Interpreting and Technology:
Is the Sky Really the Limit?

Gloria Corpas Pastor[1-2]

[1] University of Malaga, Malaga 29071, Spain
[2] University of Wolverhampton, Wolverhampton WV1 1LY, United Kingdom
gcorpas@uma.es

Abstract. Nowadays there is a pressing need to develop interpreting-related technologies, with practitioners and other end-users increasingly calling for tools tailored to their needs and their new interpreting scenarios. But, at the same time, interpreting as a human activity has resisted complete automation for various reasons, such as fear, unawareness, communication complexities, lack of dedicated tools, etc.

Several computer-assisted interpreting tools and resources for interpreters have been developed, although they are rather modest in terms of the support they provide. In the same vein, and despite the pressing need to aiding in multilingual mediation, machine interpreting is still under development, with the exception of a few success stories.

This paper will present the results of VIP, a R&D project on language technologies applied to interpreting. It is the 'seed' of a family of projects on interpreting technologies which are currently being developed or have just been completed at the Research Institute of Multilingual Language Technologies (IUITLM), University of Malaga.

Keywords: Automation, Computer-assisted Tools, Interpreting Technologies

1   Introduction

Despite the actual interdependence between technology and the human skills in interpreting (Jekat, 2015), interpreting technologies are reportedly scarce (Costa, Corpas Pastor and Durán Muñoz, 2014) and have entered the profession only in recent years (Fantinuoli, 2018). Some studies suggest that interpreters are still largely unaware of them or even reluctant to use them (Corpas Pastor and Fern, 2016). Major concerns are the loss of quality and the dehumanisation of interpreting that allegedly tend to accompany technological developments (Jourdenais and Mikkelson, 2015).

However, there is a growing interest for language technologies and digital resources in the field of interpreting. See, for instance, the number of related papers presented in relevant conferences and workshops (e.g., the 8th AIIC Interpreters for In-

terpreters Workshop, 2017[1]; the two editions of HiT-IT[2], 2017, 2019; and all editions of Translating and the Computer since 2017[3]). Interpreters' attitude has also evolved in recent years. New generations of interpreters seem to be ready to embrace technology (cf. Corpas Pastor, 2018), although levels of interpreting technology uptake among interpreters remain quite low (Kerremans et al., 2019; Gaber and Corpas Pastor, 2020).

While technology changes and developments have paved the way for profound transformations in the discipline, the academic debate is just starting to address these changes, their implications and the challenges that lie ahead. Suffice to mention seminal contributions by Fantinuoli (2018a, 2018b), Mellinger and Hanson (2018), Braun (2019) or Drechsel (2019), and the papers in the volume edited by Rodríguez Melchor, Horváth and Ferguson (2020).

This paper will present a R&D project aimed at providing technology solutions for the pressing needs of both practitioners and trainees. The first part of this paper will describe the project components, and report preliminary results about users' perceptions and system performance. This first project has given rise to several complementary projects that will be also briefly covered in the second part. We will conclude with a discussion of possible new avenues of research that could impact this emerging field.

2   VIP - Voice-text Integrated system for interPreters

The VIP project seeks to provide technology solutions for the pressing needs of both practitioners and trainees. VIP integrates a suite of tools to assist interpretation at all phases, plus an open catalogue of interpreting-related technologies (tools and resources). The system architecture is described below.

2.1   Portal

The VIP portal contains a catalogue of interpreting-related tools and resources. This collection of semi-structure data is surveyed by means of a relational database management system (RDBMS). This RDBMS uses the SQL (Structured Query Language) to access the catalogue database.

The catalogue can be searched by individual categories and subcategories. For more refined searches, keywords can also be entered to locate tools and resources with specific features that appear in their description field. They are first classified by their main category and subcategories, and then, further characterised through a general feature-based template. General advanced searches query the database through the categories and subcategories and the basic feature template (platform, languages, license available) and keywords. Specific advanced searchers can be performed for Terminology management systems and Cloud-based interpreting systems.

The VIP portal is an open and collaborative database that includes language technology intended to enhance interpreter's performance, technology aimed at delivering

---

[1]   https://aiic.de/event/8-dolmetscher-fuer-dolmetscher-workshop/. (Last accessed: 05-04-2021)
[2]   http://rgcl.wlv.ac.uk/hit-it2019/. (Last accessed: 30-06-2021)
[3]   https://www.asling.org/. (Last accessed: 13-04-2021)

an interpretation remotely, technology intended to replace human interpreters, computer-assisted interpreting training tools, resources and e-learning platforms, and miscellaneous resources.

Computer-assisted interpreting (CAI) tools. The first category of CAI tools includes terminology management tools designed for and intended to be used by interpreters (Intragloss, InterpretBank, Interplex UE, Interpreters' Help, Flashterm, etc.). Besides the basic features mentioned above, specific information is also provided for each tool (documentation, available, export and import formats, author, further information). Note-taking applications have not been specifically designed with interpreters in mind, but they are increasingly being used in digital consecutive interpreting and in hybrid interpreting modalities (SimConsec and SightConsec). They are further divided into standalone software and smart pens (e.g., Evernote, Livescribe).

Speech-to-text applications (also termed S2T and ASR, automatic speech recognition) are currently being used as a central component of CAI tools, either bundled or standalone (cf. Gaber, Corpas Pastor and Omer, 2020). Only S2T standalone applications are included in VIP (e.g., Voice Dictation, Dragon NaturallySpeaking, etc.).

Remote interpreting (RI). Unlike telephone- and video-mediated interpreting, cloud-based RI usually involves two main components: (a) the Interpretation Management System, designed to schedule and manage interpreting assignments, and (b) the Interpretation Delivery Platform, designed to support the delivery of the interpretation. Some examples are Kudo and Interprefy.

Machine interpretation (MI). Although it cannot be considered interpreting-related technology proper, MI is expected to have a significant impact on professional interpreters' work environment (akin to MT in translation). They usually involve several complex language technologies in a three-phase cascade: (a) speech-to-text conversion, (b) machine translation, and (c) text-to-speech synthesis. Some examples are SpeechTrans and VoiceTra4U.

Computer-assisted interpreting training (CAIT). This category encompasses various training materials (oral resources, digitised interpretations, videos, transcribed speeches, portals, research projects, institutional multimedia repositories, etc.), and virtual training platforms (e.g., IVY, Virtual Interpreting Environment or Melissi VS).

Miscellanea. In addition, a broad category of miscellaneous resources that can aid interpreters is also included. They encompass terminology management tools used by translators that could be useful for interpreters (e.g., SDL Multiterm, TermSuite), units converters (e.g., Converto, Units), and other relevant speech technologies, like voice recording (e.g., Voice Pro, Audacity). Corpus tools appear as a separate category in the catalogue due to their increasing importance for the preparation phase of an interpreting assignment and most specifically for glossary building. They encompass tools for building, tagging, parsing and managing comparable corpora (Bootcat, SketchEngine) and multi-lingual parallel corpora (ParaConc, ParaVoz).

2.2 Modules

VIP also comprises a modular system that includes various functionalities and technologies. Initially, components were allocated modules according to their main purpose: to help interpreters prepare for an interpreting job, to assist interpreters during an interpretation, or to train interpreters. In the second phase of this project (VIP II), components have been grouped according to their complementary nature or combined functionalities (cf. Section 3).

Module I. Designed to be used in the preparation for an interpreting assignment, this module comprises four main functionalities: (i) corpus management, (ii) glossary management, (iii) named entity recognition (NER), and (iv) automatic text summarisation. Corpus management offers different functionalities related to corpora: automatic and user-assisted corpus compilation (webcrawling), uploading of corpora, and corpus query (concordances, right/left sorted KWIC, n-grams, patterns, candidate terms). Glossaries can be created from corpora or manually compiled. Dictionary and glossary management allows users to create, upload and delete glossaries, perform external searches to locate translation equivalents or, else, translate terms automatically by using machine translation and post-editing. Automatic bilingual glossary creation of multiword terms and postediting through external searches is also possible.

Named entities (NEs) can be extracted automatically (NER) by pasting a text or uploading a corpus. NEs are then retrieved according to a set of predefined categories: location (LOC), person (PER), organisation (ORG), etc., and highlighted within the text or, else, as tables. They can be also added automatically to a given glossary.

Text summarisation allows users to produce a domain survey on any topic automatically, either by uploading texts or by selecting several documents. The domain survey can be also downloaded as a corpus and managed as such. This option is particularly relevant to extract key terms for a specific topic. Key terms can be then used recursively as seed terms for automatic corpus compilation or added automatically to glossaries, etc.

Module II. Intended to be used when delivering an interpreting job, it includes: (i) automatic note-taking, (ii) machine translation and (iii) glossary query. Automatic note taking incorporates speech recognition and automatic transcription. The system detects NEs and numbers, including physical magnitudes (e.g., *25 tons*) and common nouns (e.g. *2 rockets*). Glossaries created in Module I can be searched with Glossary query. This functionality provides instant access to terms, NEs and multiword expressions (MWEs), and to their translation equivalents (either by typing the first three letters or orally through the microphone). Machine translation is provided in case an instant draft equivalent is needed on the spot. VIP includes Translate Shell and VIP translator, a neural experimental system. While automatic note-taking could be easily used in sight translation and consecutive interpretation, in simultaneous modalities it would be more convenient for the interpreter's booth mate.

Module III. Primarily designed for training student interpreters or for life-long learning purposes, it includes (i) a training module with exercises that are automatically generated, and (b) symbols for practising note-taking (experimental). Anticipation exercises enable users to practice terminology and phraseology from selected corpora.

Exercises with numbers are also customisable as regards range (e.g., from 1 to 1000), decimals, and language (Spanish, American English or British English). Based on ASR, this type of exercises enables users to practice listening/reading random numbers and then typing or saying the answers. The system indicates mistakes and provides the correct answers.

Sight-translation exercises also make use of ASR techniques, combined with parallel corpora management. The user is presented with a fragment of the source language subcorpus and produces the sight translation orally in the target language variety selected. Then, the system recognises the speech automatically, transcribes the user's spoken utterances automatically and checks the accuracy of his/her output against the aligned subcorpus 2 in the target language. The accuracy rate is approximate as it compares user's output with the actual bitext in the target language. Synonyms, term variations, syntactic transformations or paraphrases are considered errors by the system, similarly to standard translation memory systems. In addition, some errors could be due to the ASR system in place.

The fourth type of exercises in this category are intended to practice terms, multi-word terms, and multiword expressions. Glossary exercises allow selection of the glossaries, language and diatopic variety configurations, and number of exercises desired. Users can practice with one or several bilingual glossaries of their choice, either orally or by typing. The system checks and provides correct answers.

Finally, note-taking training exercises combine speech technologies and artificial intelligence for image recognition. The tool displays terms/concepts randomly (spoken, but also written if the option is selected) for users to draw the corresponding symbol. The image is then processed automatically. A checking bar indicates the percentage of accuracy of the symbol with regard to the displayed concept.

2.3   Access to VIP and users' perceptions

The VIP system is freely accessible for research purposes. User licences can be requested from the Research Results Transfer Office (OTRI) of the University of Malaga (alinares@uma.es). A beta version of the VIP system has been tested on a number of occasions by various user groups. Lack of space prevents us from providing a full account. For this reason, only two user cases have been selected. The first one studied interpreters' perceptions after using VIP, while the second example replicates the experiment with translators.

A workshop with interpreters (professionals and trainees) was organised as part of TC42 (London, 2020). In this hands-on session participants could use the tool at ease to prepare for a blind interpretation and then provide feedback. In general, the VIP systems was rated either useful or very useful as a tool to prepare for an interpretation in a 0-5 Lickert scale (3=33%; 4=33%; and 5=22%), as well as quite intuitive and user-friendly (3=22%; 4=56%, 5=22%). And to the question, "which exercises do you like most?", participants selected Glossaries (42%), followed by Numbers (33%) and Anticipation exercises (25%).

Then, individual modules and functionalities were also evaluated as regards their usefulness in the preparation phase. The average rating was 4 out of 5. The results obtained are as follows: (a) corpus compilation: 3=20%; 4=40%; 5=40%; (b) corpus management: 4=55%; 5=45%; (c) text summarisation: 3=25%; 4=75%; (d) glossary creation: 3=8%; 4=58%;5= 33%.

As part of this workshop, participants were asked to provide feedback in order to improve VIP. The most repeated suggestions mentioned were adding more languages (the present version of VIP only supports English and Spanish), quick glossary search, inclusion of abbreviations, plus some technical issues, like browser compatibility, increased website capacity and future site maintenance.

A similar survey was used to test the system among translators (professional and students) that attended a postgraduate seminar at the University of Valladolid, Spain (March 2021). Participants were asked to use VIP in a mock translation project. The results are in line with the previous survey on interpreters. Most participants rated VIP useful (4=16.67%) or very useful (83,33%) for translation, and quite intuitive and user-friendly (4=66.67%). Among the main assets of the system participants highlighted the integration of various functionalities in one single platform, its simplicity and user-friendliness, and its fast performance. As to the relevance of various functionalities as an aid to translation, their preferences were, in descending order: (a) glossary creation (3=16.67%; 4=16.67%; 5=66.67%), (b) corpus compilation and corpus query (3= 16.67%; 4=33.33%; 5=50%); and (c) text summarisation (2= 16.67%; 4=33.33%; 5=50%). When asked what functionalities they found most useful, participants unanimously mentioned automatic corpus compilation and automatic glossary creation.

As to possible ways to improve VIP, most respondents mentioned adding more languages (French, German, Italian, Russian), and more functionalities specifically designed for translation purposes (for instance, filtering by language variety and specialised domain).

Our findings suggest that VIP could be equally useful to prepare an interpretation or translation assignment, although the specific needs of both user groups might be rather different. Further studies need to be conducted on the perception and usability of our system for different tasks and mediation modalities.

3 The R&D cluster on interpreting technologies

VIP was the first research project on interpreting technologies we were granted (ref. no. FFI2016-75831-P, 2016-2020). It opened several paths to explore the impact of technology in interpreting training and research. For instance, *INTERPRETA 2.0: application of ICT tools for the teaching-learning process of interpreting* provided teaching tools and resources for undergraduate students in order to foster autonomous learning and tech-savviness (ref. no. PIE 17-015, 2017-2019, University of Malaga). Other related initiatives worth mentioning were *Training Network on Language Technologies for Interpreters* (ref. no. EUIN2017-87746, 2017-2020, Spanish Ministry of Economy and Competitiveness), and *Application of Advanced NLP Techniques to the Field of Translation and Interpreting Technologies* (Ref. PIE 17-015, 2018-2020), already completed; as well as the *European Masters in Technology for Translation and Interpreting* (Ref. 599287-EPP-1-2018-1-UK-EPPKA1-JMD-MOB, European Commission) and the *Research network INTEC: Interpreting and New Technologies* (University of Malaga, 2021-2022), both of them still on-going.

The former are just training initiatives or networks of various kinds. In this section we will provide a brief overview of some of the research and/or transfer projects on interpreting technologies conducted within our research group (PI: Prof. G Corpas).

The first two investigate the possibilities of Natural Language Processing (NLP) and neural networks for the automation of interpretation, while the other two are extensions of VIP.

3.1 Multilingual dialogue systems using neural networks for apps in the healthcare domain: the triage (Spanish-English/Arabic)

The *triage* project (ref. no. UMA18-FEDERJA-067, ERDF, 2019-2021) falls within interpreting for the public services. It follows the digitalisation and technologicalisation trends in the sector. Specifically in the field of public services, so-called remote interpreting is beginning to be introduced, which allows the service to be offered by telephone or video conference software. This modality reduces the cost of interpreting, although it still requires a high investment on human interpreters. In a healthcare context, effective communication with the patient is essential for adequate and quality care. However, since interpreting services are expensive, not all hospitals and health centres can afford to treat foreign patients in their own language. This situation is particularly complex in the case of the Andalusian health system, due to migratory movements and the influx of tourists.

The ultimate goal of this project is the development of a multilingual system to automate triage. The term 'triage' refers to the process by which people are selected based on their need for immediate medical treatment when available resources are limited. Our focus is emergency triage, as it is the scenario that requires the fastest reaction time and has the least time to resort to external interpretation services. Central to the project is the design and implementation of an app for smartphones and other mobile devices, such as iPad, to enable effective communication between the healthcare professional and the patient which will allow patients to be assessed and ordered according to the severity and urgency of the case. The system is based on multilingual dialogical models and multimodal neural automatic translation (speech-text-speech). The system allows automatic translation/interpretation for the language pairs Spanish-English and Spanish-Arabic. In addition, patient' medical records can be generated in the three languages and stored in/retrieved from the hospital database. The Agencia Sanitaria Costa del Sol and the Hospital Costal del Sol collaborate in this project.

3.2 MI4ALL - Machine Interpretation for All Through a Deep Learning API

The advances of recent years in Artificial Intelligence (AI) are making it possible to develop applications that improve people´s lives at different levels. Within AI, the most widely used technique is Deep Learning (DL). Compared to other learning methods, this technique stands out for its high performance when solving problems and addressing tasks related to language and communication. Some examples with excellent results are found in machine translation (MT), automatic speech recognition (ASR), text generation, question-answering systems, and many other areas of NLP.

The MI4ALL project (UMA-CEIATECH-04, 2020-2022) aims at providing an automatic interpretation software platform that will combine DL and corpus linguistics to facilitate public services for foreigners and immigrants, allowing them to communicate in a language in which they are fluent. The platform consists of a REST API that

offers automatic translation, speech recognition and voice synthesis functionalities that can be accessed through an app or integrated into other systems and can be retrained for other fields and languages. Multilingual and multimodal corpora are being compiled for Spanish, English, and Arabic. The data will be used to train the system and to assess its performance. A technology company (Intelligenia S.A.) is also involved in the development of the system.

3.3 Voice-text integrated system for InterPreters: Proof of Concept

Nowadays, technology permeates the translation industry The VIP project aims at improving interpreters' performance and work conditions by taking advantage of current developments and technology in a similar way. The resulting system is fit to purpose: it brings to the current state of the art a novel technological solution that can benefit interpreters and their work environment before, during and after an interpreting assignment.

VIP technology is mature enough to meet both functional and non-functional requirements and to be considered as a functional system. However, given that the system has been developed within the framework of a research project, it is not orientated towards obtaining an industrial and marketable tool. For this reason, the objective of this proof-of-concept project (ref. no. E3/04/21, 2021-2022) is to evaluate the system in terms of usability, effectiveness, security and robustness in order to establish clearly and precisely what changes and improvements the system needs to become a marketable product (end-up product). To this end, two companies have entered the project: Kudo and EL Translations.

3.4 VIP II - Multi-lingual and Multi-domain Adaptation for the Optimisation of the VIP system

The VIP II project II (ref. no. PID2020-112818GB-I00, 2021-2025) seeks to continue and extend the pioneering work carried out on the previously funded project (see Section 2 and subsections). The VIP system (version 1.0, henceforth VIP1) represents a new-generation of interpreting-related technologies that is based in interdisciplinary cutting-edge research. VIP1 has filled several major research gaps that were identified at the time of its submission, also during the project timeframe. To the best of our knowledge, it is the first open-source purpose-built integrated system designed to fulfill interpreters needs and requirement, that is intended to provide support to both interpreters and trainees.

VIP II intends to improve the system further. While internal and external evaluations of the tool have shown very good results in general, the findings also point out to the existence of areas in need of further research (e.g., ASR, multimodal corpora), functionalities that should be improved (e.g., note-taking, machine translation), new desired features, more languages, new integrations and improved functionalities, etc. Besides, it is also necessary to adapt to new interpretation scenarios in terms of technology uptake, real needs, and degree of automation in a rapidly changing world.

Our main aim is two-fold: to improve VIP1 to better accommodate the needs and requirements of interpreters (professional and trainees), and to establish the feasibility and impact of achieving automation in real interpreting scenarios. To this end, five

specific objectives have been set up: to survey interpreters needs and technology up-take to enhance their performance, in general and for specific domains; to study the technical configurations of VIP1 with a view to developing an improved system (VIP2), taking into account the findings reported and in light of the latest research; to reuse and compile multilingual data (written, multimedia, oral) to increase the number of resources integrated into the system and to enhance adaptation (multilingual and multi-domain); to perform intrinsic and extrinsic evaluation of VIP2, the latter being indicative of the impact on users; and, finally, to design user cases to establish useful-ness and benefits of VIP2 according to different interpreting modes, modalities, lan-guage-pairs, scenarios, and purposes.

4   Conclusion

By and large, interpreters seem to view technology as a key asset. Their attitude to-wards interpreter-related technology has undergone a positive development in recent years and they are willing to use technology in their daily practice. The question now is whether academia and developers are prepared to tap into interpreters' needs and provide them with the appropriate tools and resources.

At this stage, a handful of thought-provoking questions could be a good starting point. Below there is tentative list, including, but not limited to, the following: (i) Is interpreting-related technology considered friend or foe?; (ii) Will interpreting tech-nologies replace interpreters?; (iii) Are interpreting technologies a key asset? If so, for whom?; (iv) What type of technology do interpreters need?; (v) How do interpreting stakeholders interact with technology?; (vi) Are interpreting technologies the "new normality" in the sector?; (vii) What is the role of artificial intelligence (AI), auto-matic speech recognition (ASR), speech-to-text (S2T), big data, neural networks, etc. in the future of the profession?

Too many questions and almost no answers to set off on a journey into the un-known. Perhaps the sky will be the limit… or maybe not.

References

1. Braun, S.: Technology and interpreting. In O'Hagan, M. (Ed.). The Routledge Handbook of Translation and Technology. Routledge (2019).
2. Costa, H., G. Corpas Pastor, Durán Muñoz, I. Technology-assisted Interpreting. MultiLin-gual Computing 143, 25(3), 27-32 (2014).
3. Corpas Pastor, G.: Tools for Interpreters: the Challenges that Lie Ahead. Trends in Trans-lation Teaching and Learning E, 5, 157-182 (2018).
4. Corpas Pastor, G., Gaber, M: Remote interpreting in public service settings: technology, perceptions and practice. Skase Journal of Translation and Interpretation 13 (2), 58-78 (2020).

5. Drechsel, A: Technology literacy for the interpreter. In: Sawyer, D. B., Austermühl, F., Enríquez Raído, V. (eds.) The Evolving Curriculum in Interpreter and Translator Education: Stakeholder perspectives and voices, pp. 259-268. John Benjamins, Amsterdam (2019).

6. Gaber, M., Corpas Pastor, G., Omer, A.: Speech-to-Text technology as a documentation tool for interpreters: A new approach to compiling an ad hoc corpus and extracting terminology from video-recorded speeches. TRANS: Revista de Traductología, 24, 1-18 (2020).

7. Fantinuoli, C.:Computer-assisted Interpreting: Challenges and Future Perspectives. In: Corpas Pastor, G.,Durán Muñoz, I. (eds.) Trends in e-tools and resources for translators and interpreters, pp 153-174. Brill, Leiden. (2018a).

8. Fantinuoli, C.. Interpreting and technology: the upcoming technological turn. In: Fantinuoli, C. (ed.). Interpreting and Technology, pp. 1-12. Language Science Press, Berlin. (2018b)

9. Jekat, S: Machine Interpreting. In: Pöchhacker, F. (ed.) Routledge Encyclopedia of Interpreting Studies, pp. 239-242. Routledge, London and New York. (2015).

10. Jourdenais, R., Mikkelson, H.: Conclusion. In: Mikkelson, H., Jourdenais, R. (eds.) The Routledge Handbook of Interpreting, pp 447-450. Routledge, London/New York. (2015).

11. Kerremans, K., Lázaro Gutiérrez, R., Stengers, H., Cox, A., Rillof, P: Technology Use by Public Service Interpreters and Translators: The Link Between Frequency of Use and Forms of Prior Training. FITISPos International Journal 6(1), 107-122 (2019).

12. Mellinger, C. D., Hanson, T.A: Interpreter traits and the relationship with technology and visibility. Translation and Interpreting Studies 13 (3), 366-392 (2018).

13. Rodríguez Melchor, M. D., Horváth, I., Ferguson, K. (Eds.): The Role of Technology in Conference Interpreter Training. (New Trends in Translation Studies 31). Peter Lang, Berlin, Bern, Brussells, New York, Oxford, Warszawa, Wien. (2020).