

Egan, Thomas and Dirdal, Hildegunn (eds). 2017. *Cross-linguistic Correspondences: From Lexis to Genre*. Amsterdam/Philadelphia: John Benjamins. Studies in Language Companion Series 191. 295 pages. <https://doi.org/10.1075/slcs.191>

Cross-linguistic Correspondences: From Lexis to Genre is a recent edited volume in John Benjamin's *Studies in Language Companion Series*. It is framed as showcasing a dynamic and expanding area: contrastive linguistics applying corpus-linguistic methodology. The volume comprises ten empirical studies, presented in a pre-conference workshop on contrastive corpus linguistics in 2015 (ICAME 37). English is involved in all ten studies and six other languages are part of the comparisons: Norwegian, Swedish (three studies each), Czech, Lithuanian (two studies each), German and Spanish (one study each).

The introductory chapter by the editors Thomas Egan & Hildegunn Dirdal is called *Lexis in Contrast Today*, signalling the centrality of lexis in the volume. When applied to lexis, contrastive linguistics can “not only provide insight into the similarities and differences between the lexis of two or more languages, but also contribute to the understanding of the meaning and use of lexical items within each of those languages, and to the understanding of lexical systems more generally” (p. 2). These three contributions are reviewed and exemplified in the introduction. The introduction also discusses the key question of how to go about comparing like with like when studying a given phenomenon across different languages, through the notion of tertium comparationis.

Even if corpus-based contrastive linguistics is a relatively young field that emerged in the 1990s, the editors stress the widening scope of the area, which can be witnessed through (a) the adoption of a greater range of theoretical approaches, (b) the number of languages compared increasingly including more than the prototypical two, (c) the types of lexical items investigated moving beyond adverbs, verbs and discourse particles, (d) the development of “contrastive phraseology, which acknowledges the importance of multi-word units in natural language” (p. 3) and (e) the incorporation of genre in study designs. All these developments are illustrated not only in the introduction (where they are supported by a substantial list of references), but also in the chapters themselves.

The volume is organised into three parts: part one revolves around lexis from a predominantly semantic perspective; part two around lexis from a mostly structural and typological perspective; and part three “contains studies where genre is an important dimension, either because genre differences are investigated or because one particular genre or text type has been chosen for investigation” (p. 22).

Chapter 2 opens the first part and summarises an extensive study by Åke Viberg on *SAYING, TALKING and TELLING: Basic verbal communication verbs in Swedish and English*. The semantic field of verbal communication is investigated through the most frequent verbs in the two languages, based on the English-Swedish Parallel Corpus (ESPC). More than 9,000 tokens per language are investigated. As background, the overall distributions of the verbs in focus are also considered in large monolingual corpora. The description of ‘who said what to whom about what’ is based on frame semantics, but other models—relating to speech act theory, evidentiality and metonymy—are also brought in to account for and generalise about various aspects of the rich data.

Chapter 3 is co-authored by Anna Cermáková and Lucie Chlumská and is called *Expressing PLACE in children’s literature: Testing the limits of the n-gram method in contrastive linguistics*. The notion of PLACE in “its widest sense” (p. 87) is examined through a data-driven analysis of four-grams in corpora of children’s literature in original English and translated Czech. (This is a follow-up on a study of comparable corpora or the two languages, and previous findings also serve as useful background data.) In the English data, these include nodes such as *side, end, back, top, edge*, etc. The English four-grams denoting PLACE are first identified (a total of 125 units), then the translation equivalents into Czech are recorded. The corresponding units in Czech have an overall tendency to be more condensed, with certain nodes being translated into single prepositions and with a generally high proportion of omissions, indicating that “translators may in many cases consider the exact expression of PLACE somehow redundant” (p. 88)—even if it is an important category in this text type. Even if this is to be expected from analytical English and inflectional Czech, the authors state that the extent to which there is a mismatch in correspondence “may still be surprising” (p. 92). Their n-gram research has implications for the idiom principle, which is said to operate to different extents in the two languages.

Chapter 4 by Hilde Hasselgård deals with *Lexical patterns of PLACE in English and Norwegian*, focusing on the English locative noun *place* and the corresponding *plass* and *sted* in Norwegian. The data come from the fiction part of the English-Norwegian Parallel Corpus (ENPC). The first step involves an examination of “what correspondence patterns reveal about similarities and differences” (p. 98) between the nouns in the two languages (amounting to some 2,000 examples). The second step involves considering the noun’s lexical surroundings to explore “what recurrent word combinations [...] can reveal about the ways in which PLACE is referred to” in the two languages. The different meanings found are neatly summed up in a table (p. 117) including the features [+home], [+indefinite], [+part of larger location], [+purpose], [+sufficiency], [demonstrative/deictic], [+designated/individualized]. The Norwegian nouns are shown to be almost in complementary distribution. It is found that not only do the two languages select different recurrent word combinations, but the three nouns are also found to have different ‘selectional preferences’ (p. 111).

Another perspective on locative meaning is given in Chapter 5, co-authored by Thomas Egan and Gudrun Rawoens, called *LOCATIVE ‘at’ seen through its Swedish and Norwegian equivalents*. It is the preposition ‘at’ that is investigated when it encodes physical location (N=506), using the fiction parts of the ESPC and ENPC. The general research question is “whether the Swedish and Norwegian translation correspondences of the English preposition can aid us in mapping its semantic network” (p. 121). The data analysis leads the authors to propose a semantic network for ‘at’, based on cognitive linguistics, which they refer to as ‘five principal types of *at*-ness’, or five different spatial relationships, which are illustrated with ideograms “intended to represent five image schemas” (p. 138). The authors also provide an interesting discussion of the structure of the networks proposed, linked to an ongoing debate in the literature.

Part two is opened by Chapter 6, called *Premodification in translation: English hyphenated premodifiers in fiction and their translations into German and Swedish*, by Magnus Levin and Jenny Ström Herold. It presents a lexico-morphological analysis of data from the ESPC and the Oslo Multilingual Corpus, totalling some 1,700 instances of translations. The analysis reports on what type of hyphenated modifiers, such as *old-fashioned*, occur in English (length in

words and construction patterns, distinguishing between left- and right-hand elements); what the translation equivalents are in German and Swedish (premodification, postmodification, other, omitted/generalized). The authors also look at what source-text constructions in German and Swedish trigger the use of hyphenated premodifiers in English translations. Lastly, they pose the typological question regarding what types of structural preferences can be identified in the three languages. It is found to be relatively common for German and Swedish to “use other types of structures than premodifiers as equivalents, which suggests that there are different structural preferences in German and Swedish than in English” (p. 172). Also, as predicted by previous research, “there are tendencies for German translations to prefer premodification and Swedish to prefer postmodification” (p. 173). The study also found evidence for translation universals, as there was less variation in the English translations than originals. The authors stress that the study being restricted to fiction texts is likely to have affected the results.

Chapter 7 is co-authored by Aurelija Usonienė and Audronė Šolienė and is called *Reportive evidentials in English and Lithuanian: What kind of correspondence?* The authors draw on data from various monolingual and translation corpora for the English hearsay adverbs *reportedly*, *allegedly* and *supposedly* and three ‘corresponding’ adverbials in Lithuanian. The study also includes bi-directionally established translations equivalents, such as comment clauses and *as*-parentheticals. The aim is to find how language-specific the realisations and conceptualisations of indirect reportive evidentiality are. The answer to the question in the title is that the mutual correspondence is very weak, even if, in terms of register distribution, both sets are most frequent in news discourse. The overall findings indicate that English prefers grammaticalized means of expression, while Lithuanian prefers lexical strategies for the category studied. It was also found that Lithuanian does not seem to have neutral evidential adverbials such as *reportedly*, and that the three Lithuanian adverbials analysed are distancing markers rather than evidentials.

Chapter 8, by Markéta Malá, is called *Non-prepositional English correspondences of Czech prepositional phrases: From function words to functional sentence perspective*. The starting point of the study is the top four prepositions in Czech (*v/ve*, *na*, *s/se* and *z/ze*) based on the multilingual parallel translation corpus InterCorp8. A random sample of

200 concordance lines per preposition is analysed, based on which hypotheses regarding divergent correspondences are formed, which are then “tested by a series of more specific queries” (p. 204). As in Chapter 3, typological differences between analytical English and inflectional Czech are explored, “reveal[ing] the consequences of the word-order principles prevalent in the two languages both at phrasal and clausal levels” (p. 199)—especially evident when the Czech adverbial prepositional phrase corresponds to an English subject noun phrase. This is also used to shed light on “the interface between syntax and information structure” (p. 215), connected to the thematic role of the subject.

The third and final part is introduced by Chapter 9, *A corpus-based analysis of genre-specific multi-word combinations: Minutes in English and Spanish*, by Isabel Pizarro Sánchez. This chapter can be said to represent English for Specific Purposes (ESP) and, more specifically, genre analysis, done in a contrastive fashion, with the aim to uncover multi-word generic lexis. A small specialised corpus of meeting minutes (25 per language), rhetorically annotated for Swalesean moves and steps, is used to retrieve n-grams. The approximately 1,000 n-grams analysed are classified as genre-specific, step-specific, field-related, function-word combination or noise. The findings show that, in the two selected structures (‘discussion’ and ‘adjourn’) and in both languages, there are clear associations between n-grams and rhetorical moves, even if the material comes from several different domains.

Chapter 10, *Citations in research writing: The interplay of discipline, culture and expertise*, by Jolanta Šinkuniene, also exemplifies an ESP-oriented study. The focus is on contrasting two academic disciplines: literary studies and linguistics. The variables culture and expertise are also relevant, as research articles by Lithuanian and British academics are examined, in addition to BA theses by Lithuanian students writing in L2 English. The corpus consists of ten samples each of research articles per discipline (i) in Lithuanian and (ii) English and (iii) BA theses in English, amounting to 60 texts in total. The author considers (a) overall frequency distribution (the total corpus frequency is about 3,500), (b) syntactic integration (integral versus non-integral) and (c) types of citations (direct quote versus paraphrase). Clear disciplinary variation was found for all three among the expert writers, but “[n]o striking cultural differences were observed” (p. 267), but this is also

hedged by reference to the small corpus. The point is also made that research articles and BA theses are different genres, which is another confounding factor.

Chapter 11, *Frequency and lexical variation in connector use*, by Sylvi Rørvik, explores textual cohesion in a cross-linguistic context. The study deals with English and Norwegian, both expert and novice genres, and is based on 100 texts each per language of newspaper opinion pieces and argumentative student essays. Occurrences of connectors are identified manually for maximum recall and all connectors are semantically classified based on a Hallidayan model. Connector frequency is calculated per 50 T-units to provide an approximate assessment of connectors per text. The lexical variation of the connectors is also investigated. No clear cross-linguistic difference is found with respect to overall frequency. The group that stands out most is the Norwegian novices, employing considerably more connectives overall. With respect to lexical variation, many cross-linguistic correspondences are found especially among the most frequent types, but many connector types are hapax legomena.

The main contribution of the work as a whole lies in the breadth it achieves in the collected studies and in the many theoretical and typological generalisations offered. The chapter authors avoid presentations that are overly bogged down in the minutiae of the languages in contrast, and consistently bring the analysis up to a level of greater abstraction. The volume thus illustrates the shift that the editors refer to in the introduction: “[a]lthough many studies within contrastive linguistics are still motivated by practical applications, such as language teaching or computer-aided translation, there has been a shift toward a more theoretical focus, where studies with applied aims bring in theoretical models of language” (p. 3).

It is exciting to witness also the broadening scope of contrastive linguistics, with studies falling into different research traditions gathered in one place. The most interesting newcomers are the two ESP-oriented studies in the genre section. (Chapter 10 (Šinkuniene), however, surprisingly does not explore the culture/language variable in any systematic way.) The volume offers not only the usual cross-linguistic comparison, but also cross-disciplinary, cross-generic and cross-demographic types of comparison. It seems that cross-linguistic comparison has much to gain from considering variables other than

language (or ‘language culture’) which may affect the data under analysis.

Despite the fact that ‘genre’ occurs in the title, the genre perspective is relatively restricted, mostly consisting in general observations about how lexical patterning or lexico-grammatical phenomena tend to vary across genre or register. Chapter 9 (Pizarro Sánchez) is a good start, but contrastive linguistics could definitely house more of this type of research. We can add to this the unfortunate situation that fiction is a privileged type of discourse for cross-linguistic studies, at least judging from this volume.

Considering the volume as a whole, my main point of criticism is the lack of connections between chapters. There is, for example, very little cross-referencing. This may be a conscious choice, as not many readers are likely to read the book from cover to cover. However, there are also instances where it becomes clear that information in one chapter would clearly have benefitted other chapters. The contributing authors seem not to be familiar with each other’s chapters, even if they were present for the conference workshop. As an example, Chapter 7 (Usonienė & Šolienė) includes a footnote that mentions a suggestion from a reviewer about an alternative method for normalising results that could have been adopted in the study (p. 186). Such an alternative method is actually used in Chapter 11 (Rørvik), where T-units are counted rather than words (p. 279). Another example of lacking intertextuality is that the statistical tests described in Chapter 11 (p. 279) would have been a useful addition to Chapter 10 (Šinkuniene) on citation practices where different variables are examined. Also, the key point in Chapter 3 (Cermáková & Chlumská) about issues in contrasting languages based on n-grams is not mentioned in Chapter 9 (Pizarro Sánchez) despite the fact that it also contrasts n-grams cross-linguistically.

This is compensated for by the introductory chapter, which provides generalisations linked to the different chapters. In fact, the area overview by the editors is impressive in its coverage and clarity and serves as an excellent introduction to current work in contrastive corpus linguistics. Taken together, the chapters testify to the momentum that the area has

254 *Reviews*

gained. The book constitutes a significant contribution to contrastive corpus linguistics and is a valuable resource for anybody interested in contrastive studies and/or corpus linguistics.

Annelie Ädel
Dalarna University, Sweden