

# Preface

The idea of editing a volume entirely focused on corpus-based interpreting studies was first discussed following the *First Forlì International Workshop on Corpus-based Interpreting Studies: The State of the Art*, which was held at the Forlì Campus of the University of Bologna on May 7–8, 2015. This event gathered more than 100 scholars from different parts of the world with the aim of sharing their corpus-based research endeavors, ranging from studies that exploited fully machine-readable corpora to small collections of texts or transcripts for manual analysis. The workshop came after other occasions in which translation and interpreting scholars presented their corpus-based research projects, though to differing extents. In 2010, a full session on “Interpreting Corpora” was included in the program of the *Emerging Topics in Translation and Interpreting—Nuovi percorsi in traduzione e interpretazione* conference organized by the University of Trieste (Straniero Sergio and Falbo 2012). The 2003 conference held in Pretoria *Corpus-based Translation Studies: Research and Applications* was probably the first one with an exclusive focus on corpus-based research in translation, though no contributions from interpreting scholars were presented. The volume that originated from that conference (Kruger et al. 2011) does include a chapter on CIS anyway, based on a paper presented at the 2006 IATIS conference (Bendazzoli et al. 2011). Going backward in time, it is clear that corpus-based interpreting studies, as an “offshoot” of corpus-based translation studies have flourished considerably and can be expected to develop even further.

This volume serves a dual purpose. On one hand, it aims at promoting the understanding of the interpretation process and product based not on anecdotal observations or small-size case-studies, but on comparatively large datasets of professional interpretations mostly stored and queried according to standard corpus linguistics methodologies. The volume showcases descriptions of and studies on major interpreting corpora available to date: the EPIC Corpus and its off-springs EPTIC (including also translations) developed at the University of Bologna, EPICG from the University of Ghent (Belgium) and the TIC Corpus from the University of Poznań (Poland); the 2249i Corpus, the DIRSI Corpus and the IMITES Corpus, again from the University of Bologna (Italy); the CorIT Corpus from the University

of Trieste (Italy); the FOOTIE Corpus created at UNINT University in Rome (Italy); the NAIST Corpus from the Nara Institute of Science and Technology (Japan) and the CEIPPC Corpus, which was built at the Guangdong University of Foreign Studies (China). On the other hand, the volume is also intended as a renewed call (after Miriam Shlesinger's first call in 1998) to the research community to further develop the field of corpus-based interpreting studies by offering scholars more corpus-based data and methodologies to compile their own corpora according to their research designs.

This volume consists of 11 chapters. The first two are meant to provide the theoretical framework of corpus-based interpreting studies (CIS), focusing on their development and on procedural-methodological issues respectively, while the other nine chapters present the insightful results obtained by analyzing professional interpreters' performances from this promising corpus-based perspective.

What follows is an overview of the contributions to the volume.

The book opens with the chapter "Corpus-based Interpreting Studies: Past, Present and Future Developments of a (wired) Cottage Industry" by Claudio Bendazzoli, who takes stock of nearly two decades of corpus-based studies all the way up to Web 2.0 applications and looks at the extent to which they have differently flourished in three areas of interpreting, namely research, education, and professional practice.

This is followed by a collective chapter "Building Interpreting and Intermodal Corpora: A *How-to* for a Formidable Task" by Silvia Bernardini, Adriano Ferraresi, Mariachiara Russo, Camille Collard, and Bart Defrancq, who pooled their pioneering experiences to provide an accessible step-by-step guide for corpus developers, especially those who are working with European Parliament (EP) data, and an appraisal of available technologies to cater for different research questions. The ultimate goal is to harmonize procedures in order to expand EP interpreting and multimodal corpora through a collective effort.

An example of corpus-based study is offered by Bart Defrancq and Koen Plevoets' "Over-uh-load, Filled Pauses in Compounds as a Signal of Cognitive Load", which opens the series of the chapters investigating interpreter strategies during simultaneous interpreting (SI). Based on their EPICG Corpus, a French-Dutch-English-Spanish corpus, the authors contrast SI data in Dutch with non-mediated Dutch linguistic production to ascertain the increased cognitive load associated with the production of compound lexemes between languages with compound parts in reverse order *versus* the source language (French).

A similar research question was investigated by Binhua Wang and Bing Zou in "Exploring Language Specificity as a Variable in Chinese-English Interpreting. A Corpus-based Investigation". Based on the Chinese-English Interpreting for Premier Press Conferences Corpus (CEIPPC), the authors studied the cognitive load associated with interpreting in the consecutive mode between two languages with major differences in cultural conceptualizations and linguistic structures, Chinese and English. In particular, they focus on the processing of the following asymmetry:

attributive modifying structures which are typically front-loaded in Chinese and modifying structures which are typically back-loaded in English.

The language of the professional interpreter or *interpretese* is the subject of study of the two following chapters. Guy Aston in his contribution “Acquiring the Language of Interpreters: A Corpus-based Approach” discusses the value of memorized formulae to produce fluent speech and, based on his 2249i Corpus, stresses the potential of corpora to detect fixed expressions in professional simultaneous interpreters and store them for the benefit of trainee interpreters. Marta Kajzer-Wietrzny’s chapter “Interpretese vs. Non-native Language Use: The Case of Optional *That*” describes the functions of the optional complementizer *that* and compares its use and the *zero* variant in interpreted, non-native and native English discourse at the European Parliament, collected in her TIC Corpus, in order to detect the prevailing linguistic patterns among simultaneous interpreters.

A novel field of research in corpus-based interpreting studies concerns gender. Mariachiara Russo’s chapter “Speaking Patterns and Gender in the European Parliament Interpreting Corpus” investigates a number of parameters, i.e. speaker’s mode of delivery, input speed, language combination, and topic, in relation to target speech length revealing some statistically significant differences among female and male interpreters.

The subsequent three chapters analyze simultaneous interpreters’ strategies when dealing with very challenging speech acts: the use of figurative language by the speaker and fast adversarial exchanges during a political debate and a press conference.

In her contribution “Studying Figurative Language in Simultaneous Interpreting: The IMITES (*Interpretación de la Metáfora entre Italiano y Español*) Corpus”, Nicoletta Spinolo classifies the linguistic behavior of interpreters faced with 1135 figurative expressions and identifies the nature of those most difficult to translate.

Eugenia Dal Fovo’s study “European Union Politics Interpreted on Screen: A Corpus-based Investigation on the Interpretation of the Third 2014 EU Presidential Debate” is based on her EUDEB14 Corpus, a subcorpus of CorIT, the world largest TV interpreting corpus developed by our late colleague and friend Francesco Straniero Sergio at the University of Trieste. The author contrasts interpreting norms and ethics between SCIC interpreters and free-lance interpreters for the same event, the debate among candidates to the Presidency of the EU Commission, to ascertain to what extent the composition of the interpreting team and the equipment influence the representation of the interaction. Her results reveal that the freelance interpreters, who were TV interpreting experts displayed a more telegenic style, in keeping with the spectacularization principle and the confrontational dynamics of televised political debates.

Interactional dynamics were also analyzed by Annalisa Sandrelli in a totally different setting. In her chapter, “Interpreter-mediated Football Press Conferences: A Study on the Questioning and Answering Strategies”, she observes how the target language versions closely mirrored the source language Q&A functions, despite the

fast changes in turns at talk, overlapping speech and the psychological pressure on the interviewees.

The eleventh and last chapter adds another perspective to corpus-based interpreting studies, that of technologists Graham Neubig, Hiroaki Shimizu, Sakriani Sakti, Satoshi Nakamura, and Tomoki Toda who are interested in understanding the difficulties faced by human interpreters and the possibilities of creating systems that help interpreters overcome these difficulties through the creation of assistance tools or speech translation (ST) technology. Based on the NAIST Japanese-English Corpus, they compare the interpreted output of interpreters of varying degrees of experience with the translated output of the same source speeches. Their chapter “The NAIST Simultaneous Translation Corpus” describes the collection of source language materials, interpretation processes, recording, and transcript of resulting data.

We hope that the richness of approaches and results offered by the present volume may inspire other scholars to join efforts and resources with a view to expand interpreting corpora and validate interpreting hypotheses on larger datasets. Likewise, we hope that also trainee interpreters may benefit from being exposed to such a wide range and abundance of professional interpreting styles and successful strategies.

Forlì, Italy  
Torino, Italy  
Ghent, Belgium

Mariachiara Russo  
Claudio Bendazzoli  
Bart Defrancq

## References

- Kruger, Alet, Kim Wallmach, and Jeremy Munday. eds. 2011. *Corpus-based Translation Studies: Research and applications*, London/New York, Continuum.
- Shlesinger, Miriam 1998. Corpus-based Interpreting Studies as an offshoot of corpus-based Translation Studies. *Meta* 43(4): 486–493.
- Straniero Sergio, Francesco, and Caterina Falbo. 2012. *Breaking ground in corpus-based Interpreting Studies*. Bern: Peter Lang.

Making Way in Corpus-based Interpreting Studies

Russo, M.; Bendazzoli, C.; Defrancq, B. (Eds.)

2018, XVI, 215 p. 35 illus., 17 illus. in color., Hardcover

ISBN: 978-981-10-6198-1