Straddling the divide between contrastive and translation studies

(Article begins on next page)

INTRODUCTION

Straddling the Divide between Contrastive and Translation Studies

Adriano Ferraresi and Silvia Bernardini

University of Bologna (Italy)

Contrastive studies and translation studies have much in common, particularly in terms of data and methods, and much to learn from each other, since the other's findings are of immediate relevance for the interpretation of their own. Yet journals, book and conferences bringing the two perspectives together in a balanced manner are still relatively infrequent, two exceptions being the journal that hosts this special issue, and the Conference that provided the initial spark from which it originated, *Using Corpora in Contrastive and Translation Studies 6* (*UCCTS6*).

While the relative distance between contrastive and translation studies might seem somewhat puzzling, there are a number of reasons behind it, some historical, some conceptual, which we would like to briefly discuss in this introduction. First of all, the rapprochement between the two fields is quite recent, dating back to the 1990s, when both disciplines underwent rather radical 'turns'. On the one hand, contrastive linguistics was drifting away from applied linguistics and foreign language pedagogy and approaching descriptive linguistics (Granger, 2003; Mair, 2018). Second language acquisition research had by then confirmed that a direct causal link between interlinguistic differences and learning difficulties was oversimplistic (Corder, 1981), and the communicative approach (Nunan, 1989) left precious little space for form-focused contrastive instruction in the foreign language classroom. At the same time, the importance of intercultural communication was growing globally, and with it the interest in the descriptive analysis of similarities and differences across languages at a variety of levels, including the pragmatic one (House and Blum-Kulka, 1986).

On the translation side, the newborn discipline of Translation Studies (Holmes, 1988) conceptualized translations as facts of the target culture (Toury, 1995) and translation as an activity embedded and reflecting social practices and structures (Reiss and Vermeer, 2013(1984); Bassnett and Lefevere, 1992). The ensuing radical change of paradigm took the study of translation away from comparative literature departments, where the study of translation was conceptualized as the philological, equivalence-focused analysis of canonical literary texts and their translations, and into translation departments, where

translation studies were now approached as revolving around the descriptive, situated linguistic analysis of contemporary translation practices (Toury, 1995). In epistemological terms, the two disciplines thus came to occupy a neighbouring, if not completely overlapping area, despite their quite distant points of departure.

This alignment was no doubt favoured by the emergence, at approximately the same time as the two turns occurred, of corpus linguistics. As research areas focusing on the situated descriptive analysis of contemporary texts, contrastive and translation studies have indeed found their perfect complement in corpus methods. We would argue that the impressive development of both fields is largely *thanks to* the simultaneous growth in scholarship in corpus linguistics, which has made available new tools, resources and concepts. Analysing contributions to *Languages in Contrast* from the first issue in 1998 until 2018, Hasselgård (2020) finds a high proportion of corpus-based papers, which has even increased in recent years: as many as 83 articles out of the 100 published between 2010 and 2018 are indeed corpus-based or at least corpus-informed.

The corpus resources used by the two disciplines partly overlap, with parallel corpora being equally central to both. Contrastive studies relying on parallel corpora aim to observe translation-mediated correspondences to "eventually arrive at a clearer notion of what counts as equivalent across languages" (Johansson, 2007: 5). Conversely, studies of translation rely on the same corpus type to investigate translation correspondences and shifts, since these can "furnish indications of the translational norms adopted by the translator, [his or her] interpretation of the original text and the strategy applied during the process of translation" (van Leuven-Zwart, 1989: 151). In other words, contrastive studies try to isolate cross-linguistic similarities and differences using the translator's work as an indirect access point or tertium comparationis. Translation studies instead try to isolate translation choices that lead to systematic differences between source and target texts, *excluding* cross-linguistic or other effects.

In both fields, there is an awareness that correspondences across source and target texts do not tell the whole story. In fact, criticism has been levelled at parallel texts from both fields: contrastive studies scholars have pointed out that translation provides a "distorted picture" (Teubert 1996: 247) of the language it represents, and translation scholars have suggested that comparing source and target texts is a "long-standing obsession" that should be abandoned (Baker 1993: 237). For these reasons, both types of studies nowadays often include in their design a collection of genre- and topic-comparable texts in the target language, which does not include translated data. Translation research typically compares this subcorpus with the target language one in a *monolingual comparable* corpus comparison, to *isolate* translation-specific features. Contrastive research instead compares it with the source language one, through a *bilingual comparable* corpus comparison, to *exclude* translation-specific features.

Beyond the study design, however, an awareness of, and ability to grapple with, findings from the sister field is essential to arrive at meaningful interpretations of parallel data. To give an example from translation studies, Evert and Neumann (2017) compare German and English texts and their translations into the other language, looking for evidence of *shining-through*, which they define, following Teich (2003), as "cases where the diverging frequencies of options existing in both languages are adapted in translated texts to those of the source language, thus resulting in a frequency difference between translations and

comparable non-translated texts in the target language" (Evert and Neumann, 2017: 49). Their multivariate analysis relies on indicators of underlying functions drawn from an extensive previous contrastive comparison of the two languages (Neumann, 2013). By selecting only those indicators that have been found to be comparable across languages, any differences in their frequencies between originals and their translations can be meaningfully related to translation. The analysis reveals that "shining-through is more pronounced for translations from English into German than for the opposite translation direction, pointing towards a prestige effect in this language pair" (Evert and Neumann, 2017: 47).

These findings belong to a growing body of research looking into the typical features of translation and other forms of "constrained" communication (Hansen-Schirra *et al.*, 2012; Volansky *et al.*, 2015; Kotze, 2022). Taken together, these studies provide solid evidence of quantifiable differences between mediated and unmediated instances of language production. While not always easy to interpret or generalize, these findings do contribute to establishing expectations about translation effects one is likely to find, and therefore should control or account for, in any study employing translated data, including contrastive ones.

Given their methodological proximity and their epistemological distance, it is particularly important that translation and contrastive studies scholars engage with each other. This is the rationale behind the UCCTS conference series, that since 2003 has brought together researchers using corpora for contrastive linguistics and translation studies. In line with current developments in both fields, the 6th edition, which took place in Bertinoro (Italy) in September 2021, made a specific call for solid applications of quantitative methods, the cross-fertilization of product- and process-based approaches, and the development of adequate theoretical models. Beyond the traditional concerns of both disciplines, it also encouraged participants to join forces to investigate the common ground between translation-mediated cross-linguistic influence and other language contact situations as a powerful, though no doubt complex, means of attaining higher-order generalizations about language use in a world in which linguistic superdiversity is becoming the norm.

The selection of contributions in this special issue testify to the willingness (and, we would argue, the success) of the UCCTS community in turning this agenda into research that really straddles the divide between contrastive and translation studies. Despite obvious differences in foci and approaches, positioning works closer to one side of the split or the other, the kind of corpus setups and corpus data, as well as the techniques of analysis employed in the following contributions once again reinforce the idea that the two disciplines share more than they differ.

All of the contributions make use of a combination of parallel and comparable corpora, with some of them (e.g. work by Janebová/Martinková, and by Henkel) relying on bilingual/multilingual comparable *and* parallel corpora following in the wake of Johansson's (2007) influential corpus model, and others complementing insights from comparable corpora of translated and non-translated texts with data obtained from large reference corpora of the languages under investigation (e.g. work by Oster and by Marco Borrillo/Peña Martínez).

In terms of texts types, most contributions draw on corpora of literary texts, for the most part novels, with two exceptions. The work by Maekelberghe/Delaere takes into account several genres/registers and factors in this variable as a potential cause or co-cause for the differences observed across

languages and across translated vs. non-translated texts, while Janebová and Martinková validate their results obtained from literary texts resorting to a corpus of spontaneous conversation.

When it comes to methods of analysis, a mixture of quantitative and qualitative analyses is observed in each contribution. The quantitative dimension involving frequency counts of lexical items, translation strategies, etc., is in every case followed up by statistical testing, which reflects an increasingly widespread tendency in corpus linguistics at large. The statistical techniques employed range from simple chi-squared and/or Fisher's tests of overall corpus frequencies (e.g. Oster), to Wilcoxon-Mann-Whitney tests taking into account internal corpus variation (Henkel), to advanced multivariate techniques like Hierarchical Configural Frequency Analysis, conditional inference trees and random forest analysis (Maekelberghe/Delaere). What is especially notable is that, despite differences in statistical techniques and the degree of automation involved in the extraction of the corpus data under investigation, e.g. by perusal of unannotated text (e.g. Janebová/Martinková) or drawing on part-of-speech and lemma annotation (e.g. Marco Borrillo/Peña Martínez), the analyses share a substantial component of fine-grained manual scrutiny and categorization. This is most notably true of parallel concordances, which are known to hinder any attempt at automatizing their analysis, but it is also true of the monolingual comparable analyses: for example, even the most corpus-driven analysis in the special issue, the one by Maekelberghe/Delaere, required the authors to manually annotate more than 6,000 instances of verbal gerunds, which might ultimately be taken as a blurring of the very distinction between the labels of qualitative and quantitative.

The first article in this special issue, by **Markéta Janebová** and **Michaela Martinková**, is the most prominently contrastive. The study investigates demonstratives expressing similarity in Czech and English, focusing on the pair of dictionary equivalents *such* and *takový*. While the latter has received attention in the literature, especially due to its use in discourse to create *ad hoc* categories, few studies have been carried out on the former, and none has tried to establish their actual degree of interlinguistic correspondence. The authors draw on the English-to-Czech and Czech-to-English component of InterCorp (Čermák and Rosen, 2012) to determine the degree of mutual correspondence between *such* and *takový*, and then zoom in on the syntactic structures they participate in and the discursive functions they perform. This analysis, carried out on a bidirectional comparable and parallel translation corpus of fiction texts, is complemented by a pilot study based on corpora of spontaneous spoken language, the Spoken BNC2014 (Love *et al.*, 2017) for English and ORAL2013 for Czech (Benešová *et al.*, 2013). In trying to get a more comprehensive picture of the usage patterns of these demonstratives, the authors thus also factor in a type of language data that is rarely explored in contrastive analyses. Results pertaining to original fiction texts suggest that in a majority of cases *such* and *takový* are used with a similar textual function, i.e. *ad hoc* categorization based on antecedent context. Despite the similarity, however, the two demonstratives are not often found as translations of one another. According to the authors, this is due both to other demonstratives being used with a textual function in the respective languages (e.g. *takový* is often translated by *like that*), and because of cross-linguistic mismatches in textual functions, most notably the absence of a recognitional/approximative use of *such*,

which is instead frequent for *takový* (as in *takové tři měsíce* 'about three months'). These mismatches are confirmed by the analysis carried out on the spoken corpora, where the recognitional/approximative uses of *takový* further point to a heightened intersubjective use of this demonstrative, which conveys speakers' attempt at negotiating "discourse referent tracking" for the hearer (Ghesquière and Van de Velde, 2011: 792). The authors interpret these results in the context of current discussions on intersubjectivity and *ad hoc* categorization, but they also underline the importance of testing hypotheses across as many and as varied sources of data as possible.

The article by **Daniel Henkel** is a translational and contrastive analysis of the perception verbs *see* and *hear* in English, and their equivalents *voir* and *entendre* in French. The author extensively reviews previous work on linguistic realizations of perception in English, French and several other European languages, and identifies a lack of studies investigating how cross-linguistic differences in this area might impact translation choices. To fill the gap, Henkel draws on a purpose-built bidirectional comparable and parallel corpus of 19th century fiction, combining multiple analytical perspectives allowed for by this corpus setup. His aims are to test whether translated texts differ from originals in the same language in terms of use of perception verbs (monolingual comparable perspective), and whether differences, if any, might be attributed to source language or source text features (parallel perspective). Quantitative analyses reveal that translations in both languages differ from comparable originals, with English translations featuring significantly fewer occurrences of *hear*, and French translations featuring significantly more occurrences of *entendre*. Similar patterns of over- and under-use are observed for *see* and *voir*, but differences are not significant in this case. In the parallel component of the study, the author looks for a possible explanation of these findings through a meticulous categorization of translation equivalences and shifts. Results point to substantial mutual correspondence between *hear*/*entendre* and *see*/*hear*, with verbs of cognition appearing as the most frequent alternative solution for verbs of perception in translation. Through a final small-scale analysis of translators' styles, Henkel suggests that patterns of addition, omission and literal rendering of perception verbs might in fact be closely linked to individual preferences: only a minority of translators seem to display patterns of use of these verbs that are "closest to target language norms", which would possibly explain the differences across translated and original texts. Such differences, the author notes, can only be revealed and accounted for by a combination of quantitative, frequency-based analyses and qualitative painstaking scrutiny of micro-textual choices.

In the third article of this special issue, **Ulrike Oster** presents a contrastive analysis of the prototypical lexemes associated with the concept of anger in German and Spanish, and explores whether and how translation choices impact the portrayal of this emotion taking texts translated from German into Spanish as a case in point. As argued by the author, the linguistic expression of emotions is widely researched in several branches of linguistic enquiry, including contrastive and cognitive linguistics, which have highlighted (a)symmetries across languages deriving from embodiment-, cognition- and culture-related factors. The paper aims to contribute to this line of enquiry by adding an explicitly translational perspective, focusing on the rendition of conceptual metaphors, physical effects and consequences of anger. For her contrastive analysis, Oster relies on two large reference corpora of German and Spanish, DWDS (Geyken, 2007) and Corpus

del Español (Davies, 2016), respectively. She looks at co-occurrence patterns – and especially metaphorical uses and semantic prosodies and preferences – of the lexemes *Wut*, *Zorn*, *Ärger* in German and *ira*, *rabia*, *enojo* in Spanish. The translational analysis is carried out on the German-to-Spanish component of the multilingual, parallel and comparable corpus COVALT,[1] and focuses on all anger-related lexemes observed in German, for which Spanish parallel segments are then extracted and perused. Oster's findings point to differences across the way German and Spanish conceptualize and represent anger, whereby the former places more emphasis on physical manifestations of the emotion, whereas the latter conceptualizes it as an autonomous force that acts independently of the experiencer. These differences are used as a backdrop to explain several "marked" translation solutions in COVALT: specifically, the author observes that Spanish translations feature several conceptual mappings that are either unusual or unusually frequent in this language, such as emphasis on physical manifestations of anger or on the behaviour of the angry person (e.g. turning red or breaking objects). The paper concludes with a reflection on how such cross-linguistic differences, which are more likely to go unnoticed and rendered literally than grammatical or lexical patterning, can contribute to altering the "emotional makeup" of characters in literary translation, with clear implication for readers' perception of a text.

The next article, by **Josep Marco Borrillo** and **Gemma Peña Martínez**, aims to test Halverson's (2017) revised Gravitational Pull Hypothesis (GPH) focusing on the distinction between imperfective and perfective verb aspect in two language pairs: English-Catalan and French-Catalan. The GPH was proposed by Halverson as a model to explain patterns of over- and under-representation of linguistic features in translated language and has received increasing attention in (corpus-based) translation studies. The authors argue that confirmation of the GPH can only derive from accumulation of evidence concerning multiple language combinations and multiple linguistic features of a varied nature, which is what they aim to contribute through their work. Their specific hypothesis is based on a wide-ranging review of work on the GPH, on literature on the imperfective and perfective aspect in bilingual theory and contrastive linguistics studies, as well as on a preliminary corpus-based analysis of original Catalan, English and French texts in the COVALT corpus, complemented by data from web-derived reference corpora. The authors hypothesize, in broad terms, that since the imperfective/perfective aspect distinction is more grammaticalized in Catalan and French than in English (which lacks a grammaticalized imperfective form), perfective forms will be overrepresented in texts translated into Catalan from English, while texts translated from French might or might not display such tendency due to higher isomorphism between source and target language. The main analysis is carried out by counting frequency of occurrences of POS-defined verb patterns in the Catalan target texts of the English>Catalan and French>Catalan components of COVALT. This step is followed by a) statistical testing of frequency differences and b) manual identification of source text triggers of target text verb patterns. The same procedure is then followed in the opposite direction, starting from English and French source texts and analysing their Catalan renditions. As predicted by the GPH, results point to a significant overuse of verbs in the perfective form both in English>Catalan translations and

---

[1] http://www.covalt.uji.es

in French>Catalan translations. This result is explained by way of reference to the salience of the perfective aspect in the target language, Catalan, which counteracts the high degree of isomorphism and distributional overlap between French and Catalan verbal aspect patterns. The article concludes with a reflection on the benefits involved in bringing together theoretical insights from various disciplines, as well as a thought-provoking methodological note on the appropriateness of univariate statistical methods in the context of GPH testing.

In the final contribution, **Charlotte Maekelberghe** and **Isabelle Delaere** investigate the English verbal gerund in translated and non-translated English, as well as its rendering in translation from English into German and Dutch. English verbal gerunds represent, as the authors argue, an interesting testbed to explore the intersections between contrastive linguistics and translation studies: this is both due to their hybrid lexico-grammatical status between nominal and clausal structures, and their status as 'unique items' (borrowing Tirkkonen-Condit's (2004) term), i.e. linguistic forms/structures for which no obvious translation equivalents exist in a target language. German and Dutch are chosen as a case in point because of their lack of hybrid forms: since they only feature nominal or verbal structures, observing obligatory grammatical shifts involved in translations of English gerunds into these languages is a possible way of making the nominal or clausal use of the original form explicit. The two authors rely on two comparable and parallel corpora, i.e. the CroCo corpus (Hansen-Schirra *et al.*, 2012) for the English-German language pair and the Dutch Parallel Corpus (Macken *et al.*, 2011) for English-Dutch, from which they draw *ad hoc* subsets to maximize cross-corpus comparability. Analyses make use of sophisticated statistical techniques: the monolingual comparable analysis is based on Hierarchical Configural Frequency Analysis (Gries, 2004), which makes it possible to identify usage profiles of verbal gerunds in a bottom-up, corpus-driven way, and thus reveal whether different nominal or clausal profiles emerge for gerunds, and whether these profiles are distributed differently in translated and non-translated English. The parallel component of the study adopts another set of advanced statistical techniques, i.e., a conditional inference tree and random forest analysis (Tagliamonte and Baayen, 2012). These are used to shed light on the factors which best predict whether a given English gerund is translated with a nominal or clausal translation equivalent. Results from the first analysis provide empirical confirmation of the distinction between nominal and clausal profiles (plus a third, less frequent profile involving gerunds with an explicit subject), and that original and translated English display diverging usage patterns in terms of these profiles. The parallel analysis further suggests that nominal and clausal profiles tend to be translated with equivalent nominal and clausal structures. Interestingly, in both analyses text genre is found to have a substantial impact, whereby language-specific effects are deeply intertwined with genre-specific ones. This result, which matches observations made by Janebová and Martinková in the first article of this special issue, but which also emerges, if less explicitly, throughout all contributions (as well as a wealth of other works; see e.g. Kruger and van Rooy, 2012), spotlights a methodological issue that can no longer be overlooked in contrastive and translation studies alike, particularly if based on corpora.

In closing, we would like to briefly consider what lies ahead for corpus-informed contrastive and translation studies. Above we have claimed that unawareness of the impact of contrastive differences on translation and of translation choices on language contrasts would make the results from studies in both disciplines equally uninterpretable. However, one might wonder if it still makes sense at all to postulate the existence of clear boundaries between the two disciplines, and whether such boundaries will continue to exist in the future. Translation, along with text creation in general, is facing major transformations, linked not only to progress in artificial intelligence, but also to the demise of the printed artefact paradigm (Gambier, 2022). Prototypical professional translation is thus likely to become increasingly marginal to society (Moorkens, 2017), and to be supplemented by a range of activities characterized by variable degrees of professionalism, flexible multilingualism (Bowker and Buitrago Ciro, 2019), and computer-driven language generation and processing (Wang and Sawyer, 2023).

The dilemma for translation studies is whether to embrace these new hybrid forms of translation, adapting their established research paradigms accordingly, or to stick to the old paradigm, at the risk of irrelevance. But contrastive studies will be affected too: it will be increasingly difficult to say whether a text is a translation, and to pair it with a corresponding 'original': "translation without a source text" (Davier and van Doorslaer, 2018) could soon be the norm. The boundaries between the two disciplines would thus effectively be eroded, to an extent that is hard to predict at the moment. We concur with Hasselgård (2020: 201) that "[c]onsidering the increasing amount of global media, migration and international travel, the need for insights into cross-linguistic matters is unlikely to diminish". But we would like to suggest that these matters can only be effectively approached if translation and contrastive studies scholars join forces.

Acknowledgements

References

Baker, M. 1993. Corpus Linguistics and Translation Studies. Implications and applications. In *Text and Technology. In Honour of John Sinclair*, M. Baker, G. Francis and E. Tognini-Bonelli (eds), 233–250. Amsterdam: Benjamins.

Bassnett, S. and Lefevere, A. 1992. *Constructing Cultures. Essays on Literary Translation*. Clevedon: Multilingual Matters.

Benešová, L., Křen, M. and Waclawičová, M. 2013. *ORAL2013: Reprezentativní korpus neformální mluvené češtiny*. ÚČNK FF UK, Praha. Available at http://www.korpus.cz.

Bowker, L. and Buitrago Ciro, J. 2019. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley: Emerald Publishing.

Čermák, F. and Rosen, A. 2012. The Case of InterCorp, a Multilingual Parallel Corpus. *International Journal of Corpus Linguistics* 17(3): 411–427.

Corder, S.P. 1981. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.

Davier, L. and van Doorslaer, L. 2018. Translation Without a Source Text: Methodological Issues in News Translation". *Across languages and cultures* 19(2): 241–257.

Davies, M. (2016) *Corpus del Español: Web/Dialects*. Available at `http://www.corpusdelespanol.org/web-dial/`.

Evert, S. and Neumann, S. 2017. The Impact of Translation Direction on Characteristics of Translated Texts. A Multivariate Analysis for English and German. In *Empirical Translation Studies. New Theoretical and Methodological Traditions*, G. De Sutter, M.-A. Lefer and I. Delaere (eds), 47–80. Berlin: Mouton de Gruyter.

Gambier, Y. 2022. Revisiting Certain Concepts of Translation Studies through the Study of Media Practices. In *The Routledge Handbook of Translation and Media*, E. Bielsa (ed), 91–107. London: Routledge.

Geyken, A. 2007. *The DWDS Corpus: A Reference Corpus for the German language of the 20th Century*. In *Collocations and Idioms: Linguistic, Lexicographic, and Computational Aspects*, C. Fellbaum (ed), 23–41. London: Bloomsbury.

Ghesquière, L. and Van de Velde, F. 2011. A Corpus-Based Account of the Development of English *Such* and Dutch *Zulk*: Identification, Intensification and (Inter)Subjectivity. *Cognitive Linguistics* 22: 765–797.

Granger, S. 2003. The Corpus Approach: A Common Way Forward for Contrastive Linguistics and Translation Studies? In *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, S. Granger, J. Lerot and S. Petch-Tyson (eds), 17–30. Leiden: Brill.

Gries, S. Th. 2004. HCFA 3.2 — A Program for Hierarchical Configural Frequency Analysis for R for Windows.

Halverson, S. 2017. Developing a Cognitive Semantic Model: Magnetism, Gravitational Pull, and Questions of Data and Method. In *Empirical Translation Studies. New Methods and Theoretical Traditions*, G. de Sutter, M.-A. Lefer and I. Delaere (eds), 9–45. Berlin: Mouton de Gruyter.

Hansen-Schirra, S., Neumann, S. and Steiner, E. 2012. *Cross-Linguistic Corpora for the Study of Translations*. Berlin: Mouton de Gruyter.

Hasselgård, H. 2020. Corpus-based Contrastive Studies. Beginnings, Developments and Directions. *Languages in Contrast* 20(2): 184–208.

Holmes, J. S. 1988. The Name and Nature of Translation Studies. In *Translated! Papers on Literary Translation and Translation Studies*, J. S. Holmes, 67–80. Amsterdam: Rodopi,

House, J. and Blum-Kulka, S. 1986. *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Narr.

Johansson, S. 2007. *Seeing through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam: Benjamins.

Kotze, H. 2022. Translation as Constrained Communication: Principles, Concepts and Methods. In *Extending the Scope of Corpus-based Translation Studies*, S. Granger and M.-A. Lefer (eds), 67–98. London: Bloomsbury.

Kruger, H. and van Rooy, B. 2012. Register and the Features of Translated Language. *Across Languages and Cultures 13*(1): 33–65.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. 2017. The Spoken BNC2014: Designing and Building a Spoken Corpus of Everyday Conversations. *International Journal of Corpus Linguistics* 22(3): 319–344.

Macken, L., De Clercq, O. and Paulussen, H. 2011. Dutch Parallel Corpus: A Balanced Copyright-Cleared Parallel Corpus. *Meta* 56(2): 374–390.

Mair, C. 2018. Contrastive Analysis in Linguistics. In *Oxford Bibliographies in Linguistics*, M. Aronoff (ed). Oxford: Oxford University Press. Online.

Moorkens, J. 2017. Under Pressure: Translation in Times of Austerity. *Perspectives* 25(3): 464–77.

Neumann, S. 2013. *Contrastive Register Variation. A Quantitative Approach to the Comparison of English and German*. Berlin: Mouton de Gruyter.

Nunan, D. 1989. *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.

Reiss, K. and Vermeer, H. J. 2013(1984). *Towards a General Theory of Translational Action: Skopos Theory Explained. Translated from the German by Christiane Nord*. London: Routledge.

Tagliamonte, S. and Baayen, H. R. 2012. Models, Forests, and Trees of York English: Was/were Variation as a Case Study for Statistical Practice. *Language Variation and Change* 24(2): 135–178.

Teich, E. 2003. *Cross-linguistic Variation in System and Text. A Methodology for the Investigation of Translations and Comparable Texts*. Berlin/New York: Mouton de Gruyter.

Teubert, W. 1996. Comparable or parallel corpora? *International Journal* of *Lexicography* 9(3): 238–264.

Tirkkonen-Condit, S. 2004. Unique Items – Over- or Under-represented in Translated Language? In *Translation universals. Do they exist?*, P. Kujamäki and A. Mauranen (eds), 177–184. Amsterdam: Benjamins.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam and Philadelphia: Benjamins.

Van Leuven-Zwart, K. 1989. Translation and Original. Similarities and Dissimilarities, part I. *Target* 1(2):151–182.

Volansky, V., Ordan, N. and Wintner, S. 2015. On the Features of Translationese. *Digital Scholarship in the Humanities* 30(1): 98–118.

Wang, P. and Sawyer, D. B., 2023. *Machine Learning in Translation*. London: Routledge.

Address for correspondence

Adriano Ferraresi
Department of Interpreting and Translation
Corso della Repubblica 136
47121 Forlì (FC)
Italy

adriano.ferraresi@unibo.it
https://orcid.org/0000-0002-6957-0605

Co-author information

Silvia Bernardini
Department of Interpreting and Translation

Adriano Ferraresi and Silvia Bernardini

Corso della Repubblica 136
47121 Forlì (FC)
Italy

silvia.bernardini@unibo.it
https://orcid.org/0000-0003-0750-4861