

## Translating On the Go? Investigating the Potential of Multimodal Mobile Devices for Interactive Translation Dictation



Julián Zapata  
Founder & President  
InTr Technologies  
Ottawa-Gatineau, Canada

### ABSTRACT

This article provides a general overview of interactive translation dictation (ITD), an emerging translation technique that involves interacting with multimodal voice-and-touch-enabled devices such as touch-screen computers, tablets and smartphones. The author discusses the interest in integrating new techniques and technologies into the translation sector, provides a brief description of a recent experiment investigating the potential and challenges of ITD and outlines avenues for future work.

**Keywords:** interactive translation dictation, multimodal interaction, voice recognition, ubiquitous computing, mobile computing, cloud computing

### RESUM

Aquest article proveix un panorama general sobre la traducció dictada interactiva (TDI), tècnica de traducció emergent que implica interactuar amb dispositius multimodals activats amb la veu i el tacte com ara els ordinadors de pantalla tàctil, les tauletes i els telèfons intel·ligents. L'autor examina l'interès d'integrar noves tècniques i tecnologies al sector de la traducció, proveix una breu descripció d'un experiment recent que investiga el potencial i els reptes de la TDI, i conclou indicant algunes avingudes per a la recerca futura.

**Paraules clau:** traducció dictada interactiva, interacció multimodal, reconeixement de la veu, computació ubicua, computació mòbil, computació en núvol

### RESUMEN

Este artículo provee un panorama general sobre la traducción dictada interactiva (TDI), técnica de traducción emergente que implica interactuar con dispositivos multimodales activados con la voz y el tacto tales como los ordenadores de pantalla táctil, las tabletas y los teléfonos inteligentes. El autor examina el interés en integrar nuevas técnicas y tecnologías al sector de la traducción, provee una breve descripción de un experimento reciente que investiga el potencial y los retos de la TDI, y concluye indicando algunas avenidas para investigaciones futuras.

**Palabras clave:** traducción dictada interactiva, interacción multimodal, reconocimiento vocal, informática ubicua, informática móvil, informática de nube

### 1. Introduction

Not long ago, many translation professionals worldwide dictated translations instead of typing them. Just as some famous writers did in the past (Jurafsky & Martin, 2009: 285), either due to preference or health issues, some translators opted for dictation: they spoke their texts into a voice recorder for later transcription, or even directly to a transcriptionist, who transcribed the words as they were being dictated (Jiménez Ivars & Hurtado Albir, 2003).

Translation dictation was a very popular, efficient and effective working technique in the 1960s and 1970s (Gingold, 1978), but started to fade away as professional translators'



workstations experienced the massive influx of typewriters and personal computers in the early 1980s (Laroque-Divirgilio, 1980, 1981). Today, translation dictation is rather an uncommon practice because translators are able to type their translations directly into the computer, thus “saving” the extra time and money needed to transcribe their dictations or to have them transcribed by a third party. Now, with the increasing global demand for top-quality translation, which currently exceeds human translators’ typing capacity, and the changing nature of human-computer and human-information interactions, things are bound to change—yet again. The tremendous advancements in computer science, and particularly in speech technologies, multimodal interaction and ubiquitous, mobile and cloud computing, provide an excellent opportunity to bring dictation back into the profession in a new form, using new devices and applications.

This article provides a general overview of interactive translation dictation (ITD) (Zapata, 2012, 2016), an emerging translation technique that involves interacting with multimodal voice-and-touch-enabled devices such as touch-screen computers, tablets and smartphones. The article examines the interest in introducing new techniques and technologies to the translation sector, provides a brief description of a recent experiment investigating the potential and challenges of ITD and outlines avenues for future work.

The reader must know that no physical keyboard was used to prepare this paper.

## **2. Interacting with Computer Devices in a More Natural Way**

Since the early 1980s, personal computers have allowed professional translators worldwide to type translations, access and store information and exploit a number of technological aids and applications. Since the turn of the millennium, however, the translation industry seems to be surprisingly lagging behind the evolution of writing and information tools; new tools are now being designed with end users in mind. Only a handful of researchers have consulted with translators when designing and implementing applications. Most efforts have focused mainly on improving the efficiency and performance of tools, with little or no consideration for humans’ changing needs and preferences for accessing, using and communicating information.

Now, while most peripheral tasks in the translation process may be performed with the aid of web-based applications and information, the central task of the translator remains to produce texts (i.e. in written form). Until very recently, this seemed possible only by dictating the translated text to a transcriptionist, recording it for later transcription, or most commonly, by typing it on a personal computer using a keyboard and a word-processing application. For several decades, however, computer scientists have strived to demonstrate that machines can process human language for transcribing speech and performing a number of tasks. Many experiments have proven the efficiency of voice input versus keyboard input, or simply that speaking is much more efficient than typing. Many texts, including this article, have been produced entirely without a physical keyboard.

Although early works implicitly stated that voice recognition was expected to completely replace text-input devices, it was soon proposed that voice input often performed better in combination with other input modes (Pausch & Leatherby, 1991). Thus, the computer-science research community has progressively gained interest in multimodal interaction (Oviatt, 2012). Multimodal interfaces are characterized by their ability to process two or more combined user input modes (such as voice, stylus, touch, gaze and gestures) in parallel and in a coordinated manner. Examples of commercially available multimodal interfaces are tablets, smartphones and other voice-and-touch-enabled devices such as touch-screen computers. Given that the major impetus for developing multimodal interfaces has been the practical aspects of mobile use, most commercial devices have primarily been designed for this purpose. However, the interest in multimodal interaction goes well beyond mobility, as Oviatt (2012: 422–423) explains:



Ultimately, multimodal interfaces are just one part of the larger movement to establish richer communications interfaces, ones that can expand existing computational functionality and also improve support for human cognition and performance [...]. One major goal of such interfaces is to reduce cognitive load and improve communicative and ideational fluency.

These interfaces, which are becoming increasingly robust and available and allow for a more natural interaction between humans and technologies, are significantly challenging traditional keyboard-and-mouse interfaces. In truth, we are just beginning to move towards a new computing era.

### **3. Life Beyond the Keyboard**

There is life beyond the keyboard, as translation researchers realized not long after the arrival of personal computers. In the mid-1990s, the first efforts to adapt voice recognition technology to human translation were made. These developments focused on minimizing misrecognition errors by combining voice recognition and machine translation (e.g. the TransTalk project described in Dymetman et al. (1994) and Brousseau et al. (1995)). Such hybrid systems have access to the source text and use probabilistic machine translation models to improve voice recognition. Similar efforts have continued over the years, highlighting the various challenges of voice recognition/machine translation integration (Désilets, Stojanovic, Lapointe, Rose, & Reddy, 2008; Reddy & Rose, 2010; Rodriguez, Reddy, & Rose, 2012; Vidal, Casacuberta, Rodríguez, Civera, & Martínez Hinarejos, 2006), and the potential benefits of using voice input for human translation purposes (Garcia-Martinez et al., 2014; Mesa-Lao, 2014). Likewise, translation trainers and researchers have made further efforts to evaluate the performance of students and professionals when using off-the-shelf voice recognition systems for straight dictation (Dragsted & Hansen, 2009; Dragsted, Mees, & Hansen, 2011; Mees, Dragsted, Hansen, & Jakobsen, 2013); to introduce them to this technology (Romero-Fresco, 2012); and to assess and analyze professional translators' needs and opinions vis-à-vis voice recognition technology (Ciobanu, 2014; Zapata, 2012, 2016).

In summary, research on voice recognition for translation purposes has existed for many years, but has been rather hesitant due to various challenges that have impeded the success of this technology. These difficulties include the inability of voice recognition systems to flawlessly transcribe continuous speech in the 1990s (De Schaetzen, 1995: 685; Kurzweil, 2013: 72–73, 122–123), the inability of stand-alone systems to simultaneously process two or more languages and perform translation-specific tasks (Gouadec, 2002: 133) and the lack of translation dictation and voice recognition courses in translator-training programs (Benis, 2002; Gouadec, 2007: 286; Mees et al., 2013; Zapata & Quirion, in press). Overall, studies on integrating voice recognition into the translation profession have been cautiously optimistic, and both scholars and developers have emphasized the need for additional research.

### **4. From Stand-Alone Voice Recognition Systems to Multimodal Mobile Devices in Translation: A Preliminary Experiment**

Outside the translation sphere, many challenges have also been associated with stand-alone voice recognition systems. Back in the 1990s, Shneiderman (1998: 328) argued that voice recognition is “the bicycle of user interface design: it is great fun to use and has an important role, but it can carry only a light load”. Although computing has advanced significantly since the 1990s, voice recognition systems still require a great amount of computational power, i.e. the user must have a powerful computer for the voice recognition system to work properly. Likewise, training a voice recognition system is a computationally expensive process and training commercial systems can only occur on servers or clusters. That being said, trained models can be embedded in a personal computer or used on a cloud-based server via mobile devices such as tablets or smartphones (Zapata & Kirkedal, 2015).



By introducing multimodal interaction to translation, my research is intended as a first step in overcoming the current challenges and limitations of stand-alone voice recognition systems, which are generally designed for traditional keyboard-and-mouse interfaces such as personal computers. A preliminary experiment was designed that capitalized on the unparalleled robustness of commercially available voice-and-touch-enabled devices. The investigation aimed to observe and analyze the translator experience (TX) with off-the-shelf multimodal interfaces with a view to designing and developing a functional prototype of an ITD environment. TX is defined in my research as a translator's perceptions of and responses to the use or anticipated use of a product, system or service (Zapata, 2016).

A physical prototype of an ITD environment was built and tested, and compared to a keyboard-and-mouse environment. Participants in the experimental group (n=7) were asked to translate a 100-word text from English into French by interacting with the prototype exclusively through voice, touch and stylus. The physical prototype consisted of two multimodal interfaces: a 10" Android tablet and a 27" Windows touchscreen computer. Figure 1 below shows the experimental setup, hardware and software. The model of each device seen in the figure is the actual model used in the experiment.

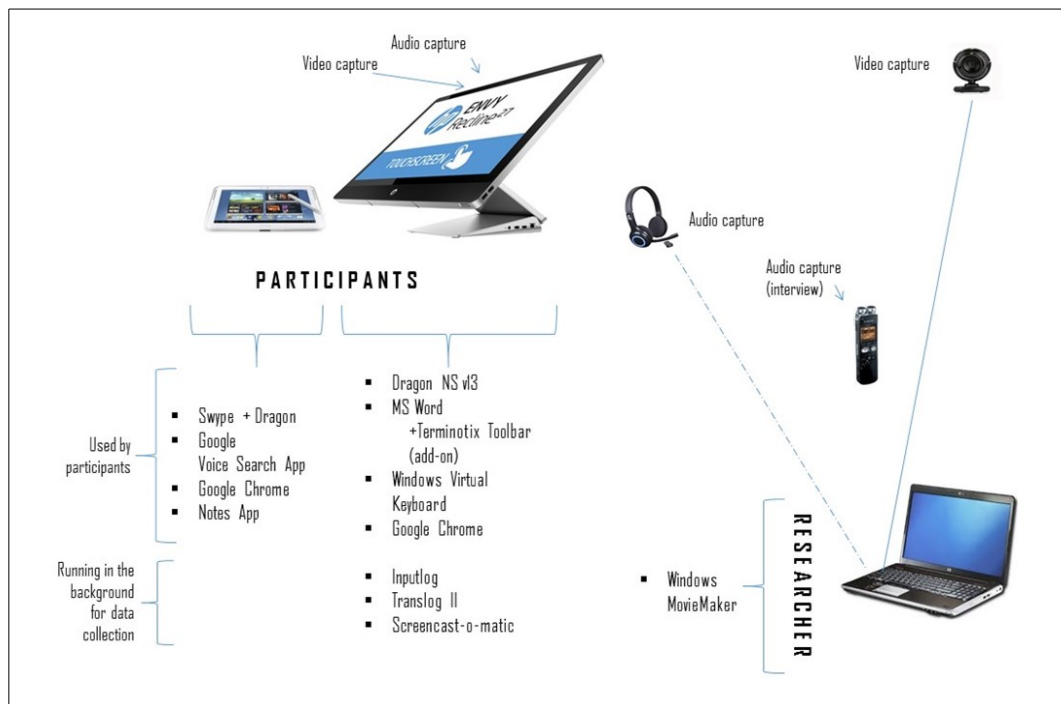


Figure 1. Experimental setup, hardware and software

A tablet was included in this experiment for three main reasons: (1) it provides the user with a different approach to VR technology (i.e. beyond device-embedded systems; see Zapata & Kirkedal (2015)) and virtual keyboard input, as I will describe below; (2) it allows the user to reflect upon the advantages and challenges of using (smaller) mobile MIs and cloud-based applications to carry out translation tasks; and (3) it is an external platform for the user to carry out any research in any source on the web, or store any information as necessary. The tablet was equipped with a cloud-based speaker-independent voice recognition system integrated in a fast-typing virtual keyboard application (Swype by Nuance). Participants used the tablet to perform searches online (i.e. by manually using Google Search with the Swype keyboard or using the voice search feature) and to access online tools, such as terminology databases, dictionaries and corpora, while preparing and/or revising their translation.



The touchscreen computer was equipped with a text-processing application (Microsoft Word 2013), with a special toolbar installed (an add-on for Word developed by Terminotix Inc. to allow users to access information resources directly from the Microsoft Word interface); a state-of-the-art commercial speaker-adapted voice recognition system (Dragon NaturallySpeaking 13 by Nuance); and the Windows 8 software keyboard. Participants used this interface to read/prepare the source text and to dictate and/or type (with the software keyboard) the target text and make corrections or changes. In addition, by means of voice shortcuts, participants were able to perform searches in two information resources directly from the source document in Word using the Terminotix toolbar. All participants in the experimental group used both devices (tablet and touchscreen computer) during the translation task.

A control group was also recruited. Participants in the control group (n=7) translated the same text using conventional input modes (i.e., a keyboard and a mouse on a desktop computer) and were instructed to translate the text “as they normally would” at work.

Both quantitative and qualitative data were collected. This experiment proposed an unprecedented translator-computer interaction scenario and a ground-breaking data-collection approach combining active observation of users with video recording using two camera feeds, screen recording using Screencast-O-Matic, input logging using Translog II (Carl, 2012) and InputLog (Leijten & Van Waes, 2013), and semi-structured interviews. The sample, as well as the tools and methods used, were sufficient for the purposes of this preliminary study, which was primarily to build a foundation for future research in the field. It also aimed to collect initial user data to design and develop a functional ITD environment prototype. A longer-term study with a larger sample and a bolder methodological design is currently being undertaken.

## **5. A Sneak Peek at the Translator Experience with Multimodal Interfaces**

The initial study was performed with the following research question in mind: Does the physical prototype of an ITD environment consistently provide translators with a better TX than a traditional keyboard-and-mouse environment, when preparing, producing and revising a translation? Presenting the entire methodology and results of the experiment outlined above is well beyond the scope of this article. The reader is referred to Zapata (2016, Chapters 5 and 6) for more details. Nonetheless, it must be said here that, based on the experimental results, multimodal interaction and ubiquitous, mobile and cloud computing appear to be promising avenues for translation technology research.

As mentioned above, both quantitative and qualitative data were collected. In addition to quantitative data, qualitative data from real users, collected in real-life or simulated scenarios, can indeed help tool designers and developers in making informed decisions about various aspects of the user interface of existing and new tools. Quantitative data were analysed first. As an example for the purposes of this article, looking at averages alone, the following could be concluded from the first data-analysis step: when considering total translation times, participants in the control group performed better than those in the experimental group; they took 25% less time to complete the task with the keyboard-and-mouse environment than participants with the ITD environment. However, only a minor difference in terms of objective quality could be observed (the translations were assessed by three expert evaluators). Furthermore, the different input process measures provided by the InputLog software indicated that the technical effort needed to produce a translation with a traditional keyboard-and-mouse environment always exceeds the technical effort needed with the prototypical ITD environment (in terms of number of characters typed, which includes spaces and deletions, number of switches between keyboard and mouse, time spent typing in active writing mode, time spent with the mouse, number of switches between windows, etc.).

Now, are these data enough to affirm that the prototypical ITD environment provides (or not) a better TX than the keyboard-and-mouse environment? While most of the objective measures



seemed to point towards a better quality-in-use of the experimental environment, the need to complement such data with qualitative data seemed to be necessary. As Swallow et al. (2005) put it, “qualitative data provides a richness and detail that may be absent from quantitative measures.” Indeed, the qualitative data from the interviews informed us about the participants’ perceptions of and responses to the use of a traditional keyboard-and-mouse workstation versus the prototypical ITD environment, and the anticipated use of an optimized version of the ITD environment. Although no generalizations could be formulated based on my sample, the multimodal environment appeared to provide translators with a better TX than the traditional one, despite the occasional cases of lack of enthusiasm, the rare negative comments and the expressions of task difficulty by some participants.

In sum, the combined analysis of objective and subjective usability measures collected in my preliminary experiment *does* suggest that the overall TX was better in the experimental environment than with a traditional keyboard-and-mouse interface. In other words, the results of my study indicated that translators generally perceived and responded to ITD more favourably. Now, in order for ITD to be fully integrated to the translation profession, much work must be done.

## 6. Looking Ahead: Some Avenues for Future Work

Moving forward, translation technology research must centre on human translators and their needs. New realities lead to new tools and new contexts of use (e.g. home technology, ubiquitous computing, cloud computing, online learning). In turn, these new contexts require new measures and new data collection and analysis approaches to adequately capture what is considered important for particular contexts and users. The next challenge in translation technology is designing for different interaction channels (i.e. different types of devices and user interfaces), different users and different contexts of use. To achieve this, researchers must follow the conceptual design path (Parush, 2015).

Ubiquitous, mobile and cloud computing are changing the way humans interact with computers and information. My preliminary experiment was performed using a reclining all-in-one desktop computer and a tablet, with both devices connected to a wireless Internet network. Apart from the computer’s power adapter, no wires were involved in the experimental setup. No headset was needed for voice recognition on any device. In addition, even though the desktop computer’s voice recognition system was installed on the device, this technology is already moving into the cloud and boasts comparable accuracy to that of embedded systems (Zapata & Kirkedal, 2015). Thus, a fully wireless-and-mobile experiment similar to the one outlined in this paper is already possible.

Future investigations will also need to consider the integration of translation memories and machine translation technology into an ITD environment. These technologies may indeed augment the workflow and improve the efficiency of the voice recognition by using a hybrid approach similar to those explored in the 1990s and 2000s, cited in section 3 above. Lastly, translation studies scholars will need to explore the pedagogical potential and challenges of ITD. One of the challenges identified in previous work is the lack of formal training in oral translation techniques and in the use of emerging technologies such as VR and mobile devices (Mees et al., 2013: 152; Zapata & Quirion, in press).

## 7. Conclusion

Professional translators—and humans in general—are increasingly dependent on computer tools and devices. In the age of mobile and cloud computing and ubiquitous information, research on speech technologies and multimodal interaction will become increasingly important in translation. Experiments that explore voice, touch and stylus input (and even other emerging interaction modes such as gaze, gesture and brain input) will play a crucial role in the design and development of new user-friendly tools and devices that are



adapted to translators' needs and to the changing reality of the industry in the twenty-first century. Shneiderman (2015: ix) affirms that:

[t]echnology designers who shape user experiences seek to smooth the path for novices and serve the demanding needs of experts. This was true for fifteenth-century book designers, nineteenth-century train designers, and twenty-first-century smartphone designers. Their innovative designs emerged from a deep empathy for people, sensitivity to diverse social contexts, and imaginative sparks to create new ways of thinking about technology.

This is precisely the aim of my research: to create new ways of thinking about translation technology. Perhaps the day in which translators worldwide will translate on the go, using mobile and wearable devices anytime and anywhere, is just around the corner.

## References

- Benis, M. (2002). Softly spoken or hard of hearing? *Language International*, 14(3), 26–29.
- Brousseau, J., Drouin, C., Foster, G., Isabelle, P., Kuhn, R., Normandin, Y., & Plamondon, P. (1995). French speech recognition in an automatic dictation system for translators: The TransTalk project. In *Proceedings of Eurospeech'95*. Retrieved from <http://www.iro.umontreal.ca/~foster/papers/ttalk-eurospeech95.pdf>
- Carl, M. (2012). Translog - II: a Program for Recording User Activity Data for Empirical Reading and Writing Research. In *Proceedings of the Eight International Conference on Language Resources and Evaluation* (pp. 4108–4112). Istanbul, May 21–27, 2012: European Language Resources Association.
- Ciobanu, D. (2014). Of Dragons and Speech Recognition Wizards and Apprentices. *Revista Tradumàtica*, (12), 524–538.
- De Schaezen, C. (1995). Bureautique et télé-bureautique en traduction. *Meta*, 40(4), 682–693. Retrieved from <http://www.erudit.org/revue/meta/1995/v40/n4/004040ar.pdf>
- Désilets, A., Stojanovic, M., Lapointe, J.-F., Rose, R., & Reddy, A. (2008). Evaluating Productivity Gains of Hybrid ASR-MT Systems for Translation Dictation. Retrieved from <http://www.mt-archive.info/IWSLT-2008-Desilets.pdf>
- Dragsted, B., & Hansen, G. (2009). Exploring Translation and Interpreting Hybrids. The Case of Sight Translation. *Meta*, 54(3), 588–604. Retrieved from <http://www.erudit.org/revue/meta/2009/v54/n3/038317ar.pdf>
- Dragsted, B., Mees, I. M., & Hansen, I. G. (2011). Speaking your translation: students' first encounter with speech recognition technology. *Translation & Interpreting*, 3(1), 10–43. Retrieved from <http://www.trans-int.org/index.php/transint/article/viewFile/115/87>
- Dymetman, M., Brousseau, J., Foster, G., Isabelle, P., Normandin, Y., & Plamondon, P. (1994). Towards an Automatic Dictation System for Translators: the TransTalk Project. In *Fourth European Conference on Speech Communication and Technology* (p. 4). Retrieved from <http://arxiv.org/abs/cmp-lg/9409012>
- Garcia-Martinez, M., Singla, K., Tammewar, A., Mesa-Lao, B., Thakur, A., M. A., A., ... Bangalore, S. (2014). SEECAT: ASR & Eye-tracking Enabled Computer-Assisted Translation. In *Proceedings of the 17th Annual Conference of the European Association for Machine Translation* (pp. 81–88).
- Gingold, K. (1978). The Use of Dictation Equipment in Translation. In P. A. Horguelin (Ed.), *La traduction, une profession. Actes du VIIIe congrès mondial de la Fédération internationale des traducteurs* (pp. 444–448). Ottawa: Conseil des traducteurs et interprètes du Canada.
- Gouadec, D. (2002). *Profession : traducteur*. Paris: La Maison du dictionnaire.



- Gouadec, D. (2007). *Translation as a profession*. Amsterdam: John Benjamins.
- Jiménez Ivars, A., & Hurtado Albir, A. (2003). Variedades de traducción a la vista. Definición y clasificación. *Trans Revista de Traductología*, 7, 47–57. Retrieved from [http://www.trans.uma.es/trans\\_07.html](http://www.trans.uma.es/trans_07.html)
- Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Kurzweil, R. (2013). *How to Create a Mind. The Secret of Human Thought Revealed*. New York: Penguin.
- Laroque-Divirgilio, L. (1980). Sondage sur les services de traduction dans l'entreprise. *Meta*, 25(2), 276.
- Laroque-Divirgilio, L. (1981). La traduction au magnétophone. *Meta*, 26(4), 398–403. Retrieved from <http://www.erudit.org/revue/meta/1981/v26/n4/002573ar.pdf>
- Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. *Written Communication*, 30(3), 358–392. Retrieved from <http://wvx.sagepub.com/cgi/doi/10.1177/0741088313491692>
- Mees, I. M., Dragsted, B., Hansen, I. G., & Jakobsen, A. L. (2013). Sound effects in translation. *Target*, 25(1), 140–154. Retrieved from <http://openurl.ingenta.com/content/xref?genre=article&issn=0924-1884&volume=25&issue=1&spage=140>
- Mesa-Lao, B. (2014). Speech-Enabled Computer-Aided Translation: A Satisfaction Survey with Post-Editor Trainees. In *Workshop on Humans and Computer-assisted Translation* (pp. 99–103).
- Oviatt, S. (2012). Multimodal Interfaces. In J. A. Jacko (Ed.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (3rd ed., pp. 415–429). New York: Lawrence Erlbaum Associates.
- Parush, A. (2015). *Conceptual Design for Interactive Systems: Designing for Performance and User Experience*. Waltham: Morgan Kaufman.
- Pausch, R., & Leatherby, J. H. (1991). An Empirical Study: Adding Voice Input to a Graphical Editor. *Journal of the American Voice Input/Output Society*, 9(2), 55–66.
- Reddy, A., & Rose, R. C. (2010). Integration of Statistical Models for Dictation of Document Translations in a Machine Aided Human Translation Task. *IEEE Transactions on Audio, Speech and Language Processing*, 18(8), 1–11. Retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05393062>
- Rodriguez, L., Reddy, A., & Rose, R. (2012). Efficient Integration of Translation and Speech Models in Dictation Based Machine Aided Human Translation. In *Proceedings of the IEEE 2012 International Conference on Acoustics, Speech, and Signal Processing* (Vol. 2, pp. 4949–4952).
- Romero-Fresco, P. (2012). Respeaking in Translator Training Curricula. Present and Future Prospects. *ITT*, 6(1), 91–112.
- Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Reading, Massachusetts: Addison-Wesley.
- Shneiderman, B. (2015). Foreword. In A. Parush (Ed.), *Conceptual Design for Interactive Systems: Designing for Performance and User Experience* (pp. ix–x). Waltham: Morgan Kaufman.
- Swallow, D., Blythe, M., & Wright, P. (2005). Grounding experience: relating theory and





- 
- method to evaluate the user experience of smartphones. In *Proceedings of the Annual conference on European association of cognitive ergonomics* (pp. 91–98).
- Vidal, E., Casacuberta, F., Rodríguez, L., Civera, J., & Martínez Hinarejos, C. D. (2006). Computer-assisted translation using speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 14(3). Retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01621206>
- Zapata, J. (2012). *Traduction dictée interactive : intégrer la reconnaissance vocale à l'enseignement et à la pratique de la traduction professionnelle*. University of Ottawa. Retrieved from <http://www.ruor.uottawa.ca/handle/10393/23227>
- Zapata, J. (2016). *Translators in the Loop: Observing and Analyzing the Translator Experience with Multimodal Interfaces for Interactive Translation Dictation Environment Design*. University of Ottawa. Retrieved from <http://www.ruor.uottawa.ca/handle/10393/34978>
- Zapata, J., & Kirkedal, A. S. (2015). Assessing the Performance of Automatic Speech Recognition Systems When Used by Native and Non-Native Speakers of Three Major Languages in Dictation Workflows. In *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)* (pp. 201–210).
- Zapata, J., & Quirion, J. (in press). La traduction dictée interactive et sa nécessaire intégration à la formation des traducteurs. To appear in *Babel*, 62(4).