

**This is a post-print version of the review published in Target 28(2),
2016, pages 337-340.**

Pablo Romero-Fresco. *Subtitling Through Speech Recognition: Respeaking* (Translation Practices Explained, 13). Manchester: St. Jerome, 2011. 194 pp.

Reviewed by Anna Matamala (Universitat Autònoma de Barcelona)

In our digitized society, audiovisual content is ubiquitous, but not everybody can fully understand and enjoy content transmitted simultaneously through the visual and the audio channels. Access services such as subtitling cater for the needs of those who do not understand or cannot access the audio channel. Various types of subtitling can be found, but especially challenging are subtitles that are produced live for an audience that cannot hear the audio. In this area different techniques exist but undoubtedly the one that is being increasingly used is respeaking, that is, the creation of subtitles on live content through speech recognition. This book by Pablo Romero-Fresco provides a comprehensive overview of this practice and one can say without doubt that only five years after its publication it has already become a classic of audiovisual translation.

The book begins with an introductory chapter in which the author provides a comprehensive definition of the term ‘respeaking,’ highlights the key aspects of this technique, and discusses terminological variation. An overview of the origins of subtitling for the deaf and hard-of-hearing and live subtitling, both in Europe and in the United States, is provided in the second chapter, where a description of regulatory practices is made. This second chapter also presents a taxonomy of subtitles according to various parameters: production approach (live, semi-live, pre-recorded), language (intralingual, interlingual), correction method (no correction, self-correction, parallel correction), editing policy (verbatim, reduced), display mode (blocks, scrolling), and presence of subtitling for the deaf and hard-of-hearing features (none, character identification, sound information). Romero-Fresco also compares the different methods to produce the subtitles, namely, QWERTY keyboards, Velotype (a syllabic keyboard), dual keyboard, stenotype, and speech recognition.

The volume continues with an analysis of respeaking on television and respeaking training in a wide array of countries (UK, Flanders, Switzerland, Denmark, France, Italy, Canada, US), offering the reader an international picture of this practice at the time the book was published, a scenario that would need updating today but which nonetheless remains

useful in acknowledging a key moment in the development of respeaking. The focus in this chapter is on television, whilst later, in chapter 9, other settings such as museums, arts venues, conferences, churches, live webcasts, telephones and schools are described.

Chapter 4 maps the skills required from respeakers, pointing to those that are shared with other translating practices (interpreting, subtitling) and those that are specific to respeaking. This mapping was much needed to acknowledge respeaking as a professional practice with its own specificities. A summary of these skills is included in a very illustrative final table, even if the graphics could be improved. The following chapters take a closer look at the various individual skills needed before, during and after the process. These include software-related skills, multitasking skills, delivery skills and assessment skills, to name just a few.

Chapter 5 explains how speech recognition systems work, thankfully bearing in mind that readers interested in respeaking are not necessarily expert engineers. It also describes how speech recognition systems interact with subtitling software. Chapter 6 deals in more depth with the process of preparing the software, focusing on Dragon Naturally Speaking 10. In this rapidly developing technological area, the market has experienced changes and new versions have been released. Yet, the basics remain the same and the information included in the book is still useful for anyone interested in the topic.

Chapters 7 and 8 are central for those willing to respeak because it is here that Romero-Fresco presents – in a very clear style that does justice to the title of the series (“Translation Practices Explained”) – what is expected from a respeaker. Special attention is paid to the fact that respeakers carry out various simultaneous tasks, and specific instructions are given on how to deal with punctuation, rhythm or speed, taking into account the various genres involved (sports, news, debates, chat shows, etc.).

The book provides a wealth of real-life examples to illustrate the author’s statements, and does not shy away from discussions in which different stakeholders hold opposing views. This is the case of the speed of subtitles (number of words per minute) and their presentation (verbatim versus edited). Romero-Fresco presents the diverging approaches to the issue, reviews the existing literature, and clarifies the terminology, as well as providing research results. This is especially relevant at the end of the book, in chapter 11, where he examines the reception of respoken subtitles. Scientific results from research projects using various methodological tools, from questionnaires to eye-trackers, allow him to discuss and even challenge existing practices. Romero-Fresco is interested in how viewers understand respoken content at different speeds, how they perceive scrolling versus block subtitles, and what they

think about subtitle quality. For instance, the author demonstrates that scrolling word-for-word subtitles cause very chaotic reading patterns which impact negatively on comprehension, which is a reason for favouring block subtitles. This type of research is of the utmost importance, especially in a scenario where statements in favour of certain speeds or presentation formats are too often made by interested stakeholders without taking into account previous investigations. At this point one could argue that Romero-Fresco fails to provide a statistical analysis of the data proving their significance. But then, strictly speaking this volume is not intended as a research paper but is “first of all a textbook for trainers and students wishing to master the fundamental aspects of respeaking, both from a theoretical and a practical point of view” (xix). As such it perfectly fulfils its function, providing not only a thorough description of the technique but also additional food for thought based on sound research results.

An important aspect that the author also addresses in chapter 10 is how to measure the accuracy rate in subtitling. The book presents the NERD model, a proposal that departs from the principles of Word Error Rate (WER) metrics and takes into account the specificities of respeaking. Clear examples of the application of this model are presented to illustrate its implementation. Again, the basics are included in the book, although this model has already been developed by the same author (Romero-Fresco and Martínez 2011) into more advanced metrics: the NER model. A specific tool has been designed (NERstar tool, see <http://www.speedchill.com/nerstar/>) and the system is even promoted by OfCom, the communications regulator in the UK, which shows the impact in society of Romero-Fresco’s work beyond the academic sphere.

All in all, this book should unquestionably be on the compulsory reading list of any audiovisual translation course. It establishes the foundations of respeaking with a multi-faceted approach. Foregrounding its didactic purpose, it includes a series of discussion points and exercises at the end of each chapter, often linked to video clips and illustrative materials available on a DVD that accompanies the book. It is worth stressing that these materials are not only provided in English but also in other languages, which is in tune with the book’s international approach. The only criticism that could be made is that, as the author already acknowledges in his final thoughts, “any attempt to capture the present state of affairs is bound to become obsolete soon” (177), especially in a field where technology plays such a significant role. Five years after its publication, further advances have been made: European projects have investigated respeaking (SAVAS, 2012-2014), providing new research results; the international “Respeaking, Live Subtitling and Accessibility” symposia have offered an

active forum to exchange research in the field, and new and complementary approaches to the topic have been published (Eugeni 2012, Van Waes et al. 2013, Remael et al. 2014). The interaction of speech recognition with other technologies such as machine translation, the application of respeaking in other fields beyond live subtitling, as well as the development of interlingual respeaking are three hot topics on the current research agenda. However, the book is still a must-read for all audiovisual translation professionals, scholars and students and a seminal reference for anyone interested in the fascinating field of respeaking. Let's hope the author can continue his active research and offer us a second edition very soon.

References

- Eugeni, Carlo. 2012. "Measuring Audiovisual Translation. A Model for the Analysis of Intralingual Live Subtitles." *US-China Foreign Language* 10 (6): 1276-1286.
- Remael, Aline, Luuk Van Waes, and Mariëlle Leijten. 2014. "Live Subtitling with Speech Recognition: How to Pinpoint the Challenges." In *Media and Translation: An Interdisciplinary Approach*, ed. by Dror Abend-David, 131-148. New York: Bloomsbury Academic.
- Romero-Fresco, Pablo, and Juan Martínez. 2011. "Accuracy Rate in Live Subtitling – the NER Model." Working Paper. Accessed October 7, 2015.
<http://roehampton.openrepository.com/roehampton/handle/10142/141892>
- Van Waes, Luuk, Mariëlle Leijten, and Aline Remael. 2013. "Live Subtitling with Speech Causes and Consequences of Text Reduction." *Across Languages and Cultures* 14 (1): 15-46.

Reviewer's address

Anna Matamala
Universitat Autònoma de Barcelona
Campus UAB, K-1002
08193 BELLATERRA (Barcelona)
Spain

anna.matamala@uab.cat