

Arias-Badia, B. (2020) *Subtitling Television Series. A Corpus-Driven Study of Police Procedurals*. Oxford: Peter Lang.
244 pp. ISBN 978-1-78707-796-6

Review

Olaia Andaluz-Pinedo¹
University of the Basque Country UPV/EHU

Received: 23/03/2020

Accepted: 25/03/2020



The volume *Subtitling Television Series. A Corpus-Driven Study of Police Procedurals* by Dr Blanca Arias-Badia (Universitat Autònoma de Barcelona) reports on a study of linguistic features present in a corpus of English scripted dialogue and Spanish subtitling compiled *ad hoc*, the Corpus of Police Procedurals (CoPP). More specifically, this work, framed within descriptive Audiovisual Translation (AVT), focuses on quantitative and qualitative analyses of syntactical and lexical features in the original script as well as in the subtitled version of the TV series which make up the corpus: *Castle*, *Dexter* and *The Mentalist*. By doing so, it delves into the relevant yet underexplored area of fictive orality in subtitling through an innovative methodology in this field, the corpus-driven approach. It is a refreshing contribution which will undoubtedly be of interest for Translation Studies and Corpus Linguistics researchers.

This publication's clear structure smoothly guides the reader through its 9 chapters. Chapter 1 introduces the corpus-driven approach, the aim and research questions and the contents organisation. Regarding methodology, the use of a corpus approach for this descriptive translation study, in line with latest advances in the discipline, is completely coherent for the empirical study of the phenomena in question. In particular, the adoption of a corpus-driven methodology is proposed as a novelty. Following this approach, the aim and research questions, related to a description of the dialogue and subtitling of the

¹ At the time of writing this review, the author is a doctoral student funded by the University of the Basque Country UPV/EHU, Spain (PIF17/46).

CoPP, are formulated generally, and will be further specified in the course of the research.

After the introduction, chapters 2 and 3 provide a theoretical framework by focusing on key notions for the study. Both form and content of these parts are interesting: the way they are presented shows a selection of essential aspects oriented to the analysis, and their revision includes fundamental and up-to-date references. On the one hand, chapter 2 highlights the vital concept of *norm*. The interdisciplinary use of this notion in the three areas on which the study draws, Translation Studies, Linguistics and Television Studies, is put forward as a common thread for the research conducted. Moreover, consistently with the main frame of this work, that is Descriptive Translation Studies, it sets out to identify and describe these recurrent patterns of behaviour in the corpus' scripted dialogues and subtitles. On the other hand, chapter 3 examines relevant literature on audiovisual texts, their verbal component, linguistic features of subtitles and fictive orality. The first two points provide a characterisation of audiovisual products and look at the role of the verbal component from a translational perspective. After this context is set, attention is paid to the language of subtitling. Ideas put forward in literature about morphosyntactic and lexical features typically associated to subtitling, such as neutralisation, are revised. Furthermore, special emphasis is placed on a specific issue of subtitling language, fictive orality. Since typical traits of spoken language are considered signs of fictive orality, a summary of prototypical features of spoken and written language is provided. Although not designed to be exhaustive, these are interestingly suggested as potential objects of examination. Another remarkable aspect of this section is the in-depth review of references dealing with fictive orality in scripted dialogue and subtitling. This reveals relatively scarcity of research on this subject so far, especially regarding some translation modalities such as subtitling, a gap which brings forward the worth of this contribution.

Once the framework is established, chapter 4 presents the CoPP English-Spanish parallel corpus. To do that, it is subdivided into four parts which deal with some details of the selected series, corpus compilation, and a previous analysis of language variation and subtitling standards of the source texts. The corpus building process is especially important from a methodological perspective. First, the texts were prepared, in this case source dialogues were obtained by way of transcription, subtitling was extracted from DVDs through SubRip software and all texts were edited. Then, the parallel corpus was built and prepared for queries. Semiautomatic alignment was performed using the

tool AlignAssist, the texts were automatically lemmatised and grammatically annotated through the IULA application, and the corpus was adapted to the requirements of CQP workbench to allow corpus analysis using that tool. This procedure reveals the adoption of different software to develop the required tasks for the creation of an aligned and annotated corpus, and may be found illustrative for other studies. In regard to the corpus textual material, this chapter also provides a preliminary description of the series in terms of plot, technical information and some genre-specific aspects of police procedurals, which paves the way for interpretation of results in the corresponding chapters. Finally, the language varieties and types of interaction of the source texts are analysed, as well as their compliance with subtitling standards, an important previous study for the corpus-driven approach.

The subsequent chapters report on the results of the corpus analysis, which is divided into four parts: morphosyntactic quantitative and qualitative analysis (chapters 5 and 6) and lexical quantitative and qualitative analysis (chapters 7 and 8). The sections display a similar structure which helps the reader navigate through them in an orderly manner. To begin with, selected linguistic features and analysis techniques are described, and afterwards the findings are explained and discussed. The study of each language level is closed with helpful recapitulations which progressively begin to construct the conclusions.

The chosen features for morphosyntactic quantitative analysis are distribution of parts of speech (adjectives, nouns and named entities, verbs, and adverbs), and sentence distribution and complexity (number and types of sentence, sentence length, complexity of verbal groups, coordination and subordination). A range of statistical techniques are applied to measure the significance of differences between the source texts and the subtitling, and also among each of the series: two-tailed t-tests for source and target texts' comparison, one-way ANOVAs and discriminant analysis for series contrast. In addition, different corpus and statistical tools are used, Contawords[®], IULA, CQP workbench and SPSS[®]. For instance, significant divergences are found regarding shorter sentence length and more subordinate clauses in the subtitling than in the source texts. For morphosyntactic qualitative enquiry, attention is paid to some fictive orality features (altered constituent order, ellipsis, question tags and number disagreement) as well as subtitle segmentation considerations (two-line subtitles and sentences across subtitles). The procedure used in this case is manual annotation of the features. Regarding the mentioned traits of fictive orality, a general tendency of the target texts to have less occurrences than the

source texts is observed, with the exception of tags. In relation to subtitle segmentation, the results show that it complies with standards.

On the other hand, for lexical analysis, the examined features from a quantitative perspective are aboutness, lexical density and vocabulary richness, information load and terminological density. The same statistical techniques and software as for the morphosyntactic quantitative study are applied, adding the tool Microsoft Access[®]. Among the results, those about vocabulary richness and terminological density could be highlighted as they show that there are not significant differences with the source texts, which contrast with the subtitling tendency to lexical neutralisation suggested in literature. From a qualitative approach the lexical features analysed are offensive and affective lexicon as well as creative lexicon (anomalous collocates, metaphor, coercion, hyperbole, euphemism, and expressive neologism). As for the qualitative morphosyntactic analysis, manual annotation is also carried out. In addition, the Corpus Pattern Analysis methodology is adapted to the study. Regarding the first two features, the translation choices observed either opt for a Spanish offensive or affective term, or omit the source text instances, leading to neutralisation. Analysis of the creative lexicon reveals that it is similarly low in the source and target texts, and translations include instances of language exploitation as well as standard language.

Finally, the conclusions (chapter 9) provide a thorough overview of the research conducted and its main results. The summaries of findings provided in the format of clear and concise tables are remarkable, as well as their discussion in relation to previous literature. In particular, three main issues related to the research questions are revisited in light of the results: subtitling neutralisation, the police procedurals genre and operative norms in the corpus. Previous descriptions of subtitles as displaying lexical neutralisation are contrasted with the data found in the corpus, where subtitles include most of the fictive orality signs of the source texts. Apart from the source and target texts, the different series are compared with regard to the notion of genre as a unifying criterium for the corpus, and similarities and divergences among them are pointed out. Lastly, norms regarding morphosyntactic and lexical features are formulated based on regularities observed in the results. All in all, findings obtained at different stages of the analysis are coherently interrelated in regard to these points. Moreover, a number of interesting paths for advancement in future studies are suggested.

Overall, this volume offers a fascinating journey into different linguistic features of the CoPP by means of innovative techniques. It deals with the relevant challenge of fictive orality in subtitling, together with other aspects associated to subtitling segmentation. A corpus-driven approach is adopted to study the corpus empirically, and specific techniques are chosen for different enquiries. Thus, a compound of quantitative and qualitative analyses of several morphosyntactic and lexical issues is offered. The results provide new insights on the selected features, which are contrasted with previous literature and formulated into norms of the CoPP. To conclude, this volume means a valuable step in the research line of corpus studies dealing with audiovisual translations, and more specifically the language of subtitling.