Contrastive Corpus Studies of Pragmatic Markers

Gisle Andersen | ORCID: 0000-0002-9585-9779
Norwegian School of Economics
Bergen, Norway
Gisle.Andersen@nhh.no

Abstract

The articles in this special issue account for the use of pragmatic markers in a variety of languages by taking a contrastive approach and applying corpus methodology. This introduction article describes these notions and the way they are manifested in each contribution. The studies contribute to the fledging field of contrastive pragmatics, which encompasses investigations of pragmatic phenomena by comparing usage across or within languages. Contrastive studies aim to uncover differences and similarities observable in language use along different dimensions, most notably across languages or varieties of the same language, or across structurally or conceptually different usage domains (or combinations of these). The studies of pragmatic markers explore how contextual factors have a bearing upon language use and variation across the domains that are compared.

Keywords


This special issue of Contrastive Pragmatics contains selected papers from a panel held at the 17th International Pragmatics Conference (IPrA) in Winterthur, Switzerland in 2021. Due to Covid restrictions, IPrA was organised as a wholly digital conference, including our panel. The panel was co-convened by Maria Teresa Musacchio and Sara Gesuato, both from the University of Padua, and the current author. Despite the digital format the organisers were satisfied with the conference presentations and the web-based discussion of them.
As suggested in the title, the thematic focus of the panel and special issue is delimited by some conceptual and methodological factors. Each contribution accounts for the use of pragmatic markers by taking a contrastive approach and applying corpus methodology. Our conceptualisation of these notions and the way it is manifested in each contribution is described in the rest of this article.

The studies contribute to the fledging field of contrastive pragmatics, which encompasses investigations of pragmatic phenomena by comparing usage across or within languages. In accordance with the editorial scope of this journal, contrastive studies aim to uncover differences and similarities observable in language use along different dimensions, most notably across languages or varieties of the same language, or across structurally or conceptually different usage domains (or combinations of these). In effect, studies within this field can either be mono- or bi-/multilingual and may compare the use of similar forms across languages or consider how pragmatic items are translated across languages. The aim of such studies is generally to explore how contextual factors have a bearing upon language use and variation across the domains that are compared.

In selecting pragmatic markers as our topic in focus, we are well aware that this is a multifaceted concept that covers a variety of forms that serve a multitude of functions in discourse. Intentionally, we did not want to restrict the focus to a particular subclass but allow for a fairly wide range of phenomena. The contributions in this special issue thus investigate forms that are traditionally described as (pragmatic) particles, such as toch in Dutch (Buysse this special issue) and also in German (Labrenz this special issue). They also cover pragmatic markers of epistemic stance (Capriario, Castello this special issue) and certain prolific and much-studied pragmatic markers in English, such as well (Quinci and Musacchio this special issue) and in fact (Gesuato and Ackerley this special issue). The contributions also account for discourse markers with planning and hesitation functions, such as uh and um (Tonetti Tuebben and Landert this special issue). Furthermore, the studies explore several different aspects of discourse-pragmatic meanings/functions, reflecting the cognitive, textual and interpersonal complexity of this class. Thus, the studies account for markers that convey contrast, concession and counter-expectation, various aspects of epistemic stance such as doubt or certainty, actuality and reality, and source of knowledge, as well as hedges and softening devices, discourse-organising features and markers of repair.

Another aspect that unites the studies is their common use of corpus methodology. This includes both the use of purpose-built corpora (such as the corpus of TED talk transcripts studied by Quinci and Musacchio) and
off-the-shelf corpora (such as the British Academic Written English Corpus, BAWE, the German corpus of the Research Unit for Emerging Grammars, RUEG, and the Santa Barbara Corpus of Spoken American English). The studies dealing with translation-related issues study parallel bilingual corpora, such as Europarl, containing the proceedings of the European Parliament. Other studies investigate learner language or classroom discourse either through multilingual corpora such as the Louvain International Database of Spoken English Interlanguage (LINDSEI) or a monolingual corpus such as the TOEFL 2000 Spoken and Written Academic Language Corpus (T2K-SWAL).

The article by Lieven Buysse deals with two Dutch particles, eigenlijk and toch, and how translations of these are manifested in a parallel Dutch-English corpus considering both translation directions. He explores the various functions of the particles in Dutch and maps their translation equivalents in the English component of the Europarl corpus. Common English counterparts of eigenlijk is in fact, really and actually, forms which capture well the meanings of contrast combined with reality and factuality that is conveyed by the Dutch particle. The particle toch, on the other hand, is very often not translated, or occurs in translations that lack an explicit lexical trigger in the source text. Buysse ascribes this to the complex duality of this marker as conveying both concessive and common-ground meanings.

Marcella Caprario considers the multifunctionality of the epistemic stance markers I mean and I guess and how their use varies across discipline and speaker role in classroom discourse. She assigns a variety of functions to the two markers, namely epistemic stance, softening, marking the speaker's viewpoint or personal perspective, discourse-organising, and marking decision-making as it happens. Her study shows that register and speaker roles have significant bearings on choice of form. For example, I guess is clearly the preferred form in the softening function, while I think is categorically used to mark a personal perspective. This suggests that these two structurally similar pragmatic markers can at best be seen as near-synonyms in English. Furthermore, as regards speaker role, students use I think at a similar rate to instructors, but they use I guess more than three times as often. There are clear differences in the degree of use of these pragmatic markers across the three disciplines studied, namely business, humanities and natural sciences.

Erik Castello looks into a set of stance markers in English spoken interaction in corpora of elicited conversations by learners and native speakers. These are drawn from the Swedish and Italian sub-corpora of LINDSEI, representing the spoken language of learners of English, and comparison with the LOCNEC corpus representing the use of native speakers. The study reveals clear differences between Swedish and Italian learners in the use of stance adverbials such
as *probably, fortunately, basically, perhaps, of course*, etc. Swedish learners have been more exposed to English than Italian learners have, and likely received a higher level of linguistic proficiency. This approach, commonly known as Contrastive Interlanguage Analysis, shows itself useful for detecting such differences between groups of learners of different language backgrounds, and it can have important pedagogical implications. In strict quantitative terms, the Swedish learner group display a native-like use of stance markers, while the Italian group differs from it in producing far less such markers. However, this picture becomes more nuanced when considering more detailed qualitative aspects such as repertoire of forms, utterance position and co-occurrence with other pragmatic items.

The article by Sara Gesuato and Katherine Ackerley focuses on *in fact* in English L1 and L2 student writing. It compares L1 data from the BAWE corpus with L2 data from a corpus of essays produced by Italian undergraduate university students. Discerning four distinct meanings of *in fact*—contrast, specification, factuality and support—the authors observe significant differences in functional distribution. In particular, *in fact* occurs almost three times as frequently in L2 essays than in L1 texts from BAWE. This can in large part be ascribed to the much higher frequency of one functional category, namely support, where *in fact* could be paraphrased as ‘for example’, ‘consequently’, ‘that is’, and the like. It appears that *in fact* in the language of L2 users is subject to interference from the Italian false friend *infatti*. This shows that pragmatic markers are challenging to learners, because of their complex multifunctionality.

Annika Labrenz considers the German particle *also* and how it varies functionally across registers and speaker groups including language contact settings. Her study compares the language of bilingual and monolingual speakers of German in Germany and bilingual heritage speakers of German in the US. The comparison shows that the societal status that a language occupies in the larger society, i.e. majority vs. heritage language, has an impact on the distribution of the different functions of *also*. Its discourse functions are described as connector, elaboration/specification, restart, correction and filler. Furthermore, she compares the use of *also* across 2x2 different registers, namely informal and formal, and spoken and written language. Labrenz’ study reveals notable differences along all these dimensions and shows that some discourse functions correlate with communicative settings and speaker groups, while other functions to not, indicating their more general nature.

The contribution of Carla Quinci and Maria Teresa Musacchio looks into the use and translation of the pragmatic marker *well*. They address the challenge of translating this highly multifunctional marker by analysing its use
in a unidirectional parallel corpus consisting of English TED Talk transcripts and their Italian translations. The data is analysed from both an intralinguistic and an interlinguistic perspective, the former analysis focusing on the functional range of the marker in this specific genre, the latter on its translations. Their findings show that *well* is used especially in connection with rhetorical questions, fictitious turn-takings between the speaker and other fictitious addressees, and quotations. Furthermore, *well* is more often translated than not, most typically as variants of Italian *bene/ebbene* but also other forms such as *allora* and *dunque*, and choice of forms tends to correlate with pragmatic function.

Finally, the article by Ilenia Tonetti Tuebben and Daniela Landert focuses on the two planners *uh* and *um* and how they are used as pragmatic markers in dialogues. They provide a contrastive perspective by considering the functions of these forms in both fictional dialogue and authentic conversation. Their data include three different genres: spontaneous conversation, scripted dialogues in television series and dialogues in improvised theatre. In line with previous literature on planners, the authors show that, in terms of function, planners go well beyond signalling hesitation or managing repair. In fact, they are similar to other pragmatic markers in having a range of cognitive, discourse management or interpersonal functions, being used, for instance, in turn management, with face-mitigating functions, etc. The study shows that functions of *uh* and *um* correlate with degree of spontaneity across the three genres studied.

Thus, the seven contributions to this special issue jointly showcase a wide set of ways of doing contrastive analyses of pragmatic markers by means of corpus methods.

**Biographical Note**

Gisle Andersen is Professor of English linguistics at NHH Norwegian School of Economics in Bergen, Norway. His research interests span topics in pragmatics (discourse markers, interjections, politeness, etc.), language contact, lexicography/terminology and Language for Specific Purposes. His recent work has focused on ways in which linguistic items that are inherently pragmatic in nature are borrowed across languages. His recent research has developed the notion of pragmatic/phraseological borrowing, an approach to language contact that investigates how the use of borrowed items is constrained by cultural, social and cognitive factors.