Corpus triangulation: Towards a new methodological framework for translation studies

Sofia Malamatidou (University of Birmingham, UK)

Triangulation, which refers to the combination of multiple (two or more) theories, data sources, methods, or investigators in one study of a single phenomenon, has been employed in a range of disciplines, including the social sciences (Denzin, 1989, 2006; Webb, Campbell, Schwartz, Serchest, & Grove, 1981), nursing and health services research (Begley, 1996; Fotheringham, 2010), education (Altrichter, Posch, & Somekh, 2008; Cohen & Manion, 2011), management (Jack & Raturi, 2006), and applied linguistics (Dörnyei, 2007; Magnan, 2006). The idea of triangulation also found fertile ground in corpus-based translation studies when scholars realised that a combination of different types of corpora is crucial if research in corpus-based translation studies is to move forward. For instance, it is argued that parallel and comparable corpora should not be seen as mutually exclusive, but rather as "complementary sources of cross-linguistic data" (Altenberg & Granger, 2002, p. 9), and that the examination of a monolingual comparable corpus of translated texts needs to be complemented with the comparison of these texts and their source texts, as well as a reference corpus in the languages involved in the comparison (Zanettin, 2000).

Despite the recognition that corpus-based translation research would benefit from the triangulation of corpora, little has been done in the direction of actually employing combined corpora in translation studies. Studies in which some sort of corpus combination has taken place do not foreground the triangulation approach and hardly any treat triangulation in a principled way. This state of affairs is evidence of the fact that corpus-based translation studies, although employing triangulation techniques, is not recognising their full potential. It also raises the important question of whether these studies have employed 'real' triangulation techniques, and not an ad hoc collation of different corpus data and, ultimately, of how exactly triangulation is understood within corpus-based translation studies. What is missing is a guided and principled account of how the principles of triangulation can be applied to corpus-based research, which could serve as a model and encourage triangulation to be used more widely but also in a more integrated manner.

This presentation aims to address this gap by providing a clear account of corpus triangulation techniques and demonstrating how these can be applied to empirical data. Following the definition of triangulation in the social sciences (Denzin 1989), where it has been extensively used, corpus triangulation is understood as the combination, in an integrated manner, of multiple (two or more) corpora in one study of a single phenomenon, for example the combination of a comparable monolingual corpus and a parallel bilingual corpus. To avoid the risk of loose interpretation and help refine further how exactly corpora can be combined, the corpus triangulation framework is associated with a corpus typology based on the idea of variables, values and attributes. This typology is introduced in this presentation together with clear examples of how it can facilitate triangulation to avoid further ad hoc attempts towards corpus triangulation. The purpose of corpus

triangulation is to increase knowledge about the linguistic and/or translation phenomenon under investigation, by approaching it from different perspectives and examining various relevant parameters. Thus, by providing corpus-based translation studies with a clear framework for corpus triangulation, which is based on a comprehensive and flexible corpus typology, the quality and rigour of research conducted in the field can be improved, as clarity and objectivity increase. This means that not only can we answer existing questions in translation studies with more confidence, and increase the reliability of research, but that we can also attempt to address new questions, which have been left unanswered until now.

To demonstrate how corpus triangulation can be used in corpus-based translation studies, the methodological framework is applied to a case study focusing on the examination of the language of translation. In particular, corpus triangulation techniques are used for the investigation of the pragmatic factors affecting the use of adversative connectives in Russian translated texts from English. The factors examined are the genre of the text (fiction, children's fiction, and non-fiction), the audience it addresses (adults vs. children), the influence from target linguistic conventions, and the influence from the source texts. Connectives are considered optional linguistic units since their presence in a text is not a necessary condition for the establishment of a link between two word groups, clauses or sentences. This optionality makes connectives a good candidate for cross-linguistic examination, especially when mediation is involved in the form of translation. Regarding function, adversative connectives (e.g. *but*, *yet*, *although* in English and *а*, *на самом деле*, *хотя* in Russian) signal that the new information will contradict the information already available as in Examples (1) and (2) below.

> (1) *Yet although* they lasted as long as the Soviet Union itself, and *although* many millions of people passed through them, the true history of the Soviet Union's concentration camps was, until recently, not at all well known.
> (2) *Хотя* горные снега и подземные воды, на первый взгляд, имеют мало общего, *на самом деле* они связаны. (*Although* mountain snow and underground water, at first glance, have little in common, *actually* they are related)

Texts include translated and non-translated Russian texts, and their English source texts, totalling a 9-million-word corpus. By combining these texts in different ways, four different subcorpora are compiled each associated with a different stage of analysis. The first stage consists of the analysis of a comparable bilingual (English and Russian) subcorpus of non-translated texts to examine how adversative connectives are distributed in Russian and English non-translated texts across different genres. Then, a comparable monolingual (Russian) subcorpus of translated texts is used to examine how adversative connectives are distributed across three translated genres in Russian. The third stage involves the examination of a comparable monolingual (Russian) subcorpus of translated and non-translated texts to investigate whether any observed differences might be related to the linguistic preferences found in respective non-translated Russian genres. Finally, a parallel bilingual (English-Russian) subcorpus of source and target texts is analysed to establish whether the observed differences might be traced back to English. Each stage of analysis provides a partial answer to the research question, and results

from each stage interact, cross-fertilize ideas and provide insights. It is only by combining the results from all three stages that a clear conclusion can be reached. A unique aspect of the corpus methodology, which is evidence of the integrated approach adopted, is that the first two stages of analysis rely more heavily on generating corpus results and examining the distribution of connectives in each subcorpus, while the last two stages of analysis rely more on comparing and contrasting the results generated during the first two stages of analysis.

The corpus triangulation techniques employed in the case study allow for meaningful comparisons to be made across different types of texts, in an effort to acquire a comprehensive understanding of the phenomenon of adversative connectives and how these are employed in translation. Results suggest that a complex interplay of factors affects the use of adversative connectives in Russian translation from English and these can be related to the influence from existing target language linguistic conventions, the source language interference, and audience, or other genre-specific, considerations. In particular, translated Russian fiction is very close to non-translated fiction in the same language, and the influence of the existing linguistic conventions in this genre is much higher than any other influence. Regarding children's fiction, translated Russian texts make a much higher use of adversative connectives than both non-translated Russian and English children's fiction. This could be considered an indication that considerations of comprehensibility have preoccupied translators, who wanted to create more accessible texts for children in particular. Finally, regarding non-fiction, translated texts seem to be somewhere in between Russian non-translated and English non-fiction, suggesting that some influence from English source texts might have been exerted. Had corpus triangulation techniques not been employed, it would not have been possible to examine the way in which these different factors interact and, ultimately, affect different genres to a different degree and level.

References

Altenberg, B., & Granger, S. (2002). Recent Trends in Cross-Linguistic Lexical Studies. In B. Altenberg & S. Granger (Eds.), *Lexis in Contrast: Corpus-Based Approaches* (pp. 3–48). Amsterdam: John Benjamins.

Altrichter, H., Posch, P., & Somekh, B. (2008). *Teachers Investigate Their Work: An Introduction to the Methods of Action Research*. London and New York: Routledge.

Begley, C. M. (1996). Using Triangulation in Nursing Research. *Journal of Advanced Nursing, 24* (1), 122–128.

Cohen, L., & Manion, L. (2011). *Research Methods in Education* (7th ed.). London and New York: Routledge.

Denzin, N. K. (1989). *The Research Act: A Theoretical Introduction to Sociological Methods* (3rd ed.). Englewood Cliffs: Prentice Hall.

Denzin, N. K. (2006). *Sociological Methods: A Sourcebook*. Chicago: Aldine Pub. Co.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.

Fotheringham, D. (2010). Triangulation for the Assessment of Clinical Nursing Skills: A Review of Theory, Use and Methodology. *International Journal of Nursing Studies, 47* (3), 386–391.

Jack, E. P., & Raturi, A. S. (2006). Lessons Learned from Methodological Triangulation in Management Research. *Management Research News*, *29*(6), 345–357.

Magnan, S. S. (2006). From the Editor: The MLJ Turns 90 in a Digital Age. *The Modern Language Journal*, *90* (1), 1–5.

Webb, E. J., Campbell, D. T., Schwartz, R. D., Serchest, L., & Grove, J. B. (1981). *Nonreactive Measures in the Social Sciences*. Boston: Houghton Mifflin.

Zanettin, F. (2000). Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis. In M. Olohan (Ed.), *Intercultural Faultlines: Research Models in Translation studies I: Textual and Cognitive Aspects* (pp. 105–118). Manchester: St. Jerome.