences in cultural expectations and understandings can have a profound impact on the group's overall performance. Cultural differences that lend to communication issues tend to manifest themselves in three major categories: the functional disciplines of the members of the group, the organizational structure of the company and, the nationalities of the group members and/or the nation in which the groups exist. Functional differences, for example, can lead to group members with different knowledge bases, reasoning abilities, motivations, and like-minded thinking approaches. [DAIM et al, 2012]

They go on to exemplify how, even though language and culture might lead people to act differently in when faced with a similar situation, characteristics such as area of study might do just the opposite, and put them into a supranational community of teachers, physicians, political scientists or engineers, for example:

“Engineers, for instance, reason differently, react differently, and are motivated differently than marketers. Any inherent differences have been reinforced over time through training and exposure to other like-minded people.” (Smith as cited in DAIM et al. 2012).

DAIM et al conclude that Global Virtual Teams tend to perform poorly due to the differences in mindset they bring from their national cultural backgrounds. But at the same time, participants who understand the company culture can use this shared set of beliefs to better understand their coworkers from different nationalities.

Ten members of the team were interviewed, interesting finding include the use of less aggressive language than the one used in face to face communication, because textual communication leaves records behind:

[the] Virtual team exhibited good communication behavior whereas co-located senior team had more conflict. People avoid or should avoid big flames in electronic exchanges to avoid misinterpretation and written record.
[DAIM et al, 2012]

Other point to initial difficulty of dealing with cultural differences, “Took a while to get to ideology understanding between various nationalities.” and show how shared corporate culture gained prominence when national cultural conventions could not be relied upon “Corporate culture was more prevalent as opposed to their local culture.”

3.1.4 What is a Better Translation? Reflections on Six Years of Running Evaluation Campaigns

Phillip Koehn is the director of the MOSES project, an open source Statistical Machine Translation toolchain, and one of the researchers at the forefront of MT development.

In Koehn (2010) he discusses the strengths and weaknesses of MT quality metrics used by the industry. Koehn defines translation as:

Machine translation, and even more so human translation, are long standing efforts that aim to re-create a document in a different language that contains the same meaning as the original language document. Since the task

involves meaning at its core, we are confronted with all the unsolved problems of representation, equivalence, and similarity. [KOEHN, 2010]

He points out that, due to the nondeterministic nature of language and the way human mind works, not only will different translator come up with different translations to a passage, but the same translators might also settle for a different translation version for the same sentence if asked at a different time.

According to him, the most popular metrics for evaluating MT output (BLEU) when trying to rate translation quality can barely distinguish human and machine translations, when able at all to do it.

The author also points out it is important to view these systems under the light of what is their purpose and how they either fail or succeed in them:

If the application of machine translation is the understanding of foreign language content, then an evaluation method may be framed as: Given machine translation output, can a monolingual target side speaker answer questions about it? [KOEHN, 2010]

3.1.5 Assessing the impact of real-time machine translation on multilingual meetings in global software projects

CALEFATO et al (2015) point out how in several emerging markets, the use of MT systems can be a quick and cheap solution to the language barrier problem. Specially in countries like Brazil, that are gaining more prominence in the IT market, but still have low number of individuals who can confidently communicate in English.

The author allude to the still less than desired quality of some MT output, but as we previous papers already stated, simple output quality not the greatest predictor of how the communication will take place:

Our findings indicate that state-of-the-art of machine translation technology is already a viable solution for multilingual group communication since it is not disruptive of the conversation flow, it does not prevent group to complete complex tasks, and it even grants discussions that are more balanced. Yet,

machine translation technology currently available is still far from 100% accurate and, as such, its adoption comes with costs. In fact, translations inaccuracies needs to be repaired by rephrasing the original content, thus causing a decrease in efficiency. [CALEFATO et al, 2015]

3.1.6 Ambiente Multiagente para Interações Síncronas

Clemente et al [2015] define the architecture of a system that can aid learners and learning guides during classes, debates, meetings and research sessions in a synchronous textual environment.

The system does so by using a community of agents that pay attention to specific aspects of a synchronous sessions and addressing the problems that come from using this medium.

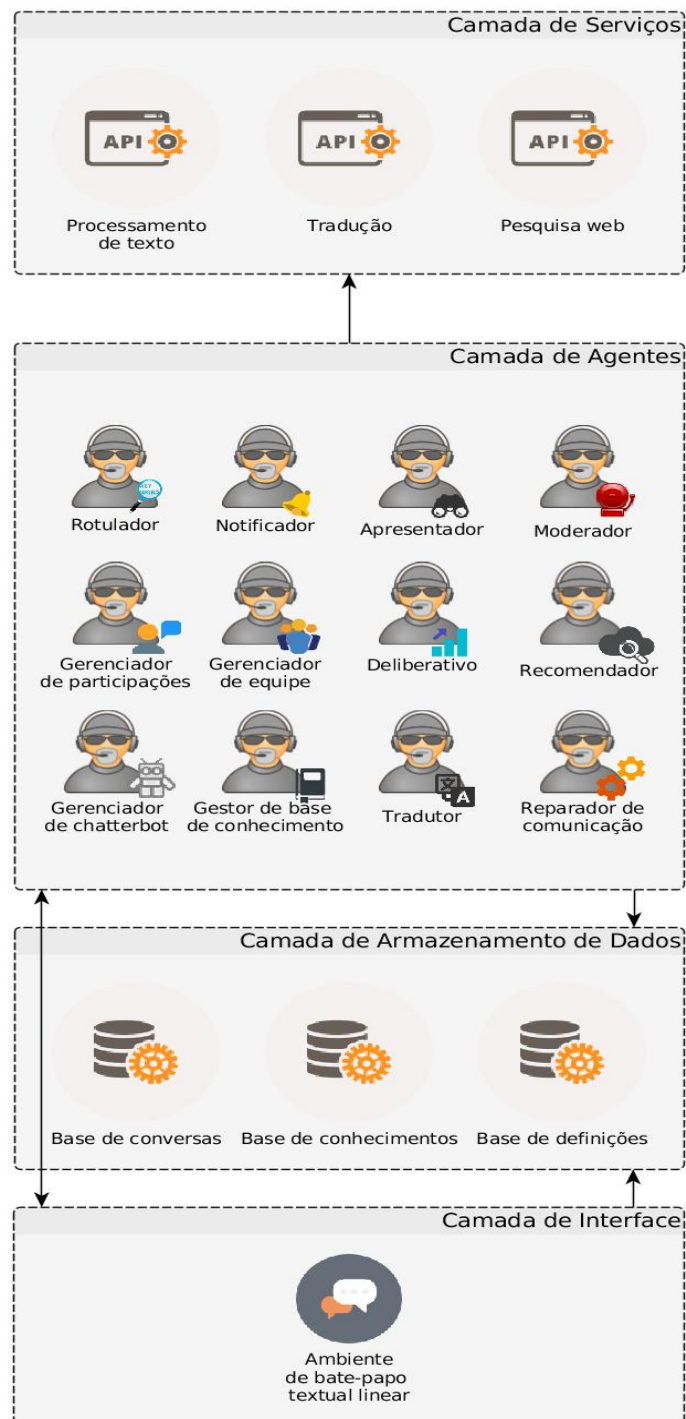


Figure 3.1 - Multilevel Architecture for a Multiagent Environment for Synchronous Interactions (Clemente et al, 2015)

For instance there are agents specialized in moderating a debate, giving each participant of group time and turn to speak; agents that check the main dialog or group dialogues and update the teacher/session-guide about the subjects being

discussed and whether they are not related to the topic of the session.

In general the community of agents can be classified in three groups: (a) those that concern themselves with helping the chat flow by guiding participants and helping the person or people in charge of the session; (b) those that aim to provide a learning equilibration disequilibration process on the basis of Pedagogical Architectures [CARVALHO et. al 2005]; and (c) the agents aid participants by defining, translating and negotiating the communication process, checking doubt and helping solve language related misunderstandings.

The APIs shown in Figure 3.1 support the community of agents according to their functions:

- Text processing: provides services like part of speech tagging, lemmatization, word classification, disambiguation, orthography correction and acronym substitution.
- Translation: machine translation, bilingual and monolingual dictionaries and technical glossaries.
- Web Search: clear websearch, dbpedia and academic paper search.

The agents also shown there are:

- Rotulador: this agent uses Natural Language Processing algorithms to classify each word in the sentences. Other agents use this classification in their decision process.
- Notificador: notifies the session coordinator about several occurrences in the room. For example, it keeps tabs on questions asked and whether they were answered or not, or alerts the coordinator when a user has not interacted for too long a period of time.
- Apresentador: generates data visualization from the chat text, such as concept maps, tree and list structures.
- Gerenciador de Participantes: tracks how participants interact and participate on the chat.
- Tradutor: attempts to enable communication between participants that do not

share a common language. The agents this dissertation describes are represented in this dissertation by the Tradutor agent.

- Moderador: helps planning and managing discussions. Also manages talk-time, using the a token, similar to the “talking stick analogy”.
- Gerenciador de Chatterbots: instantiates activity specific bots, for example a bot keeping track of users doubts in a computer programming class, this bot can take advantage of other agents services. Such as notifying the responsible teacher a group of students has a doubt.
- Gestor de Bases de Conhecimento: stores and structures subjects being discussed, keeping track of participants doubts and certainties as they evolve during the session.
- Recomendador: based on participants questions, it identifies the their interests and, if the users want to know more, suggests papers on the subject.

3.2 Commercial Projects

The idea reaching some kind of “universal communication” method is far from a new one. And several were the methods devised to achieve that, both in fiction and reality. We seem to be moving ever closer to that day, but as the head of the MOSES project puts it: *not in my lifetime*. [KOEHN, 2013]

3.2.1 Not-so-Universal Translators

Still, several crowdfunding projects promise just that, a universal translator. Some of them are mentioned in this chapter.

a. Pilot

This project was part of a crowdfunding campaign that raised US\$4,426,847 in June 25, 2016. It takes form of a wireless Bluetooth headset and claims be a universal translation system. As of the time of this dissertation being submitted, users are still complaining at the campaign page, the company has not yet started delivering the

devices.



Figure 3.2 The Pilot

b. Sigo

Another crowdfunding project that promised a universal translator but has yet to deliver any device. The last project update for the Sigo was about two years ago.



Figure 3.3 - SIGMO Talk & Understand

c. Klik

The Klik also tries to be a hardware speech translation system while also promising “amazing sound quality for music”. Like the two previous projects this one has been successfully funded and has not delivered a single unit.



Figure 3.4 - Klik

And following the same trend of promising perfect speech-to-speech translation I also came across other device named as (d) Travis, (e) SpeechLogger and (f) One2One. The one claim most of them can make is pushing the “wireless earbud” agenda before Apple did. Beyond that they are, mostly, hardware peripherals for ASR and Online MT systems.

3.2.2 Online Machine Translation

MT has largely become natural today, at least when exploring foreign texts in front of a computer, or minimally using a phone to get some vital information when abroad. This has happened because of the high availability of free to use online MT systems. Thanks to systems from companies such as Google, Bing, Yandex and Babylon.

These services, mostly offered for free and sometimes combining TTS and ASR, are the great drivers for the popularization of Machine Translation beyond professional and specialized services.

It is still hard to imagine a world where people consume content other the ones produced in their own languages and the translated version of American or European content. But with the high availability of such systems we can quickly get a good idea of what foreign article or book is saying. Just a couple decades ago that would have to be done with a dictionary and lot of guessing.

4 Brief History of Machine Translation

All languages - at least all the ones under consideration here - were invented and developed by men; and all men, whether Bantu or Greek, Islandic or Peruvian, have essentially the same equipment to bring to bear on this problem. They have vocal organs capable of producing about the same set of sounds (with minor exceptions, such as the glottal click of the African native). Their brains are of the same general order of potential complexity. [WEAVER, 1949]

The excerpt above is taken from a memorandum written by Mathematician Warren Weaver after a letter he wrote to fellow mathematician Norbert Wiener two years prior. Several of his predictions became reality. Specially he predicts the use of statistics as a predictor for the “sense” in which a word is being used, he also talks about using an intermediate form of representation to mediate the translation between a number of language pairs (which is today is known as the Interlingual method for MT).

But the human concern with being able to communicate in foreign lands could be as old as humanity itself.

Since XVII century philosophers such as Leibniz and Descartes devised languages and systems with the intention of creating “universal” forms of representing knowledge, in a logical, ordered and “beautiful” way.

Long before the first computers, the idea of “mechanical dictionaries for multiple languages” already existed. And several of the steps for a simple machine translation system can be done mechanically, or even “manually” [HUTCHINS, 2004]. Some of these systems looked like giant tables where one operator would give another a word, he would move a cursor like apparatus to the word on the table, and the mechanism would point to a fairly equivalent word in six other languages. But before getting to this point, they needed a linguist or someone well versed in the source language, to strip all words to their stems, and the output of the “translation” process would be also just stems, and in the same order as the source language. Then they’d

need someone fluent in the target language to get all stems, conjugate them and put them in the order required in the target language. Of course, that's the same a post-editor those, but imagine a post-editor working with really bad machine translation material to begin with. It might have have easier to ditch the whole machine and do the translation by hand to begin with. Anyways, it was a prototypical system.

One of the great impulses for the development of MT was the Cold War. Americans needed a way to have a peek on what the Russians were publishing, and the Russians likewise[HUTCHINS, 2005] [DORR et al, 1999]. Not only MT, but several other NLP areas saw leaps of development in those days. Much like the Arms Race was replicated in several fields, like Space Explorations, Wireless Communications, Information Storage and etc.

4.1 Important Dates in MT History

The Table 4.1 shows important dates in MT history.

Table 4.1 - Important Dates in MT

1949	Warren Weaver Memorandum
1952	First Machine Translation Symposium held at MIT under Yehoshua Bar-Hillel
1954	First rudimentary MT system developed at IBM
1954	Victor Yngve publishes the first journal on MT, with the name "Mechanical Translation Devoted to the Translation of Languages by the Aid of Machines"
1960	Parsers start to become popular in MT research
1961	Birth of Computational Linguistics. First International Conference on Machine Translation and Applied Language. [include scientists participating?]
1964	Creation of the Automatic Language Processing Advisory Committee (ALPAC) to study the viability of MT systems. The American government is intensely involved.
1966	ALPAC publishes a report pointing out how MT is a waste of money and time. This report greatly impacts the area for several years.

1970	Start of Project REVERSO (Russian Initiative)
1970	Development of SYSTRAN, a Russian-English MT system by researcher Peter Toma
1976	Creation of WEATHER system, translating weather forecasts for the general public. Part of the TAUM, a MT translation group based in the university of Montreal.
1978	Fujitsu creates its ATLAS MT system. A rule based system capable of bidirectional Korean-Japanese translation.
1982	Sharp start offering its English-Japanese MT system under the name DUET. Another rule based system.
1983	NEC develops its solution based on the PIVOT algorithm (which is based on Interlingua Translation). The system is then commercialized with the name "Honyaku Adaptor II"
1986	OKI develops PENSEE, a Japanese-English rule based translator.
1986	The Hitachi Group develops HICATS (Hitachi Computer Aided Translation System), another Japanese-English rule based translator.
1993	Project C-STAR deploys its translation system focused on the tourism industry. Its name stands for Consortium for Speech Translation Advanced Research. The system can translate to and from Japanese, English and German.
1997	Launch of Babbelfish.com (using SYSTRAN technology), kickstarting the popularization of online and free-to-use MT systems.
1998	Softissimo unveils its REVERSO translator system.
2000	Japanese laboratory ATR develops ALPH, an example based translation system for Japanese-English and Chinese-English.
2005	First Web MT system
2006	Corbí-Bellot et al. publish "An open-source shallow-transfer machine translation engine for the romance languages of Spain", which would become the Apertium RBMT system. [CORBI-BELLOT et al 2009]
2007	The METIS-II system is unveiled. It combines elements of Statistical Machine Translation, Example Based Machine Translation and Rule Based Machine Translation (SMT, EBMT and RBMT) Phillip Koehn publishes "Moses: Open Source Toolkit for Statistical Machine Translation", about the MOSES SMT project. [KOEHN, 2007]

2016	Google and SYSTRAN unveil their Neural Machine Translation projects (NMT).

Despite not being “precursors” to Machine Translation and predating the existence of computers in some cases, there is no shortage of attempts of creating a universal language or a universal translation machine. The interest in enabling communication across languages and peoples of the world was behind countless ideas, essays, and letters. Some of which were a 1629 letter in which Descartes rebukes the ideas of a “Universal Language” [1629], John Wilkins’s “An Essay Towards a Real Character and a Philosophical Language” [1668] which is more interesting as a birth to the area of taxonomy than as a treaty on language design. And, to briefly mention designed languages, there are Zamenhof’s Esperanto [BOULTON, 1960] developed to be regular and easy to learn; the International Auxiliary Language Association’s Interlingua [FALK, 1995], designed to be easy for speakers of the largest European languages to pick up; Toki Pona [LANG, 2014], with it’s total of 120 words, created as a challenge to come up with the simplest language possible; and Blissymbolics [BLISS, 1949] created aspiring to become a universal logographic language, but finding home in schools and hospitals helping aphasic children and adults to enlarge their communication repertoires. That just to mention a few.

4.2 State of the Art

There have been new developments in MT, namely the use of Recurrent Neural Networks, which Google first intended to hard language pair such as Chinese-English, but that now seem to be slowly creeping into other language pairs. As [YU et al. 2016] points out, in Google’s NMT experiments, the outputs are generally have better scores than the ones resulting from SMT.

SYSTRAN, the *de facto* standard for commercial MT is also investing in Neural Machine Translation (NMT). Albeit Jean Senellart, SYSTRAN’s CEO, says they plan on maintaining the other, more time proven, options available to their clients.

[VASHEE, 2016]

We cannot precise which of the players in the MT market are investing in NMT. But since the giants Google and SYSTRAN are, it is safe to assume the industry will follow in this direction, at least until something else is developed. Due to the nature of proprietary and closed source software it is possible that similar solutions involving NMT are already in use.

As for the Open Source community, Phrase Based Statistical Machine Translation (PBSMT) still seems the most reliable and widely spread paradigm. MOSES [KOEHN, 2007] an open toolkit for SMT has been since its release on the forefront of MT development. The system has achieved such maturity and robustness that several commercial applications and research projects rely on its infrastructure.

Another highlight from the Open Source community is Apertium, a Rule Based MT system (RBMT), which can achieve reasonable results for related languages (such as the languages of Spain, Portugal, France and Italy) with much higher speeds and lower processing cost than SMTs and NMTs.

In this section we give a short explanation of the specificities of each of the main MT paradigms.

4.3 Machine Translation Paradigms

There is a lot of debate over where the borders of certain paradigms are, specially today since all well established MT systems borrow ideas from a wide range of different paradigms. But one characteristic can be used to separate them: whether they rely linguistic knowledge or extract their rules from data (corpora, examples, etc).

The paradigms that follow are the ones that require expert linguistic knowledge as a prerequisite to setting up the translation system.

4.3.1 Rule Based Machine Translation - RBMT

Involves breaking down target and source language into sets of rules for grammar, semantics, morphology and syntax, and mapping how these structures resemble or differ from one language to another.

This approach has the advantages of allowing intervention by changing the rules, and since it does not rely on the processing of large corpora, RBMT systems are extremely fast. As a downside, rules for translation are not created automatically, it requires a lot specialized knowledge and time to create a working RBMT system.

4.3.2 Knowledge Based Machine Translation - KBMT

Systems like this one rely on Knowledge Bases for a specific domain (usually Ontologies, Wordnets), and are able to infer certain information about the world after which they are modeled. According to DORR et al (1999) "the KBMT paradigm concentrates on the development of knowledge intensive morphological, syntactic, and semantic information for the lexicon." DORR also states that this paradigm "provides high-quality, fully-automated translations", but since creating general KBMT systems would be too costly and time consuming, they are used in restricted domains. And are commonly employed as complements to LBMT and PBMT systems.

4.3.3 Lexical Based Machine Translation - LBMT

Definitions may vary, but LBMT are systems where rules are set that relate and map lexical elements from one language to another. For example:

- I'm 29 years old.
- Yo tengo 29 años de edad.

The system must account to the fact that, to express "age/how old someone or something is" the verb "be" is used in English, but that in Spanish the verb "tener" is used to the same effect.

DORR et al (1999) states that this approach “requires entire trees to be stored in the transfer dictionary for each source-to-target pair, which is significantly burdensome as the number of source and target languages begin to increase.

4.3.4 Constraint Based Machine Translation - CBMT

Uses sets of rules to map grammatical and lexical relationships among language pairs, and uses those as a base to generate translations. [DORR et al, 1999]

As opposing to the more linguistic knowledge based, MT systems called “Empiric” are the ones that extract the rules and inferences from data, building themselves the necessary insights to make the translation.

4.3.5 Statistical Machine Translation - SMT/PBMT

This method relies purely on the processing of parallel corpora to extract probabilistic rules that predict how likely it is for an (s) set of words in the source language to be represented as a (t) set of words in the target language. This particular paradigm is requires massive amounts of parallel text already translated and proven to be acceptable.

4.3.6 Example Based Machine Translation - EBMt

This method uses Case Based Reasoning, extracting from the parallel corpora sentences and cases that resemble the source text they are trying to translate. EBMt systems translate by analogy, which means they use similar sentences from the corpora to create new translations.

4.3.7 Neural Machine Translation - NMT

Recurrent Neural Networks seem to be changing the computing landscape. Being such a versatile tool, it is not surprising they are now being incorporated into MT. Conventional Phrase Based Statistical MT uses parallel corpora to learn which are

the most likely words to be used as an equivalent from source to target language. That is from segments of one to several n-grams. Thus a trained system can extract the probability of a word in the source language being represented by another word in the target language, by a purely statistical process (note that no mature system relies purely on a single paradigm, therefore most all MT systems used for real world applications are Hybrid MT systems).

What NMT does differently is having an RNN train in parallel corpora, and "learn" by itself which segments of text would be equivalent to other segments. Exactly how these artificial neuron networks learn to generate near human level translations, it is still not entirely clear.

5 Communication Strategies

In this section we explore some of the common and not so common ways in which people try to make sure they are being understood by their conversation partners. It's important to note that not all these strategies work on a textual environment, and some translate better than others.

5.1 Chunking

In several situations we tend to pour everything we have to say over the interlocutor, rarely stopping to check whether they have understood what we are trying to convey. This strategy consists in approaching the conversation as a step by step process, in which the speaker presents concepts in succession, giving opportunity for the listener to request clarification without having to interrupt the speaker.

5.2 Goal Oriented Conversation

Defining the goal at the beginning of the exchange, and making sure all participants have a clear understanding of its objectives is a way to assess how the interaction went. This framework helps all participants define the questions they must be able to answer after an exchange session and whether it was successful or not.

5.3 Explaining the Main Points

Much like having a goal defined before the dialogue starts, not losing time on unimportant details can help the conversation move forward at a pace that keeps participants engaged and prevents them from burning out and giving up on the exchange.

5.4 Talking Back

Asking participants to explain the what was discussed back to the speaker and to all other participants. Even though this strategy is not feasible to be practiced with every single member of a large group, the speaker can select small sets of participants and ask them to explain the discussed point back to them.

5.5 Avoiding Certain Words, Expressions, Constructions

There are several expressions and constructions we use daily that are not standard, they help us identify with a group or are simple habits we pick from other speakers. When talking to someone who does not have the domain of the language, trying stick to the most direct and simple variant of the language is good idea. The meaning of expressions is usually not self evident and their obscurity might derail the understanding in a conversation with people from other linguistic background.

5.6 Gesturing

Even though this strategy comes naturally, pointing at things and making gestures can help disambiguate certain meanings, but all the participants must be aware of culture specific gestures, since, just like words, some gestures assume different meanings in different communities and countries.

5.7 Use of Images

There are concepts that are easier to explain with images, which has become much easier now with the rise of mobile computing (phones, tablets, gaming devices and all sort of portable computers) and with availability of search engines with massive databases in all the languages that have records on the world wide web, although internet access is bound to be challenge depending on the setting this attempted conversation might be taking place.

5.8 Rephrasing

A strategy that is naturally used by monolinguals when they do not understand the language being used, for example when people who are not law students talk to attorneys about law dealing and have to constantly ask them “what do you mean by that?”, “what do you mean by this unnecessary latin word you just used to linguistically coerce intimidate me?”.

Unsurprisingly, bilinguals use the same strategy. Since there will always be words we have not yet listened to or whose meaning we do not fully understand, this is one of the most organic and commonly used strategy to preserve the understanding in a conversation.

Such strategies are effective in a wide range of contexts, but they can make a lot of difference when you are communicating through a textual medium while having your message processed by a device as opaque as an MT system.

6 Architecture

When dealing with pairs of groups of people who are trying to communicate without having a language in common, the agent relies heavily on MT. But there are cases in which this use is not necessary. For example in the case where a lawyer and client are talking about a case, the agent might take the burden away from the lawyer by explaining technical language the client might not know. And the same is true for any situation where a person or group of people must understand someone who has to rely on jargon or local expressions to express their thoughts.

And, beyond that, the agent also embodies strategies that are commonly used by bilinguals when having a conversation in a language they have not fully mastered. This strategy modes are initiated when the agent finds specific expressions in the conversation. Such expressions include those that show problems in the communication.

6.1 Addressing Use Case Scenarios

In this section we address the use case scenarios proposed in the first section of this dissertation and also hint at some ideas to be better discussed later into this dissertation.

6.1.1 Scenario C1

In this scenario two people are trying to have a conversation, or exchange some information, but they do not know how to speak the other's language and must rely entirely on the helper agent to communicate.

Since we are dealing with only two people, a single agent can be used, and all the mediation can be done on the main conversation thread. All word definitions and discussions can also be done on the main conversation thread.

6.1.2 Scenario C2

In this Scenario C2, in addition to the two participants and the agent, we have a third person who has some knowledge about both languages and who can intervene in case some misunderstanding happens. In this case the agent act proactively, by translating every sentence and freeing up the human helper to only interfere when something goes wrong.

6.1.3 Scenario C3

Among the proposed scenarios this is the hardest to address. More users and more languages increase the chance of mistakes happening, also the conversation flow becomes a challenge, because some users might need to wait why the others clarify their doubts about the what was said.

Several instances of the Agent can address each language group, for instance, if the person in position of a lecturer only speaks Portuguese, and among the students are Portuguese, Spanish and English speakers. Showing all translations in the same thread would make it impossible for anyone to follow the conversation without getting lost. So having an instance of the Agent helping each

6.1.4 Other Scenarios

During the course of the research several other ways to use the agent became apparent. Among which using it for language learners, targeting only the words and expressions they have not yet studied. The agent could also be useful in helping group of people explore a text that was not in their language, for example Brazilian graduate students collaboratively exploring papers in English, with the help of the agents here used to help chat participants communicate. since these ideas could be the subject of entire new research projects, they were included in session 8.3, Future Research.

6.2 Agent Strategies

In this section are presented the strategies that can be used to maximize the probability that the participants understand the messages of each other.

6.2.1 Machine Translation

In the current state of development of Statistical Machine Translation Systems (SMTs), they are able to efficiently translate most common phrases and expressions from Portuguese to English and vice versa. The same can be said for the world's most commonly spoken languages. Thus, the use of translations from SMTs is usually enough to mediate interactions in common scenarios such as greetings and basic information exchange.

6.2.2 Use of more than one MT output

Research shows that using outputs from different MT systems, that have different internal configurations and processes is more efficient than using just one of these systems. Participants are able to more accurately infer the meaning of the original message when they contrast the two translations. There is also the chance that one of the translations will be better suited to the context of the dialogue than the other.

6.2.3 Paraphrase / Paraphrasing

One of the tactics in case of a breakdown of communication between users is that the user who got lost in the conversation asks the previous user to explain what he said before with different words. Giving one more opportunity for the MT systems to generate a different output and giving more information to the recipient user of the message, so they can grasp its original meaning.

6.2.4 Dictionaries / Thesauri / Glossaries

The use of bilingual, monolingual and synonym dictionaries is another method by

which we can clarify important terms for communication, especially when these terms are regional, area specific or of little current use, and are therefore much harder for MT system to correctly translate.

In this case the use of technical dictionaries can also improve the level of understanding. Users can define which subject areas they are likely to be covering during the session, choosing features related to the subject, and specialized glossaries can be used to aid in the interaction.

Users dialogues can later be mined to enrich glossaries with new entries and definitions.

6.2.5 Wikidata / DBpedia / OpenWordnet-EN / Princeton Wordnet

They can be complementary the dictionaries and thesauri functions. Being chosen as a primary source by the user or working in redundancy to these systems, assuming the function of the previous resources if they do cannot retrieve entries related to the expression being used.

6.2.6 Collaborative Repair

In scenarios where more than two people are talking, we can rely on group consensus to advance the interaction. Even if users know little of the partner's language, the group's accumulated knowledge has the potential to produce more balanced results than in a one-to-one interaction. That strategy comes with the problem of managing user discussions, which becomes harder the more users and languages we have in the group.

6.2.7 Images

There are cases where there is simply no clear linguistic equivalent for something that is common in another language, for some of these cases an image can help tip

the interaction to the right direction.

For example, there is no "cozinha" nor "enroladinho" in the English language, just as there is no direct translation, an image can solve this type of scenario faster and in a simpler way than text can. (Example on the Figure 6.1)

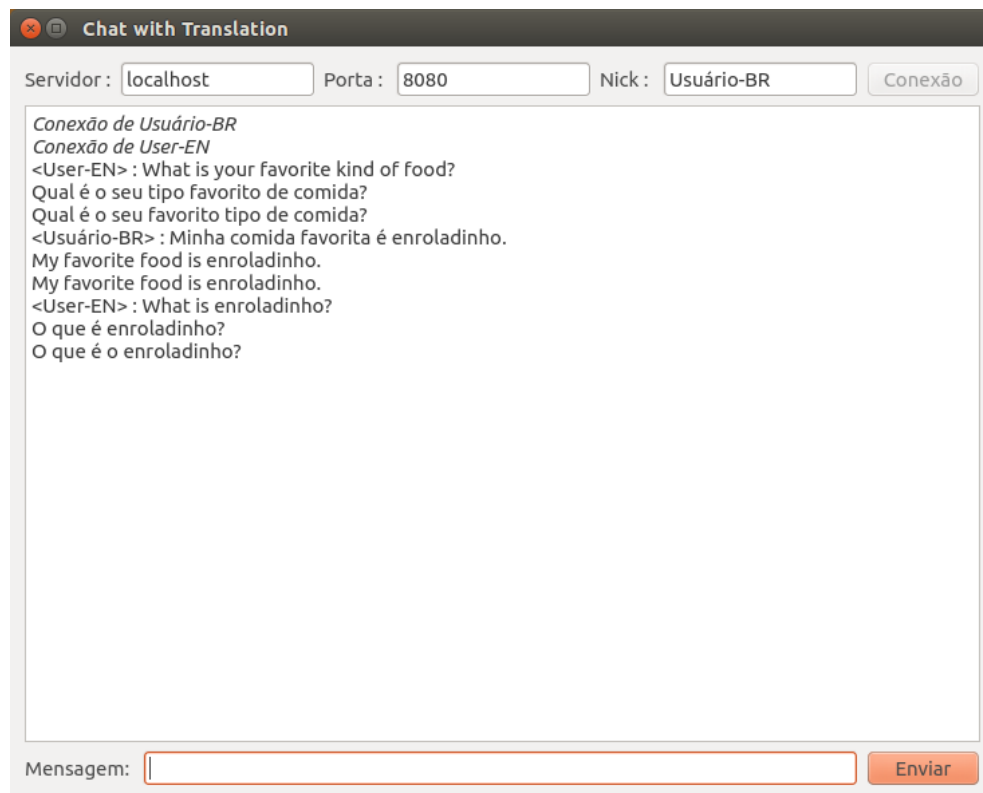


Figure 6.1 - Chat with translation - example of use

It's much faster to show the image of the food and complement that with a description than try to describe it, since food is so culture specific.

6.3 Multiagent Architecture to Support Multilingual Interaction

This section introduces a Multiagent Architecture to Support Multilingual Interaction.

The Figure 6.2 roughly shows how the work is divided. *Lang_Agent* here denotes the subsystem responsible for fetching the translations, definitions, synonyms and every

other linguistic resources needed during the interaction. *Speaker_A* and *Speaker_B* are the participants trying to have a conversation, let's assume the first speaks Portuguese and the second speaks English (as long as there are MT systems and other resources available, any two pairs of languages could be used). The *Comm_Agent* is the part of the system users interact with, either by directly asking the agent for translation, a word definition, a synonym, or indirectly by saying something or behaving in a way that prompts the agent to take part in the conversation. All the interactions are recorded with the intention of being used in future research.

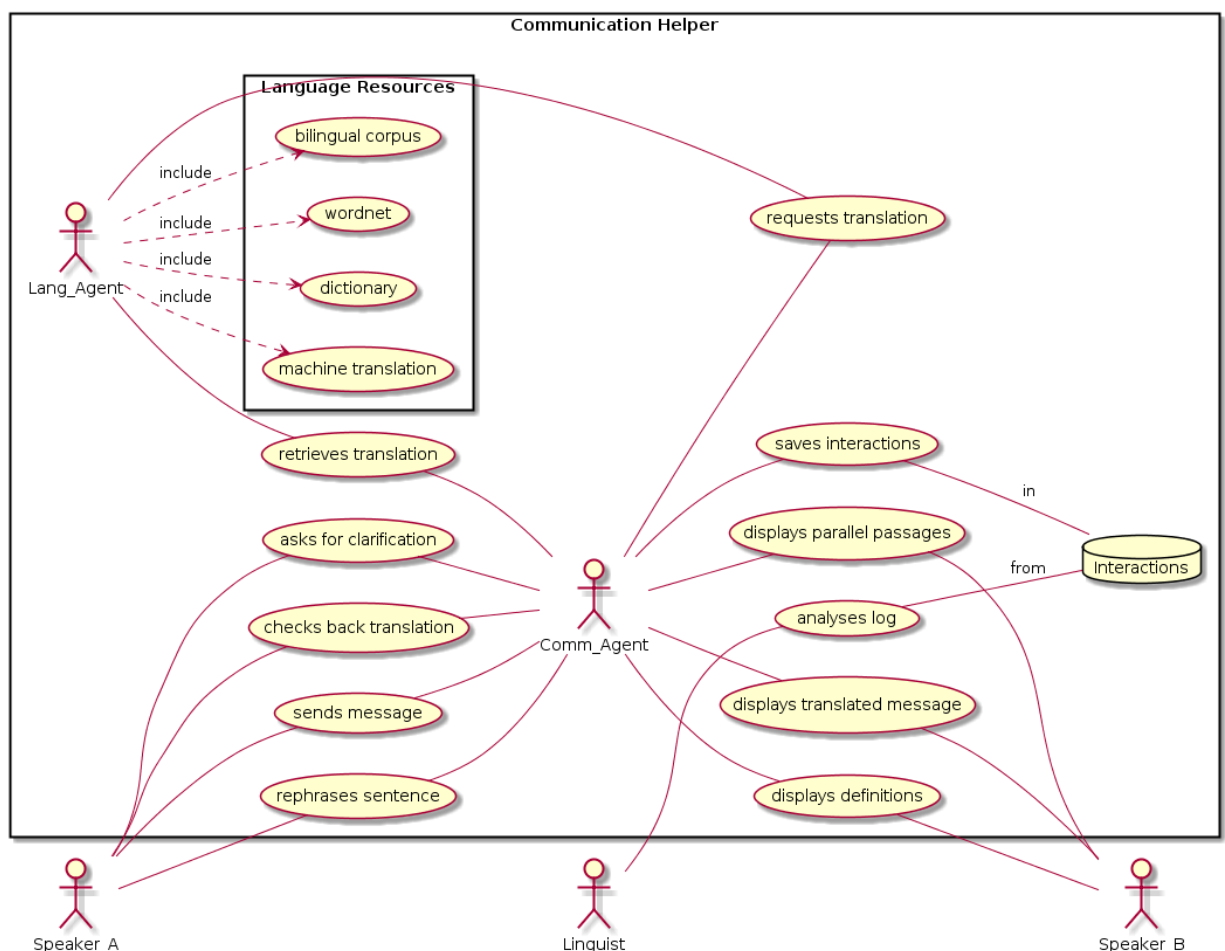


Figure 6.2 - Multiagent Architecture to Support Multilingual Interaction

The *Lang_Agent* and *Comm_Agent* collaborate to make the communication between the users possible.

Lang_Agent handles all the linguistic resources on which the strategies described in section 6.2 depend. Services such as monolingual and bilingual dictionaries, thesauri, bilingual corpora and several MT services. Giving the users and agents a way to explore the communication beyond what they are capable of without help.

The *Comm_Agent* watches the conversation, displays the translations, and monitors the users of any request of help. Such requests come as sentences like “I don’t understand” and “what do you mean by”. It also requests the user to express whether they feel they are understanding the conversation or not, and can prompt them to define what the goals of the conversation are, to break down these goals and follows the interaction are they are reached one by one.

Another function this agent has is the recording of the conversation. The use of digital resources for synchronous with additional difficulty of translation and communication repair is something the still needs to be better explored. Even though several strategies can help maximize the chance of understanding, it is hard to predict what kind of misunderstanding and problems may arise from this process. So creating a corpus of agent mediated conversation is a way of giving linguists, interpreters, translators and information scientists an extra tool to use in understanding how system like these affect our understanding, and even more importantly, where they should and should not be relied on.

For a better understanding of Multiagent Architecture to Support Multilingual Interaction, shown in Figure 6.3, all elements of this figure are described below:

1. Users:

1.1 Speaker-A: A person who is a native speaker of Portuguese and who does not speak English, this person might even have elementary understanding of the partners languages, but is not confident enough to be sure they are saying what they mean to say.

1.2 Speaker-B: Like speaker A, but inverted. If speaker A is a native Portuguese

speaker and does not speak English fluently, speaker B is a native speaker of English and has little, or even no knowledge of Portuguese vocabulary and grammar. Portuguese and English have only been chosen because one is the official language in Brazil (where this dissertation was written) and the other is seen as a global language. But any other language pair could be used in here, as long as the needed resources are available for the chosen pair.

1.3 Linguist: the strategies and resources used here aim to make communication easier when no common language is available for people to have a conversation. But language is a subtle and delicate thing, and there's no easy way to translate a sentence, as there are several potentially correct ways, but all of them are bound to leave some meaning or connotation behind. Therefore, collecting these agent-mediated dialogs in a corpus is a way recording the cases in which the Agent makes conversation possible, but also those cases in which the strategies and resources used are not enough to avoid misunderstandings. With this corpus, language researchers can study how this technological apparatus affects language, then propose improvements or new ways to aid in this mode of communication.

2. Synthetic Agents:

2.1 Lang-agent: the Language Agent manages all the tools used by the Communication Agent (Comm_Agent) to mediate the dialog. The most used tools are MT output from several translation services, bilingual and monolingual dictionaries, wordnets, translation corpora, usage examples for specific words or expressions, glossaries for conversation specific needs (that can related to technical areas, slang, regional expressions, etc).

2.2 Comm-Agent: this is the agent that presents the users with definitions, translations and sentence examples. He also keeps track of users roles in the conversation, doubts, questions, and can intervene in the conversation when a user asks for clarification. The strategies the this agent uses are modeled on strategies people use while trying to communicate in a normal conversation. For instance, asking a speaker to rephrase what he said, or to define goals for a conversation, and

later check whether these goals have been achieved. And finally this agent takes commands from users, depending on their roles in the session. Commands can signal the agent to be silent, ignore (stop translating) a specific user or language, change the translation service being used, fetch a definition or usage example, show word classification, or fetch information from other services.

3. Functions:

3.1 Retrieves translation: when users have practically zero domain of the other participants languages, the Comm_Agent translates all the messages, sentence by sentences, requesting the Lang_Agent for a translation. Which can be done from a single or multiple sources.

3.2 Asks for clarification: when a user signals he is not following the conversation, the agent requests the other speaker to explain rephrase, to repeat what he said with other words, giving the confused participant a chance another chance to try to understand what is being said, which the agent checks by asking the participants. The normal flow of the conversation is resumed either when the participants signal they understood the message or when the agent runs out of clarification strategies.

3.3 Checks back translation: When understanding is poor, and participants seems not to be able to communicate with simple or double translations, the Lang_Agent can request the sentences to be translated to the target user language and then back again to the source users language (a process which is known as back translation). This gives the speaking user a chance to check which parts of his message are more likely to be misunderstood, and address that.

3.4 Sends message: in a normal conversation flow, users are likely to only send their messages. It is only when understanding problems arise that they are likely to need to interact with the agent more directly.

3.5 Rephrases sentence: when problems appear, the Comm_Agent notifies the user that his conversation partner has not understood the last message, and suggests that

the user repeats what they were saying with a different wording. It is important to note that the agent cannot force users to act accordingly, but we count on them to do so if they are genuinely interested in making their message get through.

3.6 Requests translation: the Comm_Agent requests all translations and definitions to the Lang_Agent, which defines a common interface for all services used. A user can request the Comm_Agent for another translation if he suspects the last one displayed was not good enough, the Comm_Agent in its turn will pass the request to the Lang_Agent that will select and a new service, retrieve the translation and return it to the requesting agent.

3.7 Saves interactions: while interacting with the participants, the agent instances also store the logs for the conversation for future use. We expect to learn from these logs, by understanding how the system interferes with communication, and how users adapt to it or fail to adapt to it, and how they make use of the tools available. The information learned from this logs can be used to guide the evolution of the system, how agents should be adapted, which tools should be used more or even be abandoned.

3.8 Displays parallel passages: when being requested to search for a word or expression in parallel corpora services the Comm_Agent uses the Lang_Agent to retrieve the passage containing the desired expression and shows it to the users as a message, with added explanations.

3.9 Analyses logs: this Linguist is intended to symbolize researchers that can use the logs to gain insights and point new strategies that can be used to minimize communication problems. But beyond that, mining agents can process the logs to extract valuable statistics, word definitions, expressions not yet recorded. These can be incorporated to the glossaries and dictionaries used by the Lang_Agent.

3.10 Displays translated message: the Comm_Agent show the participants messages translated by one or two of the MT systems that interface with the Lang_Agent.

3.11 Displays definitions: the Comm_Agent also displays word definitions are requested by the users.

4. Resources:

4.1 Bilingual corpus: provides real usage examples of sentences in both the target and source languages. This resource forces the user to read passages of text from documents translated by human translators. It can be time consuming to for users to this, but if done to key concepts for the conversation, can guide the dialogue to better results. Specially when the machine translation fails to render these key concepts right (which is more likely to happen when the concept being discussed is not a common expression).

4.2 Wordnet: can be used as alternatives for dictionaries, specially if the relationship between words is an important feature for the conversation. Since these resources encode relationships of hierarchy, size, feelings and other features that are related to the entries.

4.3 Dictionary: for general communication purposes, monolingual dictionaries can be used to display definitions and have these translated to the other language(s) being used in the conversation. Bilingual dictionaries can do the same without the need of MT. Other purpose specific dictionaries can be used depending on the goals and topic of the conversation, for example: Etymological dictionaries and technical glossaries, depending on the area being discussed.

4.4 Machine translation: one of the key elements of to enable the dialog when participants do not speak the same language. Even though this resource is also the where most mistakes and problems appear, it makes conversation much faster for simple messages. Translation quality also varies a lot depending on service and language pair being used.

4.5 Interactions: in this scheme denotes the logs, the database where all interactions

among users and agents are recorded for later processing and study.

6.4 Agent State Model

The state model guides how agent instances react to users. They contain the users name or id (nickname in case of the prototype), user classification, basic numbers about the state of the conversation and current mode (Figure 6.3).

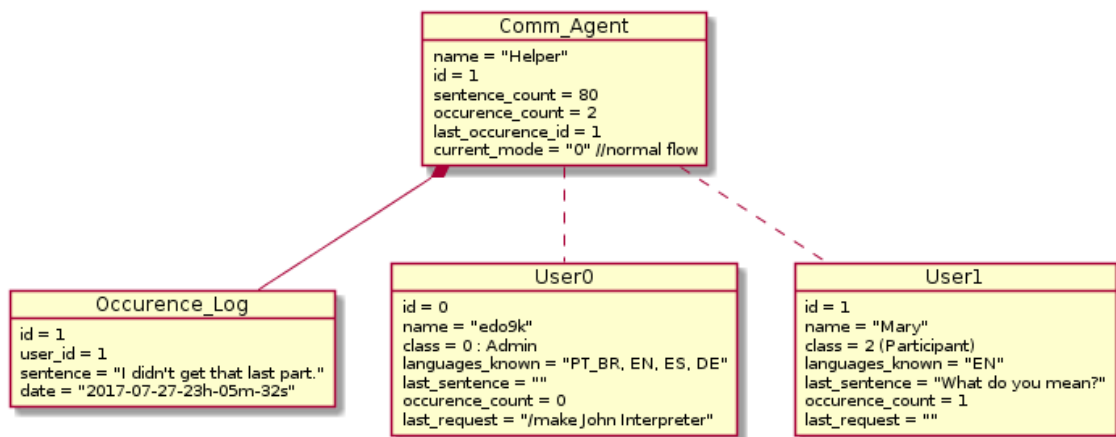


Figure 6.3 - Agent State Model

Users are classified as:

- Participant: this is the default category when the agent is not supplied any extra information about the user. The agent then tries to define which languages the user speaks based on the ones they use on the conversation, adding these to the list of languages not to translate to this specific user.
- Guide: usually a teacher, or lecturer, this participant is supposed to talk more than other users, being the one who is exposing some concept or idea.
- Interpreter: a participant that speaks both languages being used, and therefore can check the quality of the translation delivered by the Comm_Agent, spot misunderstandings, direct the agent and help users clarify their doubts. The agent, by default, does not translate what an interpreter says, letting them explain themselves.

- Admin: any user tagged as administrator will have control over the agent instances, being able to change settings, reset behaviors, and modify user classifications.

The agent stores some basic number like, how many messages have been sent, how many times users requested clarification, rephrasing, alternatives translations, usage examples, synonyms, corpora entries or signaled they were not understanding something. Users information also include a time tag indicating the last moment they sent a message or requested any of the services above.

The agents uses this information to request users to direct their attention to specific points of the conversation. Being able to show user all signaled misunderstandings, and letting them analyze them with the benefit of hindsight.

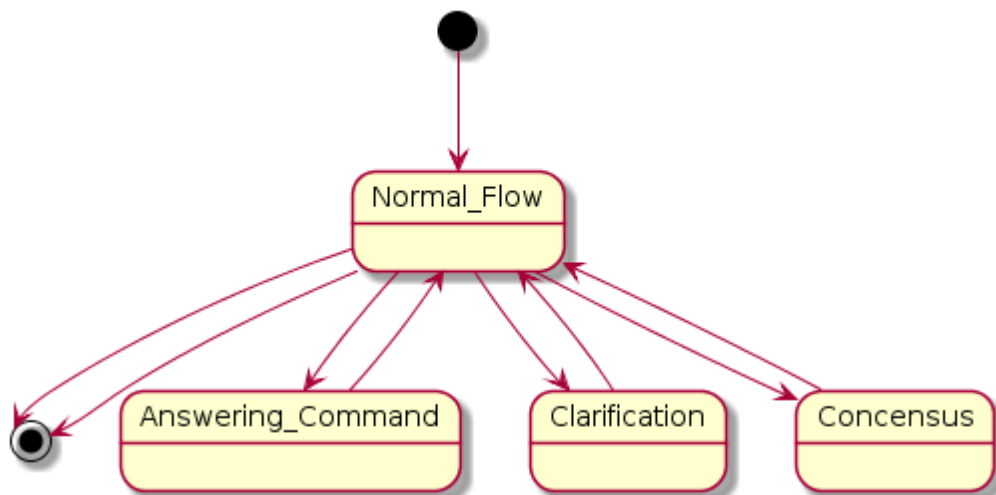


Figure 6.4 - Agent Action States

Figure 6.4 shows the communication agent states, which are detailed below. These are tied to specific moments in the conversation:

- Normal flow: simplest behavior, the agent can present users with translations for every sentence, to and from the language pairs it was set to use. Or, when not working specifically with translation, present only the definitions to the words found in the glossaries. The agent still responds to direct requests from the users when in this mode.

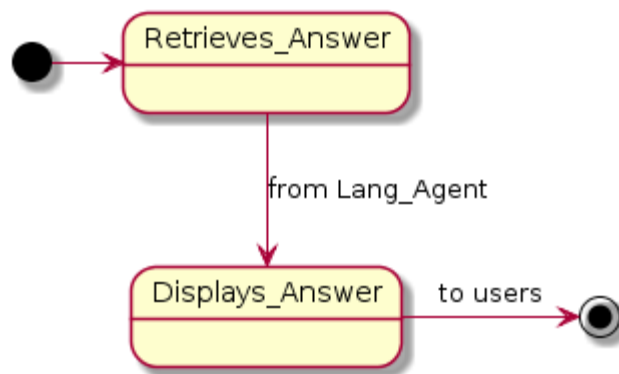


Figure 6.5 - Answering Users' Commands

- Answering commands (Figure 6.5): when a participant requests something directly to the agent, the agent will be in this state until the request is answered. Further requests received in this state will be queued and answered in arrival order, if on the main chat. Multiple requests can still be answered to privately. (IRC, the chat service used for the prototype, provides a way to send messages to specific users without taking them away from the main chat environment, if this system is available, users can request definitions and other services to the agent, and these can be delivered straight to them, without polluting the main chat flow — which is bound to happen in a multi-participant session.)

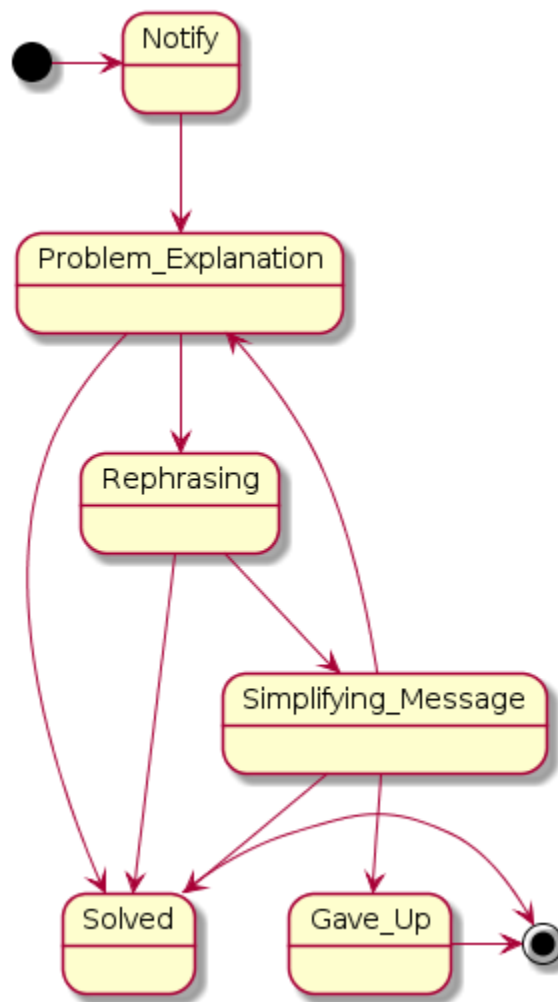


Figure 6.6 - Clarification Attempt Steps

- Clarification, Problem resolution (6.6): when a participant signals he has not understood some of the previous messages, the agent starts a back and forth process that puts the originator of the message in touch with the person who had problems understanding it, so they can solve the doubts that arose. This involves (a) asking the participants what part of the sentence was a problem, (b) asking them to express the same idea with different words, (c) asking them to simplify their message to their essentials and use dictionaries or other non-MT services to go through each of them.

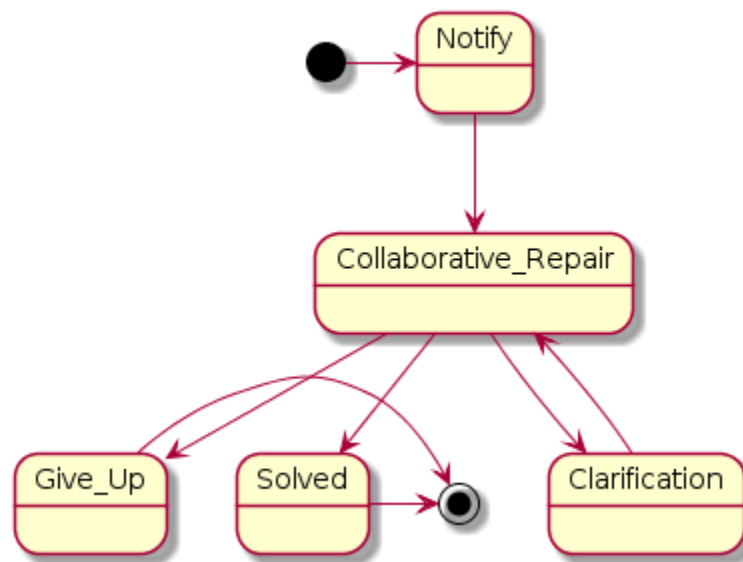


Figure 6.7 - Consensus, Collaborative Steps

- Consensus (Figure 6.7), akin to Clarification, this mode prompts the agent to notify the user who sent the message that it was not clearly understood. But, being aimed at group discussion, the users from the target language can then debate and use their collective knowledge to try to understand what was said, and ask the participant who sent the message for extra information while trying to make sense of the translated version they had trouble grasping. While in this process they can still use the strategies of Clarification to explore the problem message.
- Passive: the agent only reacts when being directly requested by the users.

On clarification, we can say the source participant is the person who sent the message that was not understood (SP from here on). The target participant is the person who received the message with translation, had some problems in understanding it and signaled it (TP from here on). The first step is notifying the SP there were some doubts in raised by his previous message, confirm with the TP which message he could not understand, show the original message back to the SP accompanied by a back translation from the message the TP had problems understanding. Then asking the TP what he did not understand and showing that to the SP, if the SP can explain that part of the sentence and the TP signals he now

understood, the clarification process is over. If the TP signals they have not yet understood, we ask the SP to rephrase the sentence, in more than one step if necessary. If the TP signals understanding, the process is over, if not, we ask the SP to simplify his message to the bare essentials, and walk the TP through each of those words as retrieved from a parallel corpus of target and source languages. This is the last resort, if the TP still signals they cannot understand the message, the agent tells the participants this occurrence is going to be stored for later review, and shows it to them again later, when they end their conversation. Of course, if they cannot go on without solving this problem, the agent can give them the chance to try again, or simply give up. If they can go on chatting, the agent will remind them of this situation by the end of their conversation, hoping that with new information acquired since this occurrence they can now understand each other better.

One resource that was thought on was having some voluntary or even paid translators on notice, and only ask them for help after exhausting the agent strategies. Or even having understanding problems sent to them by e-mail or put on a platform where speakers of both languages could help clarifying these communication problems, and this be sent to the participants later. These ideas were not pursued in this dissertation. They might be the source of future research.

```
< Speaker_A > In this case you are going to need a visa.
< Helper   > Speaker_A disse: "Neste caso, você vai precisar de um visto."
< Speaker_B > Que tipo de visto?
< Helper   > Speaker_B said: What kind of visa?
< Speaker_A > Tourist Visa. How fast can you get it?
< Helper   > Speaker_A disse: Visto de turista. Quanto rápido você pode obtê-lo?
< Speaker_B > Só depois do festival Prato da Casa.
< Helper   > Speaker_B said: Only after the festival Prato da Casa.
< Speaker_A > I didn't understand.
< Helper   > Ok, let's try to solve it.
< Helper   > Speaker_A não entendeu sua última mensagem, repita com outras palavras por favor.
< Speaker_B > Não consigo o Visto antes do festival de música aqui.
< Helper   > Speaker_B said: I can not get the Visa before the music festival here.
< Helper   > Speaker_A, is the message clearer now? (Yes/no)
< Speaker_A > Yes.
< Helper   > Ok problem solved, please go on. / Problema resolvido, por favor continuem.
```

Figure 6.8 - Rephrasing for clarification

Consensus mode tries to use the similar strategies from the previous one, but lets users collaboratively repair the message broken by the MT, while the participant who

sent the message follows this process trying to answer questions or help with extra information.

Participants can define milestones or points to discuss, and have the agent keep track of them. This way being able to guide a conversation, meeting, class or discussion, point by point, only moving to the next when signally the current one has been dealt with.

7 Prototype Implementation

The prototype was built using python. There is an effort for the use of this language at LIEd, aiming to ease the interoperability among the several projects under development. I did not have any contact with the language prior to the beginning of the research, but the language seems to be designed for simplicity, and except for minor peculiarities it did not hinder the development of the prototype.

There was a requirement of a synchronous chat system and, because of the familiarity with the IRC protocol and its openness, the large number of servers, clients for every major platform and open APIs, this service was chosen. But only as an example, because the ideas are applicable to any textual communication system. Google Hangouts, Facebook Messages, Telegram, any service that has an open API can use the same strategies (or even the same codebase). There was also the intention of making a Whatsapp version, since this service is the most popular in Brazil as the time of the writing of this dissertation. But Whatsapp does not offer an open API.

An obvious choice for the MT System was using Google Translator, but several other Online MT systems were screened. Namely Yandex, Bing and Babylon. For a better control we also studied using MOSES and Apertium, two MT systems that can be deployed not only on the cloud but also as local services. This was not done in the prototype, but could make the system much more robust, allowing offline use, and the retraining of the system based on the data acquired.

The first plans also included adding AVR (Automatic Voice Recognition) and TTS (Text-to-Speech Synthesis). Which can both be implemented as additional modules to the system, but were not addressed in this prototype due to restrictions to the Google Voice API, and the fact that all the LVCSR (Large Vocabulary Continuous Speech Recognition Systems) were not only proprietary but also not designed to work under Linux. Microsoft, Apple and Nuance have robust and flexible AVR systems that fit the requirements but, as mentioned before, neither are free to use

nor open.

Open SPHINX is an Open Source solution for AVR, but after several attempts, we could not get it work. Therefore we decided to leave this aspect out of the research and focus solely on the textual synchronous communication with the use of MT, dictionary and dictionary-like services and a helper agent.

The first versions of the prototype were developed in Python 2.7, the oldest but still supported version, after the third iteration of development the code was ported to python 3.5.2, the latest version at the time of the writing.

In addition to the hooks that connect the agent and services to IRC, a desktop chat application was developed, as way of exploring PyQt5, a technology that I was personally interested in, since I had worked almost exclusively with web technologies in the past. This application was developed on a client-server architecture. Whereas the IRC reliant version can either be run locally or on a server.

7.1 Examples of use of the prototype

This section presents some examples of use of the prototype in different usage scenarios.

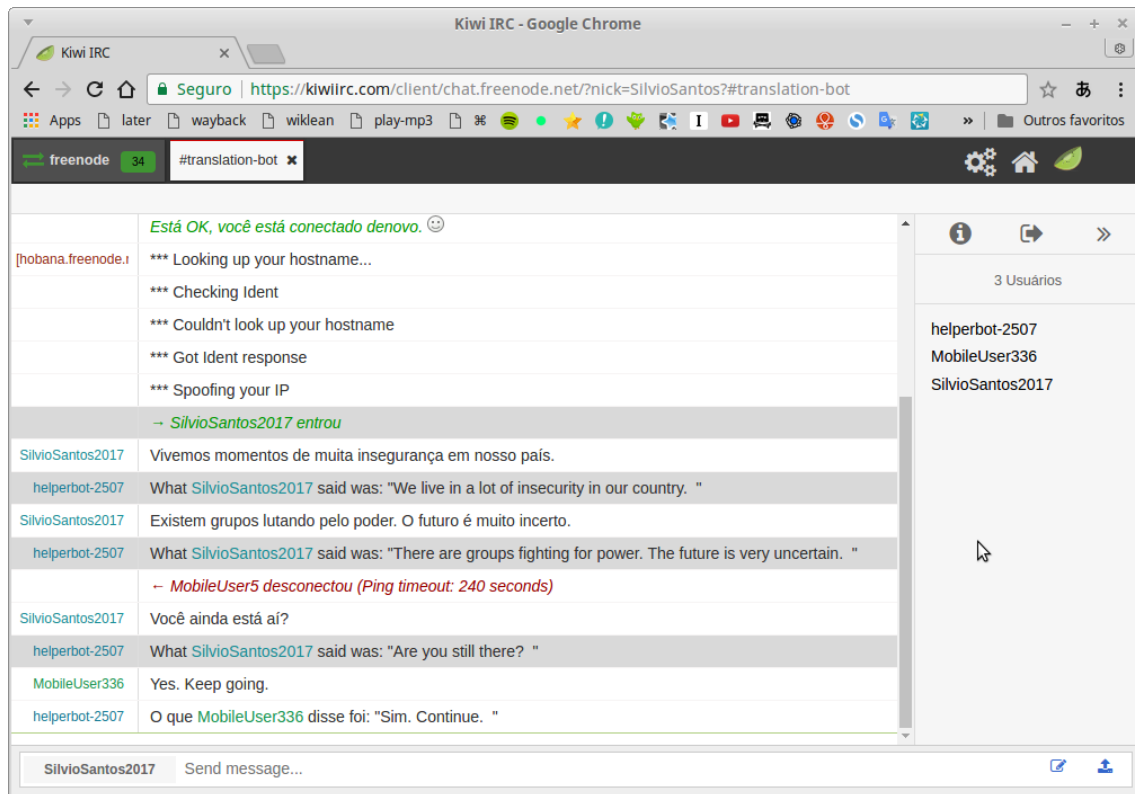


Figure 7.1 - Web session with two participants

We chose the Internet Relay Chat (IRC) protocol to demonstrate some of the ideas we discussed in this dissertation because of the ease with which agents can be implemented as bots and log into the network. Python was used here, but any other language that has a way of sending and receiving HTTP requests could have been used to implement this prototype. Figure 7.1 shows a simple interaction using an IRC Web Client.

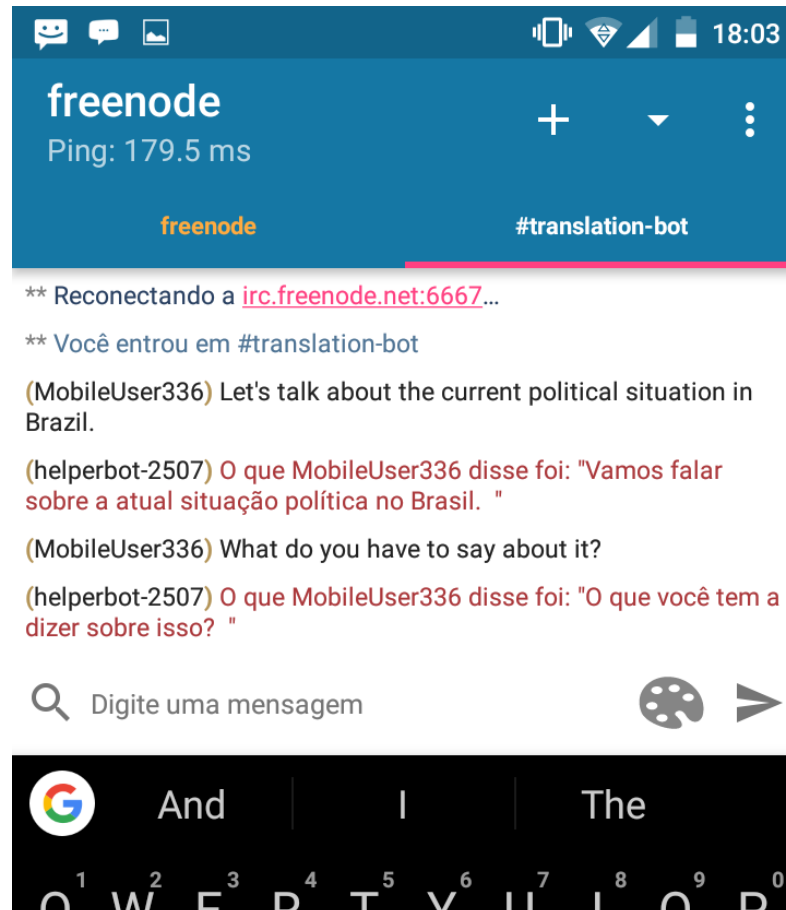
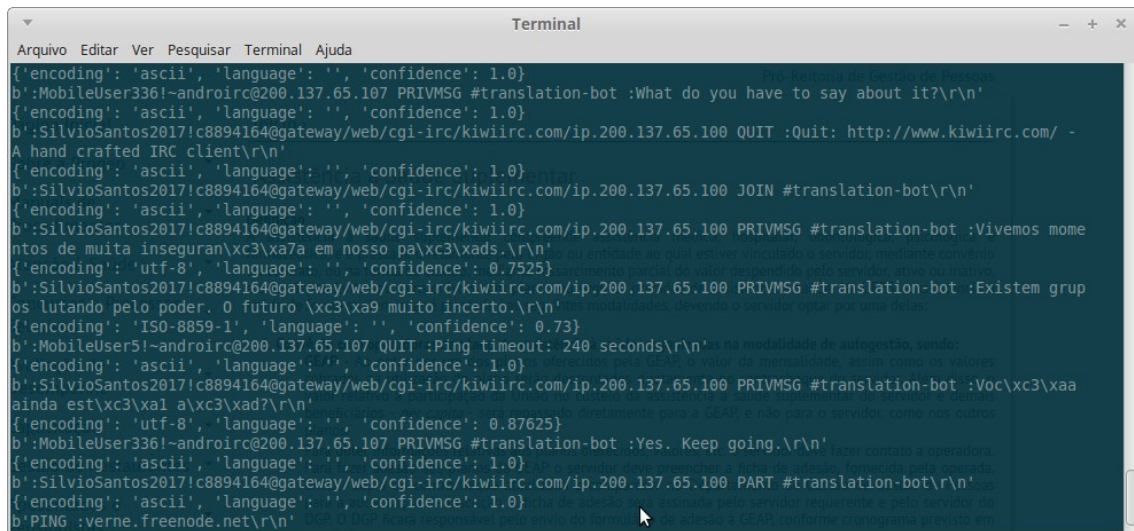


Figure 7.2 Access with a mobile client for Android

Being also a mature and well known protocol, IRC has client for all major platforms, Windows, Linux, Android, MacOS, iOS. That was another factor, due to so having clients for all platforms this system can be used with existing infrastructure, without the need of purchasing and specific hardware or paying for license fees. The Figure 7.2 shows an example of access by a mobile client.



```
Arquivo Editar Ver Pesquisar Terminal Ajuda
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':MobileUser336!~androirc@200.137.65.107 PRIVMSG #translation-bot :What do you have to say about it?\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 QUIT :Quit: http://www.kiwiirc.com/ -
A hand crafted IRC client\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 JOIN #translation-bot\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 PRIVMSG #translation-bot :Vivemos mome
ntos de muita inseguran\xca7a em nosso pa\xca9ads.\r\n'
{'encoding': 'utf-8', 'language': '', 'confidence': 0.7525}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 PRIVMSG #translation-bot :Existem grup
os lutando pelo poder. O futuro \xca9 muito incerto.\r\n'
{'encoding': 'ISO-8859-1', 'language': '', 'confidence': 0.73}
b':MobileUser5!~androirc@200.137.65.107 QUIT :Ping timeout: 240 seconds\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 PRIVMSG #translation-bot :Voc\xca9
ainda est\xca9al a\xca9ad?\r\n'
{'encoding': 'utf-8', 'language': '', 'confidence': 0.87625}
b':MobileUser336!~androirc@200.137.65.107 PRIVMSG #translation-bot :Yes. Keep going.\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':SilvioSantos2017!c8894164@gateway/web/cgi-irc/kiwiirc.com/ip.200.137.65.100 PART #translation-bot\r\n'
{'encoding': 'ascii', 'language': '', 'confidence': 1.0}
b':PING :verne.freemove.net\r\n'
```

Figure 7.3 Agent Event Log

A live log (an example in Figure 7.3) shows what the agent is doing at any moment. Since Freenode servers can use different encodings depending on the client, an encoding guessing library was used to check in which encoding the incoming messages were.

```
{'encoding': 'utf-8', 'language': '', 'confidence': 0.73}
```

That was necessary because the language prediction library required a specific encoding and would break if not provided with unicode.

The other lines show how the communication between IRC server and client is done.

```
b':MobileUser336!~androirc@200.137.65.107 PRIVMSG #translation-
bot :What do you have to say about it?\r\n'
```

The initial **b** denotes a it is a binary string, followed by the user name (MobileUser336), the client he is using (androirc), his ip (200.137.65.107), the kind of message (PRIVMSG, used both to send messages to a channel or to privately message another user), the channel (#translation-bot) and finally the message itself, with encoded carriage return and new line characters.

An operator could, without interfering on the conversation flow, adjust the Agents behavior, changing the services being used, or mode of interaction. This is especially important if several agent instances are working with different groups. From this screen they can be all controlled at once.

8 Final Considerations

The barrier language barrier is an invisible wall that prevents the exchange of ideas and collaboration in several settings. Working towards the dissemination of tools that help mitigate this problem is, in our view, a worthwhile pursuit.

Tools such as Apertium, MOSES and CMUSphinx have the potential of integrating a fully open source system to that purpose. Coupled with the use of crowdsourcing and an open development model, tools like this one can grow to become the backbone important transnational exchanges and experiments.

Less-than-optimal MT output is not, in itself, a factor that impossibilities its use as a key component in the communication system here discussed and prototyped. That said, the difference of quantity and, in some cases, quality of tools available for English and Portuguese can make the architecture less functional. And that problem becomes even worse when we have to deal with less popular languages.

Better collaboration between human and machines may be the key to overcoming the weaknesses in both. But many more tests are needed, as well as the implementation of more modules, to make the system more robust and integrated with today's more popular communication tools (such as Whatsapp and Facebook Messenger).

An extended use in a larger variety of scenarios would make it possible the assembly of a relevant corpus for agent-assisted communication. The insights from that would be the key to fixing the flaws in the system.

All above said, the popularization tools such as this one would be beneficial especially in the learning setting.

8.1 Return to research questions

In this session we take another look at the questions first proposed and address them with new insights acquired from the research and development process.

Q1. Is it possible to promote communication between monolingual speakers of English and Portuguese with the use of technologies available on the web?

For the simple dialogue scenarios we tested, yes.

But it is hard to make sure what part of the communication gets lost in the process. To account for that users have to be willing to spend a lot of time explaining every bit of what they are trying to say, planning in advance what is important focus on, and learning to adapt to the strategies and resources used.

When misunderstandings can lead to costly mistakes, it is better to be safe and use human and experienced translator.

There is also the potential for created an agent that is specialized in helping bilinguals act as interpreters, but more research is needed.

Q2. In what ways does this technology help (or hinder) communication?

MT depends on the corpus used for training. Common sentences will be represented several times in a large corpus, and since a lot of our daily communication is composed of common phrases, this works fine for greetings, casual exchanges and small talk. But when we get creative and start using language in ways that are not common in the corpus, even if completely normal for native speakers, the need to spend time explaining the message step by step may discourage participants to use this resource.

Q3. What is the ideal scenario for the use of such technology?

At the same time the strategies and resources used cannot guarantee participants will understand each other all the time, if they have to communicate and have not alternative language in common to use, or someone to act as an interpreter, they have more chances of reaching successful dialog aided by the agents than alone. But broader tests and more data are needed to address this question.

Q4. Is it possible to build a multiagent architecture to support multilingual interaction?

The tools available today can help native speakers coming from different backgrounds to communicate, even though mistakes are sure to happen. If the participants are patient enough to follow the agents' guidance, they might be able to solve some of the conversation mistakes.

8.2 Experiences acquired with the development of the environment

The will to work with a resource that dealt somehow with language or translation was present even before the idea of this dissertation took place. And this idea of working with language became even stronger when this work began mixing with ideas and research of other LIEd students.

So much so that the agents here discussed became part of a greater structure, created to aid teachers in managing synchronous chat sessions, debates and group collaboration in a text environment.

The communication helper agents here discussed, then gained an entire community of agents designed to help participants manage their doubts, look for research material and question their beliefs. An environment that, when fully realized can help bring about an alternative experience for distance learning and collaborative environments.

The other overlapping parts of this work became dissertations on their own, that of Bruno Gutierrez Clemente, dealing with knowledge basis and the questioning agents, and that of Thiago de Oliveira Tuler, that implements an environment with several tools that can aid collaborative learning. Under the supervision of our advisor, we published some of these ideas under the title "Ambiente Multiagente para Interações Síncronas" [CLEMENTE et al, 2015], cited in session 3.1.6, and present in the 2015 of TISE, a Latin American event on Information Technology in Education.

Among the several difficulties faced during research and development were my own doubts on my ability to finish this dissertation.

The greatest flaw is not having put the agents to enough stress tests, specially for environments with several participants. Some of the strategies are certain to break and new ones to come from the experience of having the agents mediate real classes or business meetings, with all their messiness and broken conversation threads.

8.3 Future Research

Upon reading the dissertation of another LIEd graduate, Bruno Clemente [2016], I realized that the *Comm_Agent* (Figure 6.2) should have been implemented using AIML. Using this specialized modeling language could have made the Agent much more versatile and natural. Implementing new commands and behaviours would also have been much easier than using basic *string compare* to check whether an expression of interest was used. Furthermore incorporating the agents discussed in this work with the ones developed by Clemente and the ones developed by Tuler [2017] in their research. Creating an agent community capable of supporting a large number of learning and interaction situations in synchronous textual environments.

Working with MOSES and Apertium, two open source MT systems, would give us better control over the MT process, as we would not only be able to adjust the settings of the system, but to also implement our own translators for languages that are not so popular on the world stage. Languages such as LIBRAS (Brazilian Sign Language), the tens of indigenous languages and the ones spoken by the descendants of immigrants (Guarani, Nheengatu, Pirahã, Pomerano, etc).

Incorporating Voice Recognition and Speech Synthesis into the prototype was part of the initial project idea, but unfortunately, I could not set up CMUSphinx [LAMERE, 2003] to work properly. And this project was the only open source candidate system to do the recognition part. This addition could make the system much more useful, as

well as several magnitudes more complex.

Having worked with translation and being a language learner, I realized having quick access to all the resources that interface with the *Lang_Agent* (Figure 6.2) can be a great help on my future translation projects, and also when trying to learn another foreign language. As well as customizing the agents translate only the words I have not yet memorized, helping on text exploration. Making the agent a companion tool to services like Duolingo [HACKER, 2014] and Anki [ELMES, 2017].

From the experience in the graduation program, I could see that even though all of Information Technology graduate students had to read and work with texts in papers and books in English, many did not feel comfortable with the task. So creating a collaborative environment for the discussion and exploration of foreign texts could benefit our graduate students. And maybe even bring to our attention interesting works that were written in languages other than English. Even though this dissertation is written in English, it was born from the very desire to do away with this kind of language barrier.

9 References

AMORIM, Marcello N. de. TSL Uma Linguagem para Especificação de Tradutores. Dissertação de Mestrado, Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo, 2007.

BAHDANAU, D., Cho, K. & Bengio, Y., 2014. Neural Machine Translation by Jointly Learning to Align and Translate. Available at: < <http://arxiv.org/abs/1409.0473> >. Accessed November 16, 2016.

BLACK, Paul E. "Zipf's law "in Dictionary of Algorithms and Data Structures, Vreda Pieterse and Paul E. Black, eds. 24 Aug. 2009. Available in < <https://www.nist.gov/dads/HTML/zipfslaw.html> >. Accessed on 3 Oct. 2016.

BLISS, Charles Kasiel. International Semantography: A Non-alphabetical Symbol Writing Readable in All Languages. A Practical Tool for General International Communication Especially in Science, Industry, Commerce, Traffic, Etc., and for Semantical Education, Based on the Principles of Ideographic Writing and Chemical Symbolism. Institute for Semantography, 1949.

BOULTON, Marjorie. Zamenhof: creator of Esperanto. London: routledge and Paul, 1960.

BREDA, Wesley Lucas. Um Ambiente de Apoio à Tradução Baseado em Conhecimento - Estudo de Caso Português-Libras. Dissertação de Mestrado, Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo, 2008.

CARVALHO, M. J. S., Nevado, R.A., Menezes, C.S. (2005) Arquiteturas pedagógicas para a educação a distância: conceitos e apoio telemático. Proceedings of the XVI SBIE, Juiz de Fora – MG – Brasil.

CATFORD, John C. (1965) A Linguistic Theory of Translation: an Essay on Applied Linguistics, London: Oxford University Press.

CHO, K. et al., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. Available at: <http://arxiv.org/abs/1406.1078> [Accessed November 16, 2016].

CLEMENTE, Bruno G. R. Um Ambiente para Construção de Chatterbot com Evolução Dinâmica da Base de Conhecimento. Dissertação de Mestrado, Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo, 2016.

CORBÍ-BELLOT, M. et al. (2005) "An open-source shallow-transfer machine translation engine for the romance languages of Spain" in Proceedings of the European Association for Machine Translation, 10th Annual Conference, Budapest 2005, pp. 79–86.

CRYSTAL, David. English as a global language, second edition. Cambridge university press, 2012.

DESCARTES, René. [Letter] November 20, 1629, Amsterdam. [to] MERSENNE, Marin. Paris. Available in < <http://ckcc.huygens.knaw.nl/epistolarium/letter.html?id=desc004/1022> >. Accessed on June 30, 2017.

DORR, Bonnie J.; JORDAN, Pamela W.; BENOIT, John W. A survey of current paradigms in machine translation. Advances in computers, v. 49, p. 1-68, 1999.

EDUCATION FIRST. Inglês Proficiency Index. Available in < http://www.ef.com.br/___/~/media/centraleftcom/epi/v4/downloads/full-reports/ef-epi-2014-portuguese.pdf >. 2014. access: June 20, 2017.

ELMES, D. ANKI - spaced repetition system. [Software]. 2017. Accessed in June 27, 2017. Available at < <https://apps.ankiweb.net/> >.

FALK, Julia S. Words without grammar: linguists and the international auxiliary language movement in the United States. *Language & communication*, v. 15, n. 3, p. 241-259, 1995.

HACKER, Severin Benedict Hans. Duolingo: Learning a language while translating the web. 2014. PhD Thesis CMU-CS-14-116, Carnegie Mellon University, 2014. 5.

HUTCHINS, John. The History of Machine Translation in a Nutshell. Available in <<http://www.hutchinsweb.me.uk/Nutshell-2005.pdf>>. Acesso em: 3 mar 2015. DORR, Bonnie J e JORDAN, Pamela W e BENOIT, John W. A Survey of Current Paradigms in Machine Translation. [S.l: s.n.], 1999. v. 49.

_____. Two precursors of machine translation: Artsrouni and Trojanskij. *International Journal of Translation*, v. 16, n. 1, p. 11–31, 2004. Available in <<http://www.hutchinsweb.me.uk/IJT-2004.pdf>>. Accessed on June 30, 2017.

JAKOBSON, Roman."On linguistic aspects of translation." *On translation* 3: 30-39. 1959.

KATIKALA, Soujanya. Google™ Project Loon. *InSight: Rivier Academic Journal*, v. 10, n. 2, p. 1-6, 2014.

KOEHN, Phillipp. HOANG, H. BIRCH, A. CALLISON-BURCH, C. Moses: Open Source Toolkit for Statistical Machine Translation. , (June), pp.177–180. 2007.

KOEHN, Philipp. Open Problems in Machine Translation. Youtube, University of Edinburgh Channel, March 25, 2013. Available in <<https://www.youtube.com/watch?v=6UVgFjJeFGY>>. Accessed in July 30, 2017.

KONSTANTIN, Adreev. The Univesal Langaue. Youtube, Dalarna University Channel. June 9, 2011. Available in <<https://youtu.be/8FD-z3JCCzM>>. Accessed on July 30, 2017.

LAMERE, P. KWOK, P. WALKER, W. GOUVÊA, E. B. SINGH, R. RAJ, B. WOLF, P. Design of the CMU sphinx-4 decoder. In: INTERSPEECH. 2003.

LANG, Sonja. "Toki Pona, the language of good." Tawhid, May 25, 2014.

MATT. Dr Sbaitso was my only friend. 2006. Accessed on July 27, 2017. Available at < <http://www.x-entertainment.com/articles/0952/> >.

MOED, HF. Country Trends - The Language of (Future) Research Communication. Research Trends. n. 31, p.1-12, 2012.

OKRENT, Arika. In the land of invented languages. NY: Spiegel and Grau, 2009.

OLIVEIRA, Gilvan M. Multilingualism in Brazil. Brasilia: UNESCO Office in Brazil / IPOL 2008. Available at: <<http://unesdoc.unesco.org/images/0016/001611/161167por.pdf>>. Accessed on June 20, 2017.

PIANTADOSI, Steven T. Zipf's law word frequency in natural language: a critical review and future directions. June 2, 2015.

ROGERS, Stephen D. The Dictionary of Made-Up Languages: From Elvish to Klingon, The Anwa, Reella, Ealray, Yeht (Real) Origins of Invented Lexicons. Simon and Schuster, 2011.

SCIENCEWATCH. Country Profiles. Available at: <<http://archive.sciencewatch.com/dr/cou/pdf/09decALL.pdf>> 2009. Accessed on June 20, 2017.

SOUZA, José Pinheiro. Theories of Translation: an integrated view. Journal Letters No. 20 volume ½ Jan / Dec 1998.

TARDY, C. (2004) "The role of Inglês in scientific communication: lingua franca or

Tyrannosaurus rex?", In: Journal of Inglês for Academic Purposes, v. 3, n. 3, pp. 247-269.

TULER, Thiago de O. Ambiente para Interações Síncronas em Aprendizagem Colaborativa. Dissertação de Mestrado, Programa de Pós-Graduação em Informática, Universidade Federal do Espírito Santo, 2017.

UMBERTO, Eco. Dire Quasi la Stessa Cosa. Milano, Bompiani, 2003.

VASHEE, K., 2016. eMpTy Pages: A Deep Dive into SYSTRAN's Neural Machine Translation (NMT) Technology. Available at < <http://kv-emptypages.blogspot.com.br/2016/09/a-deep-dive-into-systrans-neural.html> >. Accessed on November 16, 2016.

VENNERS, B., & Rossum, G. van. (2003). The Making of Python. Available at < <http://www.artima.com/intv/pythonP.html> >. Accessed on July 30, 2017.

WEIJEN, Daphne van. The Language of (Future) Scientific Communication. Research Treends. Pages 7-8. Issue 31, November 2012.

WERNER OSWALD. Sapir-Whorf Hypothesis. Concise Encyclopedia of Philosophy of Language: 76-83. 1997.

WILKINS, John. An Essay towards a Real Character and a Philosophical Language. 1668. Available in < <https://archive.org/details/AnEssayTowardsARealCharacterAndAPhilosophicalLanguage> >. Accessed on June 30, 2017.

WU, Y. SCHUSTER, M. CHEN, Z. LE, Q. V. NOROUZI, M. MACHEREY, W. KLINGNER, J. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. 2016. Available at < <http://arxiv.org/abs/1609.08144> >. Accessed on November 16, 2016.

ZUCKERBERG, Mark. "Is connectivity a human right." Facebook Newsroom 20. 2013. Available in < <https://www.facebook.com/isconnectivityahumanright> >. Accessed on July 30, 2017.