

Karin Aijmer and Bengt Altenberg (eds), *Advances in Corpus-based Contrastive Linguistics: Studies in honour of Stig Johansson* (Studies in Corpus Linguistics 54). 2013. Amsterdam and Philadelphia: John Benjamins. ISBN 978-90-272-0359-5. x + 295 pp.

Reviewed by Lieselotte Brems (University of Liège)

This edited volume brings together papers presented at the ICAME32 workshop on ‘Corpus-based contrastive analysis’ as well as a number of invited papers. Like the workshop, the book resulting from it is dedicated to the memory of Stig Johansson, a pioneer in corpus-based (contrastive) language research and corpus compiling. Many authors explicitly acknowledge and make use of his seminal work in the field, thus truly honouring Johansson.

Thomas Egan focuses on the crucial notion of equivalence in contrastive research and proposes translational equivalence as a way of measuring semantic equivalence across three languages. He starts from the Norwegian preposition *mellom*, expressing ‘betweenness’, and uses its translations into French and English to put forward 7 semantic categories (Location, Motion, Scale, Time, Relationship, Interaction, Comparison) that serve as the starting point for the comparison of French and English. The study shows similarity in the encoding of betweenness in English and French with 56% overlap between *between* and *entre*. Non-prepositional French translations, on the other hand, are used for the spatial senses.

Åke Viberg uses multilingual corpora to revisit the typological question of verb-framed versus satellite-framed languages (Talmy 1985). His case study zooms in on the expression of motion in a vehicle and compares vehicle verbs in Swedish original texts and their translation into English, German, French and Finnish. Contrasting the inventories of vehicle verbs and their semantic extensions in these languages reveals specific differences which may be related to the obligatoriness of contrasts (e.g. Swedish *gå* is restricted to motion on foot, whereas its ‘counterpart’ *go* does not have this restriction). In addition, there are differences in the use of general motion verbs, such as English *go*, versus a more specific vehicle verb in the other languages. More generally, Viberg concludes that languages seem to prefer different perspectives and hence different ways of coding situations. He shows how multilingual corpora can help gain insight in the overall lexical profile of languages.

Rosa Rabadán and Marlén Izquierdo study how affixal negation in English original texts is translated into Spanish. The inventory of translation strategies is then verified against a monolingual Spanish corpus. The differences between translated Spanish and non-translated Spanish turn out to be statistically significant.

The over- and underuse of certain constructions in translated Spanish can be partly explained by translation universals. English and Spanish turn out to have language-specific preferences to express negation. Specific English affixes furthermore prefer particular Spanish constructions (e.g. English prefix *dis-* and lexical negation in Spanish).

Using translation corpora, **Anne-Marie Simon-Vandenberg** looks into the English adverbs of essence *basically*, *fundamentally* and *essentially* via their Dutch and French (translation) equivalents. Her findings corroborate earlier monolingual research and provide new insights. *Basically* has intersubjectified into a downtoner, whereas *fundamentally* has developed an amplifier use. The same core sense of 'in essence' can hence give rise to seemingly contradictory pragmatic inferences, i.e. from being true 'in principle' to not being completely true, or to being prototypically true. Those inferences conventionalized into downtoner and amplifier uses respectively.

In order to investigate and prove semantic change, **Kate Beeching** too uses a parallel corpus approach, arguing that translation choices reflect linguistic change. She focuses on the French pragmatic marker *quand même*, which has developed adversative and relational functions. If English translations consistently differ, this is taken as evidence that semantic change has occurred. She points out a number of obvious pitfalls of this approach, such as translation universals, errors and trends in translation practice. In addition, it is well-known that pragmatic markers are often omitted in translations. Her study shows that *quand même* is synchronically multifunctional. Zero translations or translations by intensifiers, hedges or reformulations are taken as partial evidence for its more recent pragmatic function.

Anna-Brita Stenström also presents a study on pragmatic markers, viz. Spanish *vale* and English *okay*, which are both highly frequent and multifunctional in youngspeak. *Okay* turns out to be the more versatile of the two and covers all of the functions of *vale* (i.e. directive and reactive). Upper-class girls seem to use *vale* the most, whereas *okay* is mostly used by upper-class boys. Both pragmatic markers are more frequently used by teenagers than adult speakers. As the author admits, contrastive research on pragmaticalization, and involving sociolinguistic questions, depends on the availability of appropriate comparable corpora of spoken language, which are not always easy to come by.

Sylvie De Cock and **Diane Goossens** study quantity approximation co-occurring with numbers in English and French business news, as in *around 20,000 feet* or *des milliers d'emplois*. Approximation subsumes the indication of a minimum or maximum, an interval, a quantity that is either *n* or *n* (p. 147). They observe less approximation in French overall, even though the range of devices in French and English seems similar and includes adverbs, prepositions, derivational/inflectional suffixes, verbs, nouns, punctuation marks and determiners. Possible explanations

for the differences might be that French uses vague round numbers more often and the fact that French is more formal still, avoiding imprecision.

Sylviane Granger and **Marie-Aude Lefer** look into the phraseological coverage of high-frequency bundles including French *encore* and English *yet* in bilingual dictionaries by studying monolingual and translational corpus data. They identify patterns (i.e. 2 to 5-grams) in monolingual corpora and then check their translations. They conclude that there is little overlap in coverage between dictionaries and corpora, but English bundles have better coverage overall. Bundles are often hidden in dictionary entries or the quality of examples is often poor. The authors suggest corpus-informed warning boxes might be added, and they are in favour of web-based bilingual dictionaries.

Jarle Ebeling, **Signe Oksefjell Ebeling** and **Hilde Hasselgård** too have a phraseological focus and compare recurrent word-combinations in English and Norwegian. Using a bidirectional parallel corpus, they discuss three case studies of recurrent 3-word combinations in original and translated fiction texts in search of constructional, semantic and pragmatic differences. They use a bootstrapping method to generate recurrent n-grams, rather than starting from pre-selected lexical items. They argue that this new methodology can complement more traditional approaches and can reveal differences that the latter cannot. They argue that the idiom principle works differently in both languages in that, for instance, the same meaning is expressed by a 3-word combination in one language whereas the other one uses a shorter/longer phrase.

Kerstin Kunz and **Erich Steiner** use translation and comparable corpora to compare cohesive substitution in English and German. They chart the range of clausal, nominal, verbal and substitution mechanisms, the different functions of the resources and their frequencies of use. They observe that verbal and clausal substitution do not exist in German in the strict sense, whereas nominal substitution in German is more finely differentiated than in English. This helps explain why, even though substitution is less common in German than in English, German displays a greater variety of resources. English resources are often semantically bleached or highly grammaticalized.

Jennifer Herriman presents a study on extraposition of clausal subjects in original and translated Swedish and English. She notes that information structure principles are similar in these two languages, with new information typically occurring later in the clause, but there is a difference in terms of the word order onto which the information structure is mapped. *It*-clefts are more common in Swedish, which has a preference for short and light Themes and heavier Rhemes, while English allows a greater information load before the verb. Translated English shows more extraposition than original English due to interference from Swedish. Extraposition is also shown to be overused by Swedish learners of English.

Julia Lavid, Jorge Arús and Lara Moratón use a bilingual comparable corpus of English and Spanish news reports and commentaries to study clausal and discourse-thematic differences between the two languages and genres. They argue that clausal and discourse-thematic choices reflect the different communicative purposes of the newspaper genres. Language-specific differences seem less important. In news reports Sayers are typically Thematic Heads, refer to external sources and are realized by concrete NPs. In commentaries, which are less neutral, the Carrier is used to give personal opinions, and is often an abstract noun or may consist of more complex and longer NPs (as in the example *The negative stigma attached to IMF financing is a thing of the past*, p. 266). The two genres are less differentiated in Spanish.

In conclusion, the present volume presents an interesting set of contributions to the field of corpus-based contrastive studies. The authors rely on different types of corpora, i.e. monolingual, comparable and translation corpora, and often combine different types. The contributions mostly cover rather well-known and well-documented Germanic and Romance languages such as English, Dutch, German, Norwegian, Swedish, French and Spanish, but also Finnish. As the editors admit in the introduction, English typically serves as the hub (p. 3). The contributions address a fairly large range of topics, including more theoretical contributions on such notions as comparability and the *tertium comparationis*, as well as (genre-) specific case studies to do with phraseology, pragmatics, information structure, quantity approximation, etc. Several authors address the shortcomings or dearth of parallel corpora, which are often small, genre-specific or simply not available, but also show how using translations and bidirectional parallel corpora can be a way to bypass the problem of a *tertium comparationis*. The relevance of these contributions goes beyond contrastive linguistics and extends to the study of language change, language acquisition and translation studies. The questions raised in this volume will surely encourage further research in corpus-based contrastive studies and related fields.

Reference

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In *Language Typology and Syntactic Description. Volume 3: Grammatical Categories and the Lexicon*, Timothy Shopen (ed.). Cambridge: Cambridge University Press, 57–149.