https://doi.org/10.11606/issn.2317-9511.v32i0p9-31

# Ferramentas de Interpretação Assistida por Computador (CAI) e opções de automação com o Reconhecimento Automático de Fala

# Computer-Assisted Interpreting Tools (CAI) and options for automation with Automatic Speech Recognition

Luís Eduardo Schild Ortiz*
Patrizia Cavallo*

*Computer Systems Engineer. Bachelor's degree in Computer Sciences (Pontifical Catholic University of Rio Grande do Sul – PUCRS, Brazil) and Specialist degree in Translation Studies (PUCRS, Brazil). E-mail: schild.luis@gmail.com

*Translator and Conference Interpreter. Doctoral student in Terminology, Lexicography and Translation at the Federal University of Rio Grande do Sul (UFRGS, Brazil). Master's degree in Interpreting (University of Bologna, Italy) and in Comparative Literature (UFRGS, Brazil). E-mail: patriziacavallo.ita@gmail.com

*Abstract:* In recent years, several studies have indicated interpreters resist adopting new technologies. Yet such technologies have enabled the development of several tools to help those professionals. In this paper, using bibliographical and documental research, we briefly analyse the tools cited by several authors to identify which ones remain up to date and available on the market. Following that, we present concepts about automation, and observe the usage of automatic speech recognition (ASR), while analysing its potential benefits and the current level of maturity of such an approach, especially regarding Computer-Assisted Interpreting (CAI) tools. The goal of this paper is to present the community of interpreters and researchers with a view of the state of the art in technology for interpreting as well as some future perspectives for this area.

*Keywords*: Interpreting; CAI; Automation; Automatic Speech Recognition; ASR.

*Resumo*: Diversos estudos nos últimos anos apontam a resistência de intérpretes à adoção de novas tecnologias. Essas, no entanto, permitiram o desenvolvimento de ferramentas para auxiliar esses profissionais. Nesse artigo analisamos brevemente, por meio de revisão de literatura e de pesquisa documental, as ferramentas citadas por diversos autores para vermos quais delas permanecem sendo atualizadas e disponíveis no mercado. Sucessivamente, apresentamos conceitos de automação, e observamos o uso de reconhecimento automático de fala (ASR), analisando seus potenciais benefícios e o grau de maturidade atual dessa abordagem, especialmente em ferramentas de Interpretação Assistida por Computador (CAI). O objetivo desse artigo é apresentar à comunidade de intérpretes e de pesquisadores uma visão do estado da arte da tecnologia para interpretação, assim como algumas perspectivas futuras nessa área.

*Palavras-chave*: Interpretação; CAI; Automação; Reconhecimento Automático de Fala; ASR.

# 1 Introduction

There are ongoing debates regarding the so-called "disruptive technologies" (those that may cause fast or sudden paradigm shifts in some fields of society), examples of which are Netflix and online streaming, Uber and Smartphone apps, Bitcoin and Blockchain technology, as well as Artificial Intelligence and Deep Learning. On a recent TV programme, a businessman spoke about the short cycle of reinvention in business and the professionals' need to adapt in a period of increasing automation in several fields[1]. History shows that this situation is not completely new — the businessman argues — and that at least since the Industrial Revolution humanity constantly experienced periods of changes similar to those we are currently witnessing. The field of interpreting is not exempt from this process.

With the onset of simultaneous interpreting, which relied on the help of the technology available at that time, consecutive interpreters criticized and resisted the new working method (PAGURA 2010). However, the changes it brought have persisted and professionals have adapted to it. Today, some people refute the possibility of technology ever replacing translators and interpreters, while continuing to seek out computer programs that facilitate their work. In fact, there is a history of denial regarding the subject (BERBER-IRABIEN 2010: 29; DONOVAN 2006: 5; FIRMINO 2016: 2). Pym (2011: 4) argues: "Resistance to technological change is usually a defense of old accrued power, dressed in the guise of quality". Could today's simultaneous interpreters be facing the same fate as consecutive interpreters in the 20th Century? Or perhaps those professionals are unfamiliar with the tools currently available to them? Corpas Pastor (2017: 7) states that "*although most interpreters are unaware of interpreting technologies or are reluctant to use them*, there are some tools and resources already available, mainly computer-assisted interpreting (CAI) tools" (our italics).

[1] The *Entre Aspas* programme, "O impacto que novas mídias e redes sociais causam nas empresas" [The repercussions of new media and social networks on companies], Globonews, broadcasted on 02 January 2018. Last access on 06 March 2018.

The use of software designed specifically for interpreters still seems to be at an early stage: many still count on basic programs, such as Word and Excel (NEJM 2011; FIRMINO 2016: 16; CORPAS PASTOR 2017: 7), or on non-specialised tools for translators and interpreters (see the list of tools mentioned below). Even Computer-Assisted Interpreting (CAI) tools, which have been designed for interpreters, have yet to be integrated or standardised.

Farwick (2009: 72), when addressing the use of computers in the booths for consulting dictionaries, glossaries and following slides, underlines the overload they cause: "As if simultaneous translation was not enough in terms of multi-tasking, [the interpreters] are now operating a computer in parallel, either to look up information, or to follow slides, or both". In a situation where interpreters employ several computer programs to fulfil their tasks, could those devices be leading to additional overload instead of providing support?

Apart from interpreting tools, we could consider a scenario where interpreters are replaced by technological solutions. In Brazil, a study conducted by Firmino in 2016 quotes Thomas Binder, an AIIC member, who states that when interpreters are asked about Remote Interpreting and other technologies, they do not seem worried about the possibility of being replaced by them in the future (apud FIRMINO 2016: 11). In that research, a group of interpreters were asked to fill out a questionnaire and, regarding the topic of replacement by technologies, 86% of the respondents stated they do not believe interpreters will be replaced by technologies, while 14% do believe in that hypothesis (FIRMINO 2016: 19).

However, research shows interpreters are also adopting new tools due to the considerable benefits they provide (TRIPEPI WINTERINGHAM 2010; FANTINUOLI 2017b). For both academic and commercial projects, there is a growing range of solutions available for interpreters.

Thus, in light of that dynamic scenario, this paper[2] aims to briefly analyse those alternatives and provide the community of interpreters and

[2] This paper was originally written in Portuguese ("Ferramentas de Interpretação Assistida por Computador (CAI) e opções de automação com o Reconhecimento Automático de Fala") and

researchers with an overview of the available CAI tools and the use of automation technologies (especially ASR – Automatic Speech Recognition). To begin, we present some of the support tools currently available to interpreters[3]. In doing so, we hope to put ASR, which is the focus of the second part, in context and clarify its use. In compiling the list of tools, we considered studies published on that topic since 2010. Of the technologies studied in the interpreting field, we focus on ASR as an example of automation because it is the tool that most clearly represents a paradigm shift in the working modality.

# 2. Brief overview of the technologies available to interpreters

Given the rapid pace with which IT solutions (computer programs, service websites, devices, accessories, etc.) appear, evolve and also disappear, it is important to revise the CAI tools listed by researchers in the interpreting field, check their level of development and provide updated information on those tools. In 2010, Tripepi Winteringham investigated which Information and Communication Technologies (ICTs) were most useful to interpreters and under which circumstances. In a series of publications, Fantinuoli (2016, 2017a, 2017b, 2017c), analysed the practical use of technology by translators and interpreters, mainly referring to products such as *TranslatorBank* and *InterpretBank*. Both Costa, Corpas Pastor & Durán Muñoz (2014) and Firmino (2016) have presented an overview of programs used for different interpreting-related tasks. In 2015, Sandrelli conducted a similar analysis, but focusing on Computer-Assisted Interpreter Training (CAIT) tools, in which she wrote that "today professional interpreters use many tools that simply did not exist fifteen years ago, such as notebooks, tablets,

was presented by Ortiz, under the academic supervision of Cavallo, as part of the Specialization Course in Translation Studies: Theories, Practices and Technologies of the Pontifical Catholic University of Rio Grande do Sul, 2018.

[3] In this paper, the terms 'interpreting' and 'interpreters' refer to the interpreting involving spoken languages and not sign languages, since the technologies available to sign-language interpreters may be evidently different and require a separate and specialised study.

smartphones, terminology management software, and so on" (SANDRELLI 2015: 120).

As for the level of maturity of those tools, it is interesting to establish whether they were developed specifically for interpreters, which area of their work they are designed to support, and what kind of support they do, in fact, provide, i.e. how much they facilitate performing tasks or allow new tasks with higher aggregated value to be performed). Such an analysis will result in a better understanding regarding which interpreters' needs are being met and how.

For the sake of clarity, we will define categories according to which tools can be listed. Based on the logical cycle and common routine of the interpreters' work, the categories are the following: **Training**, during which trainee or beginner interpreters perform lexical and terminological searches, but also focus on the consolidation of their mother-tongue and their foreign language(s), and experienced interpreters further enhance their skills; **Preparation**, during which interpreters investigate the main topic for the interpreting service they were hired to provide, apart from preparing glossaries to consult before and during interpreting and getting familiar with the speakers' accents and biographies based on videos and online information, etc. However, most of the preparation work can be conducted by means of non-specialised tools, such as spreadsheets and text editors, and in a non-automatic way, whilst CAI tools can also be adopted. Our focus will be on the preparation involving a higher level of automation, and that level will be split into two categories, the first consisting of tools currently used for compiling specialized corpora during the so-called Corpus Driven Interpreting Preparation (**CDIP**) (FANTINUOLI 2017b), and the second of **Terminology Management Tools**. Terminology extractors and glossary management tools are both included in the latter category, since they operate in similar areas that can overlap. Whilst terminology extractors are designed to extract specialised terms from monolingual or bilingual corpora, glossary management tools are used in the place of simple spreadsheets, but in certain cases and among other functions also include more sophisticated resources for looking

up terms and facilitating that operation in the booth (for which reason we also included those tools in the "Simultaneous" category in the table below).

Afterwards, there is the interpreting phase itself. Due to the differences in the modalities adopted by interpreters — mainly simultaneous (in the booth) and consecutive (in any place the people requesting such a service may be) — two distinct categories are created, simply denominated **Simultaneous** and **Consecutive**, since some of these tools are only used in one of these modalities. Other programs that are interesting to include are those requested for **Remote Interpreting**. They provide an environment in which simultaneous or consecutive interpreting can happen. Lastly, another category with a strong automation component is that of tools whose purpose is the **Replacement** of interpreters, i.e. tools that perform, at least in part, the interpreters' work.

Another aspect considered when categorising the tools was whether they were developed specifically for the community of interpreters, either created within the academic environment or described by their developers as such. There are tools that are considered non-specific because they are designed for more general users, such as *Audacity*, *Dragon* or *Moodle*, which are used by interpreters to meet some of their specific needs. In addition, there are several programs and websites that are even less directed to interpreters, which we will not investigate, although they are used by them and mentioned by a number of authors. Among those tools, there are unit converters, YouTube channels, search engines, Microsoft Office programs (Word and Excel), as well as tools for activities related to the profession, such as platforms aimed at promoting services, CVs and building client portfolios.

The list of programs is compiled according to the criteria specified by the authors of the papers we consulted, which follow in chronological order:

- Tripepi Winteringham (2010) provides a theoretical overview of the usefulness of ICTs in interpreting practice and mentions *Voxtec Phraselator*, a speech recognition software program.
- Costa, Corpas Pastor & Durán Muñoz (2014) mainly list terminology management tools (*Intragloss, InterpretBank, Interplex UE, LookUp* and *Interpreter's Wizard*) and note-taking programs for consecutive

interpreting (*Evernote, Inkeness, Penultimate, LectureNotes, PenSupremacy* and *My BIC Notes*), but also include unit conversion tools (*ConvertUnits, OnlineConversion, Convert, Converto, Convert Units for Free, Units* and *Unit Converter*) and training programs for interpreters (*Black Box, InterpretaWeb, Linkterpreting*).

- Sandrelli (2015) lists several repositories for CAIT, in different formats, such as *Interpr-It* (CD with interviews), *Black Box* and *IRIS* (shared repository of miscellaneous materials — videos, audio files, academic papers, etc.), apart from speech banks in which students or professors can share such materials in their university training courses for interpreters.

- Costa et al. (2016) analyse terminology extractors and provide a list with the following tools: *SDL Multiterm Extract, Simple Extractor, TermSuite, Sketch Engine, Translated s.r.l.*, and *Terminus*. They also include *Frameworks*, i.e. tools that allow for the development of computer programs within certain standards and provide ready-to-use code libraries in several fields of applications, be they general and specific. However, *Frameworks* are more related to the development of computer programs than to ready-to-use tools, which is why we did not include them in our analysis.

- Firmino (2016) provides a list with training programs (*Black Box, IRIS, Marius, Audacity, Moodle, EU Speech Repository,* among others), programs related to speech recognition during interpreting (*PureVoice, IBM Via Voice, Dragon Naturally Speaking*); others for remote interpreting (*Babelverse, Capiche, ZipDx*), and still others intended to replace interpreters by speech-to-speech systems (*IBM/Mastor, Voxtec/Phraselator, NEX/Tele Scouter, IraqComm*). In these systems, a device captures human speech in a source language, searches for equivalents against pre-recorded sequences in a target language and reproduces such a sequence.

- Fantinuoli (2017a) mainly refers to terminology management tools. He divides them in two categories, i.e. 'first-generation' tools (*Interplex*, *Terminus* and *Interpreters' Help),* which manage terminology in a friendly manner without supporting any other part of the interpreting process, and 'second-generation' tools *(InterpretBank* and *Intragloss)*, which offer more advanced functions. He also mentions *Lookup* and *Multiterm* (the latter

being designed for translators), among other tools he describes as obsolete. Fantinuoli (2017b) also refers to *BootCat* and *CorpusMode*, tools used for building corpora.

In order to better describe the range of tools currently available to interpreters, a table follows below. The programs are categorised according to their main functions, whether they are specific for interpreters or not, apart from providing information on the version and date of last update, and official address of the online provider (as of June 2018). Fields with a question mark indicate information that was not found or that we were not able to confirm. Names marked by "*" indicate tools considered obsolete because they are no longer available for purchase, no longer provided by their developers, have been sold or been incorporated into other tools or were no longer available for analysis.

*Chart 1 – Complete list of tools found in the above-mentioned literature review*

| *Name* | *Category (main function)* | *Specific for Interpreters* | *Version* | *Last update* | *Official website* |
|---|---|---|---|---|---|
| *Melissi (Black Box/VIE)* | *Training material* | *Y* | *5* | *Dec/2017* | *www.melissi.co.uk* |
| *IRIS* | *Speech bank* | *N* | *N/A* | *2017* | *https://www.iris-database.org* |
| *Marius** | *Speech bank* | *N* | *?* | *?* | *?* |
| *Moodle* | *Learning platform* | *N* | *3.5* | *15/05/2018* | *https://moodle.org* |
| *EU DG SCIC Speech Rep.* | *Speech bank* | *Y* | *2* | *01/12/2017* | *https://webgate.ec.europa.eu* |
| *Linkterpreting** | *Training material* | *Y* | *?* | *?* | *linkterpreting.uvigo.es* |
| *InterpretaWeb** | *Training material* | *Y* | *?* | *26/04/2013* | *www.interpretaweb.es* |
| *Interpr-It** | *CD with interviews* | *Y* | *?* | *Jan/1995* | *?* |
| *Interplex UE* | *Glossary management* | *Y* | *2.1.1.60* | *07/03/2017* | *www.fourwillows.com* |
| *InterpretBank* | *Glossary management* | *Y* | *5.35* | *23/05/2018* | *www.interpretbank.com* |
| *Interpreters' Help* | *Glossary management* | *Y* | *?* | *2018* | *https://interpretershelp.com* |
| *Interpreter's Wizard** | *Glossary management* | *Y* | *2.0.1* | *17/02/2013* | *?* |
| *Intragloss* | *Glossary management* | *Y* | *2.2* | *2015* | *www.intragloss.com* |
| *LookUp** | *Glossary management* | *Y* | *2.3.7* | *01/08/2005* | *www.lookup-web.de* |
| *Translated s.r.l.* | *Glossary management* | *N* | *N/A* | *2015* | *https://www.translated.net* |
| *BootCaT* | *Corpora building* | *Y* | *1.07* | *21/05/2018* | *bootcat.dipintra.it* |
| *CorpusMode** | *Corpora building* | *Y* | *beta* | *2018* | *www.staff.uni-mainz.de /fantinuo/translatorbank.html* |
| *SDL Multiterm Extract* | *Terminology extraction* | *N* | *2017* | *2017* | *https://www.sdl.com* |

| Name | Category (main function) | Specific for Interpreters | Version | Last update | Official website |
|---|---|---|---|---|---|
| Simple Extractor | Terminology extraction | N | 1.1.2 | 2015 | www.dail-software.com/help/9_en/ |
| Sketch Engine | Terminology extraction | N | 3.101 | April/2016 | https://www.sketchengine.co.uk |
| Terminus | Terminology extraction | Y | 3.1 | 30/03/2009 | www.wintringham.ch |
| TermSuite | Terminology extraction | N | 3.0 | 31/08/2017 | termsuite.github.io/ |
| PureVoice* | Speech recognition | N | ? | 1999? | https://www.qualcomm.com ?? |
| IBM Via Voice* | Speech recognition | N | 10.5 | 2005 | https://www.nuance.com |
| Dragon NS | Speech recognition | N | 15 | 2017 | https://www.nuance.com |
| Evernote | Note-taking | N | 8.0 | 02/05/2018 | https://evernote.com |
| Inkeness* | Note-taking | N | ? | ? | ? |
| Penultimate | Note-taking | N | 6.2.3 | 18/12/2017 | https://evernote.com |
| LectureNotes | Note-taking | N | 2.8.1 | 05/05/2018 | https://www.acadoid.com |
| PenSupremacy* | Note-taking | N | 1.7.1 | 20/03/2012 | ? |
| My BIC Notes* | Note-taking | N | 1.0.2 | 08/04/2013 | https://www.bicworld.com |
| ZipDx | Audio conference | N | ? | 01/09/2018 | https://www.zipdx.info |
| WebEx | Audio and video conference | N | ? | 01/09/2018 | https://www.webex.com |
| Skype | Audio and video conference | N | ? | 01/09/2018 | https://www.skype.com |
| Babelverse* | Audio conference | N | ? | 2010? | http://babelverse.com |
| Capiche | Automatic Translation of Text | Y | N/A | Not available yet | https://eit.europa.eu |
| IBM/Mastor* | Speech-to-speech system | Y | ? | 2006? | https://www.ibm.com |
| Voxtec/Phraselator | Speech-to-speech system | Y | SQ.410 | 2015 | http://voxtec.com |
| NEC/Tele Scouter* | Speech-to-speech system | Y | ? | ? | ? |
| IraqComm* | Speech-to-speech system | Y | ? | 2006? | https://www.sri.com |

A first look at the table immediately reveals the large number of discontinued tools. Out of the 40 tools listed above, 16 are no longer available or have not been updated in 5 or more years. We could have included them in the analysis for historical purposes, but we decided to focus on the updated programs. The table below only contains the updated tools, with their functions listed according to the previously described categories.

*Chart 2 – List of updated tools by category*

| *Name* | *Category (main function)* | *Specific for Interpreters* | *Training* | *Prep. (CDIP)* | *Prep. (Terminology Management)* | *Simultaneous* | *Consecutive* | *Remote Interpreting* | *Replacement* |
|---|---|---|---|---|---|---|---|---|---|
| *Melissi (Black Box/VIE)* | *Training material* | *Y* | *X* | | | | | | |
| *IRIS* | *Speech bank* | *N* | *X* | | | | | | |
| *Moodle* | *Learning platform* | *N* | *X* | | | | | | |
| *EU DG SCIC Speech Rep.* | *Speech bank* | *Y* | *X* | | | | | | |
| *Interplex UE* | *Glossary management* | *Y* | | | *X* | *X* | | | |
| *InterpretBank* | *Glossary management* | *Y* | | | *X* | *X* | | | |
| *Interpreters' Help* | *Glossary management* | *Y* | | | *X* | *X* | | | |
| *Intragloss* | *Glossary management* | *Y* | | | *X* | *X* | | | |
| *Translated s.r.l.* | *Glossary management* | *N* | | | *X* | *X* | | | |
| *BootCaT* | *Corpora building* | *Y* | | *X* | | | | | |
| *SDL Multiterm Extract* | *Terminology extraction* | *N* | | | *X* | | | | |
| *Simple Extractor* | *Terminology extraction* | *N* | | | *X* | | | | |
| *Sketch Engine* | *Terminology extraction* | *N* | | | *X* | | | | |
| *Terminus* | *Terminology extraction* | *Y* | | | *X* | | | | |
| *TermSuite* | *Terminology extraction* | *N* | | | *X* | | | | |
| *Dragon NS* | *Speech recognition* | *N* | | | | *X* | | | |
| *Evernote* | *Note-taking* | *N* | | | | | *X* | | |
| *Penultimate* | *Note-taking* | *N* | | | | | *X* | | |
| *LectureNotes* | *Note-taking* | *N* | | | | | *X* | | |
| *ZipDx* | *Audio conference* | *N* | | | | | | *X* | |
| *WebEx* | *Audio and video conference* | *N* | | | | | | *X* | |
| *Skype* | *Audio and video conference* | *N* | | | | | | *X* | |
| *Capiche* | *Automatic Translation of Text* | *Y* | | | | | | *X* | *X* |
| *Voxtec/Phraselator* | *Speech-to-speech system* | *Y* | | | | | | | *X* |

It can be noted that 10 of the 24 listed tools were designed with interpreters as the target users. However, overall, there are few tools specifically directed at the interpreting process itself, either Simultaneous or Consecutive, with nine occurrences. Of those, only four have been specifically designed for interpreters (*Interplex EU, InterpretBank, Interpreters' Help* and *Intragloss* terminology management tools) and, even so, their main function is glossary management, i.e. they mainly belong to the Preparation category. As for the others, *Dragon NS* is not used directly by the interpreter, but as a

basis for the advanced Automatic Speech Recognition function of *InterpretBank* (discussed in the next section), while *Evernote, Penultimate* (both belonging to the same application suite) and *LectureNotes* are note-taking programs only. Considering that encoding, memorising and decoding are exhausting tasks for the interpreter in cognitive terms (TRIPEPI WINTERINGHAM 2010: 88) that also demand a high degree of accuracy, one would expect that a larger range of technological tools would be available to support those phases. On the other hand, if a program were capable of unburdening the human interpreter, by automating one or more processes of the interpreting task, one could envisage that programs with such a capability would, as they evolve, be directed at the hypothetical replacement of the human being by the machine.

It is interesting to observe that the programs above do not yet compose app suites. Instead, each one of them comprises few resources, obliging the users to have a number of tools available to meet all their needs. As mentioned in the introduction, it seems we have not yet reached a stage of constant professionalized development, in which companies are interested in regularly developing and marketing solutions, but rather most solutions come about as a result of academic research[4]. Costa, Corpas Pastor & Durán Muñoz (2014) corroborate that understanding:

> Although the number of these technologies is growing fast due to an increasing interest towards interpreters' needs, they are still insufficient and unable to fulfil all the necessary requirements. There is an urgent need to develop technologies that automate the process (p. 7).

# 3. Automation and the use of ASR in tools for interpreters

Interpreters' and interpreting tools can be said to be moving towards a complete IT solution when they include not only several resources in one

[4] Sandrelli (2015: 121) mentions several examples, some of which appear in the list of obsolete tools above.

program, but also automation solutions. For some time, there has been considerable expectation regarding what these technologies can do for the interpreter. Tripepi Winteringham (2010: 90) summarizes some ideas from Kelly (2008) on this matter:

> (…) computers and new technologies offer potential for easing some of the transfer burdens related to interpreting tasks, in that they can help interpreters in their real-time work providing them with quick access to a broader range of information in electronic dictionaries, databases and glossaries. These powerful technological CAI tools include terminology aids, such as laptops, notebooks, small handheld PDAs (Personal Digital Assistants) or similar instruments with Internet accessibility that may facilitate interpreters' work.

Within the last few years, PDAs have disappeared and there have been several technological advances. However, has technology really evolved sufficiently to provide that support? On this matter, Tripepi Winteringham (2010: 90-91) argues:

> Theoretically, these tools should represent the most effective information interface when interpreting, but is their practical use feasible and does rendition benefit? The main drawback of the use of these tools is that it is still considered, at least in the booth, to some extent as unnatural (DONOVAN 2006: 5), presumably because it may be time-consuming and distracting in an activity that requires concentration and fast-paced decoding and delivery. The interpreter at work may not have the time or the cognitive ability to look up a word online or in his/her electronic dictionary, or detect and choose the correct translation (...) In addition, as Veisbergs (2007: 80) states, should the right word be found it may not be possible to incorporate it smoothly in speech.

In 2017, the concern about the actual support technological tools may provide increased. A paper published by Prandi (2017) proposed a model to operationalise hypotheses concerning the use of CAI tools in the booth, which makes it possible to compare "traditional" methods of terminological search (through Word and Excel) and the use of CAI tools. According to Prandi, interpreters are sceptical regarding the use of CAI tools in the booth and question the extent to which they help them in their task, suggesting they may represent a hindrance or an additional cognitive load during the interpreting task (PRANDI 2017).

Nonetheless, thanks to developments in certain IT areas, the level of support provided by tools is improving, maybe overcoming scepticism. For example, the use of speech recognition programs in the booth to look up specialised terms and their correspondents in a target language could lighten the workload during simultaneous interpreting (TRIPEPI WINTERINGHAM 2010: 92-94). Some examples of that development are now more concrete. Fantinuoli introduces two examples of automation: the use of corpora in the interpreter preparation (CDIP) and the use of ASR during the work in the booth (FANTINUOLI 2017b, 2017c).

CDIP aims to automate the building process of a corpus, i.e. a group of texts from which the interpreter can extract specialised terminology before an interpreting assignment, therefore facilitating and speeding up his/her preparation, apart from being able to investigate the use of terms in real contexts. Based on Data Driven Learning, or DDL (BOULTON 2009), that use of corpora requires, first of all, the gathering of documents to build corpora, and then the use of a concordancer to locate language use cases. Tools such as *BootCat* and *CorpusMode* automate both procedures, allowing tasks that might take hours or days to be done in a few minutes. During the preparation phase, the use of Terminology Extractors such as *SDL Multiterm Extract*, *Sketch Engine* and *Terminus*, make the whole process much easier. Whilst CDIP can save considerable time during the preparation phase, ASR tools can be used directly in the interpreting phase.

The benefits of using ASR could be compared to having a virtual booth mate, who looks up the glossary for the interpreter. ASR is defined as the computational process of converting human speech into a sequence of words (JURAFSKY; MARTIN 2009 apud FANTINUOLI 2017c: 28), and has several applications: educational, military (as in interpreting, military projects also use ASR to relieve the workload for fighter pilots[5]), support for the disabled, speech transcription for documents etc. Even though ASR itself is not frequently found in the translation environment, ASR technology is behind the quality leap in Machine Translation (MT). The addition of Deep Learning and Neural Artificial Networks to the range of approaches adopted for creating

[5] https://www.eurofighter.com/the-aircraft. Last access on 25 February 2018.

ASR solutions and other Natural Language Processing Applications are the same used by Google Translator, which so impressed users with its leap in quality in 2016 and 2017[6]. Historically, ASR performance has been assessed through the Word Error Rate (WER), submitted to recognition tasks against pre-defined corpora. In the 1990s, the DARPA (*Defense Advanced Research Projects Administration*) agency of the US Department of Defense supported the elaboration of those tasks, *Switchboard* being one of the most challenging (JUANG; RABINER 2005). A WER of 10% for a viable ASR system was considered acceptable (JUANG; RABINER 2005; SAON et al. 2017). In March 2017, IBM announced that, after testing the ASR system developed by the Watson team against the *Switchboard* corpus, results of 5.5% WER was achieved (SAON et al. 2017; LAZZARO 2017; LEWIS-KRAUSDEC 2016). Therefore, the results point not only to its feasibility but to an approximation of the sought-after "human parity", i.e. the same error rate humans experience in the comprehension of words during a conversation, estimated by IBM at 5.1% (SAON et al. 2017). With that level of precision, the practical use of ASR in interpreting is possible. The *InterpretBank* tool described by Fantinuoli (2017c) provides a good example of that. *InterpretBank* is mainly used by interpreters to prepare specialised glossaries (via CDIP, Terminology Extractor or another editor) and easily access a glossary in the booth thanks to advanced lookup functions, such as the partial entry of terms and typing error correction[7]. In its latest versions, however, it introduces a significant new function, although on an experimental basis. In the booth, the tool, apart from allowing the interpreter to look up glossaries, receives the same audio input as the interpreter, feeding it to an ASR module (which can be configured to use one of several commercially available modules, for example, *Dragon NS*). It then transcribes the audio input into text, and displays it on the screen. The transcribed text is then forwarded to the CAI module, which searches in the terminological database for a target term, and for problematic elements such as numbers,

[6] Lewis-Krausdec (2016) describes the project that led to that quality leap, including a seven-point increase in the so-called BLEU metric, used for comparing the performance in automatic translation, in which increases of two points were already considered optimal.

[7] A detailed description of *InterpretBank* prior to the use of ASR can be found in Fantinuoli (2016a).

acronyms and proper names. The search results are also displayed, effectively automating the term lookup during interpreting.

With this example, the idea of interpreters being replaced by the total automation of their work re-emerges. Among the items in the questionnaire Firmino submitted to interpreters (see above), one refers to the belief that technology may replace the interpreter. Of the 64 participants, 86% claimed not to believe in that hypothesis, against 14% who claimed to believe in it. That result would seem to be related to aspects of the work that machines would be supposedly unable to perform:

> Among the reasons the led some respondents to claim that technology will not replace the interpreters in the future, there are: nuances, linguistic variation, non-verbal communication, accents, linguistic subtleties, emotion, understanding of the 'between the lines', flexibility of the human being's adaptation, decision-taking, reliability, culture, metaphors, intonation, irony, ambiguities, unpredictability, capability of judgment (FIRMINO 2016: 19, our translation)[8].

ASR has been under development for decades, and obstacles remain, some of which are included in Firmino's list above. However, the current technological progress should guarantee the continuity of research aimed at solving them. Fantinuoli (2017c) comments on some of these problems in the context of use of *InterpretBank*. First of all, in order to be used in different interpreting scenarios, the software requires ASR to be speaker-independent, that is, it should not need to be trained to correctly transcribe the speech of each speaker (a technique adopted by some and that increases, for that person's speech, the accuracy of recognition). That kind of training would limit its functional application. However, the development of a speaker-independent ASR, which recognises any kind of speech, faces the challenge of understanding different pronunciation variations, such as accents, thus reducing its accuracy. The use of casual speech aggravates that scenario due

[8]"Dos respondentes que afirmaram acreditar que a tecnologia não substituirá os intérpretes no futuro, as justificativas foram: as nuances, variações linguísticas, comunicação não verbal, sotaques, sutilezas do idioma, emoção, entendimento das entrelinhas, flexibilidade de adaptação do ser humano, tomada de decisões, confiabilidade, cultura, metáforas, entonação, ironia, ambiguidades, imprevisibilidade, capacidade de discernimento".

to disfluencies, such as hesitations, repetitions and interruptions during pronunciation while fast speech can make the problem even worse. Another challenge is the language ambiguity, as occurs with homophones. How can they be transcribed? In translation, new artificial intelligence approaches allow Google Translator to use context information to decide which equivalence should be used, as for the word "light", for example (LEWIS-KRAUSDEC 2016). However, Google Translator can access a full paragraph of text from which it can extract that context (are we talking of "light" as opposed to darkness, or "light" as opposed to heavy?).

ASR used in the booth, however, needs to have a fast response time (another prerequisite listed by Fantinuoli for *InterpretBank*), and cannot wait for a long speech excerpt to solve such an ambiguity. Another typical problem is continuous speech. For a machine, identifying the correct separation of words in continuous speech is not an easy task. An ASR field of application focuses on the recognition of words pronounced separately (as in the telephone automatic systems used by remote service apps), but that approach does not apply to conference interpreting, where potentially long sentences need to be identified. Other *InterpretBank* prerequisites for ASR are: supporting the recognition of a wide vocabulary range, supporting the customisation of that vocabulary to recognise specialised terms and a fast response time (required values are not specified), apart from the previously mentioned low WER.

*Dragon NS* software, used in the experiment reported by Fantinuoli (2017c), meets most of those prerequisites, although curiously it is described by its developer as a speaker-dependent ASR[9]. Apparently, technology allows for a certain degree of flexibility on that matter. According to the information obtained on *InterpretBank*[10], the transcription resource and the term lookup by ASR are only included on an experimental basis due to the technological limitations mentioned above, the lack of an ASR package or solution that is

[9] Information found at: https://www.nuance.com/dragon/transcription-solutions.html: "Dragon is not designed for use with multiple speakers or to transcribe lectures. The software does, however, support the voice writing transcription technique, where the transcriptionist repeats or "parrots" audio from other speakers" (Last access on 05 March 2018).

[10] Contact by e-mail with its main developer.

economically accessible, can be used and installed on a computer (there are online solutions, but they may incur in confidentiality violations), performs well (processing speed), and is multilingual.

However, if the performance of ASR in this tool is good enough, it is easy to see how it might lighten the interpreter's workload, and provide that discrete virtual booth mate mentioned by Fantinuoli. His paper includes a case study with WER metrics and response times for terminological searches, with results described as satisfactory under "standard" interpreting conditions (FANTINUOLI 2017c). Prandi (2017) compares interpreters' performances in scenarios consisting of 'traditional' glossary lookup (Word and Excel) versus searches conducted using CAI tools (*InterpretBank* without ASR). It would be interesting to conduct a similar study including automatic searches using ASR, so as to obtain concrete information concerning the advantages of using that approach.

# Final remarks

Based on the literature review showing the low use of CAI tools by interpreters, either due to reluctance or a lack of knowledge, we performed a brief analysis of the available tools. We found that of the 40 tools analysed, 16 were no longer available, 10 were developed specifically for interpreters, and none has been designed primarily to assist interpreters in the booth, although glossary management tools include search functions. We have presented some of the concepts related to automation in interpretation today and mainly of ASR, with an in-depth look at the example of the *InterpretBank* tool. Currently, it seems automation technology is close to being able to provide interpreters with truly useful tools, but still far from achieving the level of maturity required to replace humans.

Nevertheless, the rapid progress allows us to glimpse the future with automated interpretation. The imagination can really fly if we consider ASR being combined with other areas of research - quantum processors tremendously increasing processing capacity, with artificial neural networks

using machine learning (Deep Learning) running on these processors, in turn allowing ASR modules to function with high rates of precision and performance. If a CAI tool then combines the ASR with the Automatic Translation and Text-to-speech (TTS) technologies, all technologies currently under development, could we achieve automation at all stages of the process?

In addition to the aforementioned study to measure gains obtained with the use of ASR, other areas of study might help enhance the tools used by interpreters. The present analysis of the supply of technology-based tools can be taken further to reveal market niches that might justify greater investment in the development and improvement of those tools. Improved graphic capabilities, ergonomics, and the use of other input devices such as touch screens are examples of paths that could lead to improvement. Research on basic technologies, such as those mentioned in the previous paragraph, are especially attractive for producing more significant changes.

# References


BERBER-IRABIEN, D. *Information and communication technologies in conference interpreting*: a survey of their usage in professional and educational settings. 2010. 272 f. Doctoral Thesis – Universitat Rovira I Virgili, Department of English and German Studies, Tarragona, 2010. Available at: https://www.tesisenred.net/bitstream/handle/10803/8775/tesi.pdf?sequence=1 Last access on 20 May 2018.

BOULTON, A. Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 2009, v. 35, n. 1, pp. 81-106. Available at: https://hal.archives-ouvertes.fr/hal-00326990v1/document Last access on 4 June 2018.

CORPAS PASTOR, G. VIP: Voice-Text Integrated System for Interpreters. *Proceedings of the 39th Conference Translating and the Computer*, Geneva, pp. 7-10, 2017. Available at: https://www.asling.org/tc39/?page_id=3223 Last access on 20 June 2018.

COSTA, H.; CORPAS PASTOR, G.; DURÁN MUÑOZ, I. Technology-assisted Interpreting. *MultiLingual*. v. 25, n. 143, pp. 27-32, 2014. Available at: https://eden.dei.uc.pt/~hpcosta/docs/papers/201404-MultiLingual-published_version.pdf Last access on 30 June 2018.

COSTA, H. et al. Nine Terminology Extraction Tools: Are they useful for translators? *MultiLingual*, v. 27, n. 3. April 2016. Available at: https://eden.dei.uc.pt/~hpcosta/docs/papers/201604-MultiLingual.pdf Last access on 20 May 2018.

DONOVAN, C. Trends – Where is interpreting heading and how can training courses keep up? *Presentation at the EMCI conference: The Future of Conference Interpreting: Training, Technology and Research*, v. 30, 2006. Available at: http://www.emcinterpreting.org/?q=system/files/Where%20is%20interpreting%20heading%20Donovan.pdf Last access on 10 June 2018.

FANTINUOLI, C. InterpretBank. Redefining computer-assisted interpreting tools. *Proceedings of the Translating and the Computer 38 Conference in London*, Geneva, pp. 42-52, 2016. Available at: http://www.staff.uni-mainz.de/fantinuo/download/publications/InterpretBank%20Redefinin

g%20computer-assisted%20interpreting%20tools.pdf Last access on 30 May 2018.

FANTINUOLI, C. Computer-assisted interpreting: challenges and future perspectives. In: DURÁN, IS.; CORPAS, G. *Trends in e-tools and* RESOURCES *for translators and interpreters*, Leiden: Brill, 2017a, pp. 153-174.

FANTINUOLI, C. Computer-assisted preparation in conference interpreting. *Translation & Interpreting*, v. 9, n. 2, pp. 24-37. 2017b. Available at: http://www.trans-int.org/index.php/transint/article/view/565/294

FANTINUOLI, C. Speech Recognition in the Interpreter Workstation – *Proceedings of the 39th Conference Translating and the Computer*, Geneva, pp. 25–34, 2017c. Available at: http://www.staff.uni-mainz.de/fantinuo/download/publications/Speech%20Recognition%20in%20the%20Interpreter%20Workstation.pdf Last access on 25 June 2018.

FARWICK, J. S. New Technologies and Conference Interpretation. *Proceedings of the Translation Conference held on November 8, 2008 at the University of Portsmouth*. pp. 70-79, 2009. Available at: http://www.port.ac.uk/media/contacts-and-departments/slas/events/tr08-farwick.pdf Last access on 30 June 2018.

FIRMINO, P. M. RA. SI. Os intérpretes e a tecnologia: uma parceria fundamental. *Tópicos e contextos em interpretação*, v. 1, 2016. Available at: http://interpret2b.com/cms/uploads/publicacoes/publicacao_7/os%20interpretes%20e%20a%20tecnologia%20-%20uma%20parceria%20fundamental_Paula%20Firmino.pdf Last access on 15 June 2018.

JUANG, B.; RABINER, L. R. Automatic Speech Recognition - A Brief History of the Technology Development. *Georgia Institute of Technology*. Santa Barbara: Atlanta Rutgers University and the University of California., v. 1, pp. 67, 2005. Available at: https://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf Last access on 15 May 2018.

KELLY, N. Moving toward machine interpretation. *Multilingual*, Dec. 2008.

LAZZARO, S. 03/06/2017. *Google Is Using Artificial Intelligence to Make a Huge* CHANGE *to Its Translate Tool*. Available at: http://observer.com/2017/03/google-translate-neural-update/. Last access on 9 February 2018.

LEWIS-KRAUSDEC, G. 14/12/2016. *The Great A.I. Awakening: How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself*. Available at: https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html. Last access on 9 February 2018.

NEJM, C. C. L. *Interpretação simultânea*: *a linguística de Corpus na preparação do intérprete*. 2011. 205 f. Dissertation (*Mestrado em Estudos Linguísticos e Literários em Inglês*) – Department of Modern Letters, USP, São Paulo, SP.

PAGURA, R. J. *A interpretação de conferências no Brasil: história de sua prática profissional e a formação de intérpretes brasileiros*. 2010. 231 f. Doctoral Thesis (*Doutorado em Estudos Linguísticos e Literários em Inglês*) – Department of Modern Letters, USP, São Paulo, SP.

PRANDI, B. Designing a Multimethod Study on the Use of CAI Tools during Simultaneous Interpreting, *Proceedings of the 39th Conference Translating and the Computer*, Geneva, 2017. Available at: https://www.asling.org/tc39/wp-content/plugins/download-attachments/includes/download.php?id=3263 Last access on 15 June 2018.

PYM, A. What technology does to translating. *Translation & Interpreting*, v. 3, n. 1, pp. 1, 2011. Available at: http://www.trans-int.org/index.php/transint/article/view/121/81 Last access on 10 June 2018.

SANDRELLI, A. Becoming an interpreter: The role of computer technology. *MonTi: Monografías de Traducción e Interpretación*, pp. 111-138, 2015. Available at: http://www.redalyc.org/pdf/2651/265146985004.pdf Last access on 24 June 2018.

SAON, G. ET AL. English Conversational Telephone Speech Recognition by Humans and Machines. *Cornell University Library*, March 2017.

Available at: https://arxiv.org/pdf/1703.02136.pdf Last access on 15 May 2018.

TRIPEPI WINTERINGHAM, S. The usefulness of ICTs in interpreting practice. In *The Interpreters' Newsletter*, v. 15, pp. 87-99, 2010. Available at: https://www.openstarts.units.it/bitstream/10077/4751/1/TripepiWinteringhamIN15.pdf Last access on 30 May 2018.

Recebido em: 06/07/2018
Aceito em: 29/11/2018
Publicado em dezembro de 2018