English in Contrast: Corpus-based Approaches. Introduction

Signe Oksefjell Ebeling (University of Oslo)

Hilde Hasselgård (University of Oslo)

The articles in this special issue of the *Nordic Journal of English Studies* are concerned with different and innovative aspects of corpus-based contrastive analysis, featuring English as the pivot language. In other words, the contributions feature studies in which English is compared to at least one other language. Most of the papers were presented at the 44th ICAME conference at North-West University in Vanderbijlpark, South Africa, 17–21 May 2023.

After 30 years of existence, contrastive corpus linguistics is still thriving and has developed immensely from its early beginnings (Aijmer, Altenberg, and Johansson 1996). The field is now well established, and much insight has been gained into the languages compared, the relationships between them, and about languages in general. However, as stated in the call for papers for the pre-conference workshop, it is still the case that we need a greater diversity of multilingual corpora and expertise and creativity regarding their use (Johansson 2012). Besides the workshop theme 'Corpora crossing language borders', contributors were challenged to expand the present state-of-the-art of the field by considering potential innovations in terms of methods and corpora employed and/or topics and languages compared. In this way, it was hoped that the contributions to the workshop would be able to push the field of corpus-based contrastive linguistics further.

Many of the contributions in this special issue present such innovations. For example, Ebeling and Hasselgård outline a refined method of calculating mutual correspondence on the basis of bidirectional translation data. Curry's analysis of the rhetorical functions of questions draws on comparable corpora and reflects on ways of ensuring a sound

tertium comparationis for the analysis. Reflecting a recent type of text, Ruzaitė presents a cross-linguistic analysis of fake news articles. The studies span from lexicogrammatical comparisons to discourse analysis, and represent morphology, grammar, semantics and pragmatics.

From this variety of perspectives, as many as eight different languages are compared and contrasted with English. Below follow brief summaries of the ten articles in this special issue.

Signe Oksefjell Ebeling and **Hilde Hasselgård** discuss and expand the measures of translation bias (TB) and mutual correspondence (MC) in the light of textual variation, with the aim of better illuminating cross-linguistic equivalence. An English-Norwegian contrastive study of the cognate verbs GO and GÅ serves as a testbed for developing the notions of a *Distributive TB* and a *Distributive MC*. The authors conclude that the distributive variants of TB and MC, showing the range of correspondence within a dataset, are both more informative and accurate than the traditional ones. The study of cognates also revealed that it is important to take lexicogrammatical features into account when discussing cross-linguistic correspondence.

Michaela Martinková and **Markéta Janebová** investigate the relatively low mutual correspondence of self-agentive uses of venitive verbs in English and Czech. Drawing on bidirectional data from a corpus of subtitles, the authors uncover that Czech *při*-verbs do not seem to be limited to deictic contexts to the same extent as English *come*. While this may explain some of the instances of non-correspondence, there are also other factors that seem to restrain the use of *při*-verbs in contexts where *come* is the verb of choice in English, including the strong association of *při*-verbs with the arrival perspective, as well as their perfective nature.

Thomas Egan examines the verb *put* in English source and target texts and their correspondences in Norwegian. The focus is on caused motion constructions with *put* in the SVOA pattern, featuring the semantic roles of AGENT, THEME, PATH, and GOAL. The contrastive data suggest that THEME, encoded as the direct object, is a much better predictor of Norwegian verb correspondences than both PATH and GOAL. Another finding is the skewed relationship between the cognates *putte* and *put*, tentatively attributed to a cognate avoidance hypothesis, i.e., that 'translations into the language in which the cognate expression is less common are likely to contain fewer examples of this expression than translations into the language in which it is more common'.

Karolina Rudnicka and **Aleš Klégr** explore non-verbal plural number agreement across four languages—English, German, Polish, and Czech—through a triangulation of methods. The results emerging from the different kinds of acceptability ratings (questionnaires) complement each other and substantiate findings from other sources, notably empirical corpus material. The four languages are found to differ in their use of the distributive plural or singular in objects, PP adverbials or PP postmodifiers of plural nouns (subjects or NP heads as antecedents). The preferred choices in each language—typically distributive singular in German, Polish, and Czech and plural in English—seem to be context- and noun-related, at the same time reflecting language-specific rhetoric strategies.

Åke Viberg's starting point is a Swedish spatial marker of frontal orientation which has no direct equivalent in English, namely *fram*. This marker is used as a free word (roughly 'forward'), as a particle in phrasal verbs, and as a morpheme in compounds and derived words. The meaning potential of *fram* is described in three stages: basic spatial, extended spatial, and metaphorical. These are explored in detail in all the uses mentioned above. English translations, especially of bound forms, often lack the explicit element of frontal orientation.

Lobke Ghesquière and **Lieselotte Brems** compare English *rather* and French *plutôt*, both described as 'elusive adverbs'. The adverbs are found to have relatively similar functional profiles, acting as compromisers and markers of contrastive relationships (reformulation, preference, replacement, and antithesis). However, data from a bidirectional parallel corpus (Europarl direct), show that they are not always translated into each other. For example, *plutôt* is the dominant translation of *rather* only in the replacement use, while translations of the other senses are often more explicit regarding the exact type of contrastive meaning. The two adverbs are found to express different types of contrast in unequal proportions, with *rather* most frequently expressing stark contrast and *plutôt* tending towards weaker meanings of preference and reformulation.

Karin Aijmer studies the evidential adverb *apparently* in an English-Swedish contrastive perspective. The author sets out to disentangle the multifunctionality of *apparently*, notably its evidential and epistemic meanings. The English-Swedish contrastive data, supplemented with German and French translation data, show that in terms of evidentiality (referring to the source of information or epistemic justification),

apparently is associated with indirect rather than direct justification; thus, inference and hearsay are characteristic features in this regard. With reference to epistemic support (referring to degree of commitment), correspondences of *apparently* suggest that inference based on evidence is present. Yet other epistemic correspondences indicate that *apparently* may also be associated with partial support in the sense of probability or likelihood.

Camino Gutiérrez-Lanza and **Rosa Rabadán** investigate recapitulating discourse markers (DMs) in English and Spanish, e.g., *well*, *anyway*, *you know*, *I mean*, *after all* and their Spanish counterparts such as *en fin*, *en síntesis*, *definitivamente*, and *en resumidas cuentas*. The Spanish markers are identified via English, as translation equivalents in the parallel corpus P-ACTRES 2.0. In a second step, Spanish translated from English is compared to original Spanish from the monolingual CORPES XXI within fiction and non-fiction, focusing on markers with a low mutual correspondence value in the parallel corpus. Overall, recapitulation is signalled by DMs more frequently in Spanish than in English, and Spanish translations contain discourse markers that do not have an explicit source in the English original. The comparison of translated and non-translated Spanish reveals that fiction shows a greater tendency towards normalization than non-fiction, where the opposite phenomenon is found. Both registers show signs of deflation and equalization in Spanish translated from English.

Niall Curry studies the rhetorical functions of direct questions in online English, French, and Spanish climate crisis discourse. His comparable corpus consists of parascientific texts drawn from *The Conversation*, i.e., those that have thematic tags relating to climate discourse. Questions are found in more than half of the blogs in all these languages, but are most frequent in French. The discursive role of direct questions includes helping authors to manage texts, making arguments and framing hypothetical situations. The questions address some sub-themes equally across the languages (e.g., climate science and social/political issues), while climate scepticism is addressed more often in English and French than in Spanish, where climate communication is more prevalent.

Jūratė Ruzaitė's contribution concerns the language of disinformation articles in English and Lithuanian. The corpus consists of COVID-19-related fake news articles in English with their published counterparts in Lithuanian. The Lithuanian texts are usually not direct

translations: they are usually shorter than the source text, and the content may be adapted in other ways too. Generally, the TTR is higher in the Lithuanian texts, which may however be related to both text length and typological differences between the languages. It is also found that the Lithuanian texts retain the emphatic nature of the source texts but reduce the amount of tentativeness. Disinformation texts in both languages tend to be linguistically simple, and the tendency towards sensationalism is even stronger in Lithuanian than in English.

We believe that the sum of the contributions in this special issue represents the state of the art of contrastive corpus linguistics. At the same time the authors promote forward-looking perspectives, which bodes well for the future of the field of corpus-based contrastive linguistics.

References

Aijmer, Karin, Bengt Altenberg, and Mats Johansson (eds). 1996. *Languages in contrast. Papers from a symposium on text-based cross-linguistic studies*. Lund 4–5 March 1994. Lund: Lund University Press.

Johansson, Stig. 2012. Cross-linguistic perspectives. In *English corpus linguistics: Crossing paths*, edited by Merja Kytö, 45–68. Amsterdam: Rodopi.