



Roberta Facchinetti<sup>1</sup>

## CORPUS LINGUISTICS AND ELF

### 1. CL and ELF: A Successful Combination

Corpus Linguistics and English as a Lingua Franca (henceforth respectively CL and ELF) have been strongly interrelated since their early stages; this is largely due to the assumption shared by scholars in both fields that empirical analysis of natural text(s) is an indispensable prerequisite for theoretical discussion – and vice versa. Such prerequisite is even more cogent if we bear in mind that the majority of the world's English users are non-native speakers of English and largely exploit it as their working language. Hence, sheer intuition-based methodology cannot give adequate account of the manifold actualizations of a language that systematically trespasses its national borders, imposing deep sociological, ethnographic and cultural considerations on everyday real-world language use.

This common ground of CL and ELF has also led them to share the initial difficulty of being acknowledged and integrated into traditional, mainstream studies of English, within and beyond the field of linguistics. Over the last two decades, however, ELF “has grown from a minority interest within applied linguistics to a major field of study in its own right” (Jenkins 486); the same has happened to CL and both have increasingly challenged the assumptions and tenets of armchair linguists (Svartvik 35).

The first appearance of CL and ELF on the international scholarly arena with properly dedicated journals is relatively recent, the International Journal of Corpus Linguistics being launched in 1996, while the Journal of English as a Lingua Franca (JELF) – the first and still only journal dedicated to ELF – started its publication not earlier than 2012. With JELF, Barbara Seidlhofer and her former co-editors, Jennifer Jenkins and Anna Mauranen, have successfully met a need that had already been perceived by the CL community as well, due to the

considerable and fast-growing field of research that is concerned both with the sociolinguistic significance of English as a lingua franca as a naturally adaptive linguistic development and with its theoretical as well as applied linguistic implications, the latter particularly in the areas of language policy and education. (Seidlhofer, Jenkins and Mauranen 2012, 1)

Indeed, CL and ELF share the same fields of linguistically-related research and CL methodology is regularly exploited in ELF research. This is confirmed by the fact that in 2008, long before launching JELF, the University of Helsinki had already launched the ELFA corpus, a one-million-word database of spoken academic ELF, which has been and still is the hardcore of a number of studies in this field (and not only). One year later, in 2009, the Vienna-Oxford International Corpus of English (VOICE) was also ‘officialized’ as a project, to study how (non-)native speakers exploit English on a daily basis all over the world, in their personal, professional or academic lives. And even more recently, in early 2015, widening up the scope of ELF corpora from spoken to written discourse, the WrELFA corpus of written academic ELF was completed, “drawn from academic genres including the institutional (PhD examiner reports), professional (unedited research papers) and digital media (research blogs).”

Currently, the majority of ELF studies rely on corpora; this is the case, for example, of Cogo and Dewey's research on pragmatics and lexico-grammar in ELF communication, which is based on a corpus of naturally occurring spoken ELF interactions, compiled to identify patterns of language use and their underlying processes. In turn, Vettorel touches on interculturality in her study of an ELF blog corpus compiled to explore how this variety is employed in internationally-oriented personal blogs; her findings show how bloggers

---

<sup>1</sup> Roberta Facchinetti is Professor of English Language at the University of Verona. Her main research interests, which are supported by the use of computerized corpora of both synchronic and diachronic English, focus on media linguistics, lexicography and ESP. On these subjects she has authored, co-authored and edited various books and articles including *Dizionario Giornalistico Italiano-Inglese* (2015), “Newsroom jargon at the crossroads between corpus linguistics and lexicography” (2014), and *A Cultural Journey through the English Lexicon* (2012).



exploit an array of resources to their interactional aims, combining both global and local communicative practices.

The successful combination of CL and ELF has given rise also to Seidlhofer's pioneering publications largely related to interculturality and communication training (Jenkins, Modiano and Seidlhofer; Seidlhofer 2001, 2003); such studies have opened up to further insightful research, which has increasingly highlighted the importance of dynamic, flexible instruction, engaging learners in meaningful communication, so as to achieve mastery in professional discourse, while adapting to the diversified sociocultural communicative contexts.

So, Björkman analyses a corpus of 15 hours of naturally occurring spoken interactions to help understand diversified communicative functions in academic ELF, ranging from explicitness strategies to comprehension checks and clarification requests. Handford and Matous shed light on the intertextual role contracts play in the construction industry problem solving, by analyzing professional ELF interactions between full-time Japanese engineers and contracted Hong-Kongese foremen and engineers belonging to the same joint venture. Finally, Boshier and Stocker rely on a 24,642 word corpus to investigate the current and changing use of English for Nursing Purposes in Taiwan and the experiences of nurses using English in their workplace; their results highlight the importance of professional reading and of communicating with foreign caregivers to improve patient care.

All these studies testify to the highly adaptive nature of ELF to a variety of communicative, professional contexts; this may be even more so online, where ELF is increasingly the norm. Bearing this in mind, the second section of this paper will illustrate a case study on news-related blogs as a potential instance of online ELF (professional) communication, where mainstream journalists, independent experts and ordinary citizens meet on the online ground to provide content on the same event from all over the world.

## **2. News-related Blogs: A Case in Point for ELF?**

Since their first appearance in the late 1990s, blogs have increasingly specialized and diversified their uses, so that now they may be devoted both to personal and to public-interest topics, such as news, politics, economics, sport, health, or hobbies; moreover, they may be authored both by everyday people and by professionals and corporate organizations all over the world. Journalists and media organizations have been attracted more and more to this new format, due to its popularity and flexibility, while news-related blogs run by non-professionals have been heralded as a new type of journalism. Especially within the wide-ranging scope of this latter form of unmediated journalism, it is now increasingly the case for blogs to be written in English by non-natives and for an international audience, which makes it an interesting field of analysis for ELF.

Bearing this in mind, the data of the Corpus of News-related Blogs compiled at the University of Verona and exploited for this case study have been selected on the basis of a topic of international relevance, namely the August 2008 Russo-Georgian War, prompted by the South Ossetia and Abkhazia's declaration of independence from Georgia.

To compile the corpus, the data have been retrieved online through a keyword search on the main countries directly involved in the conflict: 'Abkhazia,' 'Georgia,' 'Russia,' and 'South Ossetia' and the keywords were searched (a) on specialized search engines for blogs, (b) on blog portals/providers, (c) on English language newspaper websites, and finally (d) by following the links to other blogs on any retrieved post. A total of 705 posts were collected, published in 65 different blogs for 424,129 running words, which were categorized as follows:

- Mainstream media j-blogs (MB), published on the websites of news media organizations operating also offline, such as printed newspaper companies and TV or radio news broadcasters;
- Independent j-blogs (IB), portraying themselves as a form of independent journalism, being authored either by freelance professional journalists or by bloggers organized as an online independent media outlet;
- Citizen blogs (CB), published by bloggers who are not professional journalists.



All categories cover speakers of English from a variety of geographical origins including, among others, international MBs with headquarters in the Czech Republic, IBs and CBs blogging from the Caucasus area, and collectively-authored IBs gathering international contributors.

The linguistic analysis of the data show that, independently of the area of blogging, these news-related blogs are diversified according to the writer and the 'channel' through which the blogger writes, be it mainstream, independent, or 'personal'. Specifically, there is a clear difference between MBs on the one hand and IBs and CBs on the other, with reference to a set of interpersonal features, including pronouns and modal verbs. Indeed, CBs and IBs were found to exhibit a lower frequency of first-person singular pronouns I and me and, in turn, a higher frequency of second-person you and first-person plural us compared to MBs.

In CB I and me are almost exclusively used to express the blogger's personal opinion, in collocation with present tense verbs or in such phrases as "it seems to me", "for me", or "I (do) believe / think / bet / can say / (do) wonder / (don't) know" and again "I am" followed by opinion adjectives, like sure, alarmed, glad. In turn, both we and us often instantiate inclusive we/us, generally referring to writer and reader together. Moreover, CBloggers use us mostly to show the bloggers' emotionally sharing of the civilians' plight and their call for universal values. More significantly, you is used extensively by CBloggers and almost exclusively when addressing readers.

In contrast, MBs exhibit a high degree of self-referentiality, in so far as the authors refer more to themselves and less to their interlocutors. These bloggers act often as travel feature writers, and particularly in MB we witness a high frequency of verbs referring to past situations or past actions carried out by the bloggers themselves, who, as is typical of travel features in mainstream journalism, report their personal experiences once back from the conflict zones. Differently from CB, here the use of you hardly shows any indication of interactivity, but is rather a different way of expressing the writer's experience, where this pronoun is used impersonally.

All this would indicate that CBs and IBs are less self-referential and more participative/interactive than MBs. In contrast, MBs exhibit a high degree of self-referentiality, in so far as the authors refer more to themselves and less to their interlocutors.

With reference to modal verbs, the data point to their gradually increasing occurrence from MBs to IBs and mostly CBs and a general tendency of tentative modals to be exploited with more frequency than their present counterparts in CBs, while in MBs they occur largely with their past indicative values. Mostly CBs (but also IBs) appear to favour epistemic possibility modals rather than deontic necessity ones, which occur more frequently in MBs, though exclusively in quotations. Indeed, while MBloggers tend to voice their media institutions and maintain a role of reporters without indulging in the expression of their opinions, but rather reporting other people's thoughts and claims, CBs are more personally-written and express freely their subjectivity. When, however, MBloggers do express their point of view, they favour very tentative modalized patterns, particularly by means of might.

Overall, MB show preference for a high degree of self-referentiality, while CB favour a low degree of self-referentiality, and IB share features of both previous categories, thus lying somewhere in between the two, as a hybrid category.

Apart from this difference with reference to the degree of interpersonality, however, no specific aspect pertaining to that may be peculiar to ELF communication (Vettorel) has been traced in the corpus under scrutiny with great prominence. For example, the following:

- zero third-person-singular marking in present tense verbs,
- this is/there is + plural,
- pluralization of uncountable nouns,
- non-marking of –s plural in nouns,
- interchangeable use of who/which

do occur in the corpus, but only to a limited extent, with no statistical significance. This may lead us to infer that, although a number of these blog posts are run by ELF users, the professional status of journalists on the one hand and the importance of providing properly constructed news-related content on the other prevail over the non-native status of bloggers, thus 'imposing' some kind of standard English to which they comply to.



### 3. Conclusion

The case study illustrated in section 2 testifies to the need to discuss and analyse more and more in detail the diversified instances of English produced online, which all too often are termed ELF, although the boundaries of ELF are still far to be traced; indeed, a number of variables come into place when analyzing it, including the professional field of reference, the sociolinguistic origin of the participants, as well as the topic and the degree of formality of the interaction.

CL has found in ELF research a minefield of study both for spoken and for written communication; indeed, the so far fruitful application of CL methodology to the ELF field can only contribute to strengthen the bond between the two, and head towards further positive developments and new findings in the study of language variation in international, intercultural contexts.

In the long run, the joint venture between CL and ELF might even lead us to rethink our conventional views and understanding not only of English as the world's working language, but also of the very concept and intrinsic nature of global (online and offline) communication.

### Works Cited

- Björkman, Beyza. "An Analysis of Polyadic English as a Lingua Franca (ELF) Speech: A Communicative Strategies Framework." *Journal of Pragmatics* 66 (2014): 122-138.
- Bosher, Susan and Joel Stocker. "Nurses' Narratives on Workplace English in Taiwan: Improving Patient Care and Enhancing Professionalism." *English for Specific Purposes* 38 (2015): 109-120.
- Cogo, Alessia and Martin Dewey. *Analysing English as a Lingua Franca: A Corpus-driven Investigation*. London: Continuum, 2012.
- Facchinetti, Roberta. "Modal Verbs in News-related Blogs: When the Blogger Counts." *English Modality: Core, Periphery and Evidentiality*. Eds. Juana I. Marin-Arrese, Marta Carretero, Jorge Arús Hita, and Johan van der Auwera. Berlin: Mouton de Gruyter, 2013. 359-377.
- Handford, Michael and Petr Matous. "Problem-solving Discourse on an International Construction Site: Patterns and Practices." *English for Specific Purposes* 38 (2015): 85-98.
- Jenkins, Jennifer. "English as a Lingua Franca from the Classroom to the Classroom." *ELT Journal* 66.4 (2012): 486-494.
- Jenkins, Jennifer, Marko Modiano, and Barbara Seidlhofer. "Euro-English." *English Today* 17.4 (2001): 13-19.
- Seidlhofer, Barbara. "Closing a Conceptual Gap: The Case for a Description of English as a Lingua Franca." *International Journal of Applied Linguistics* 11.2 (2001): 133-158.
- Seidlhofer, Barbara. "A Concept of International English and Related Issues: From 'Real English' to 'Realistic English'?" 2003.  
<http://www.serwis.wsjo.pl/lektor/1783/R21%20Seidlhofer%20International%20English.pdf>. Last visited 20/09/2015
- Seidlhofer, Barbara, Jennifer Jenkins, and Anna Mauranen. "Editorial." *Journal of English as a Lingua Franca* 1.1 (2012): 1-3.
- Svartvik, Jan. *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlin: De Gruyter Mouton, 1992.
- Vettorel, Paola. *English as a Lingua Franca in Wider Networking, Blogging Practices*. Berlin: De Gruyter Mouton, 2014.