Features of English as a lingua franca in academia

Anna Mauranen
University of Helsinki
Email-address: anna.mauranen@helsinki.fi

Abstract
Co-constructing communicative effectiveness is often challenging in English as a lingua franca (ELF): speakers have considerably less to go on in terms of shared expectations of cultural knowledge and linguistic norms. A university environment provides a convenient backdrop for sharing at least academic conventions – although these vary more than might be surmised from the uniform labelling of such event types. This paper looks into some discourse and lexicogrammatical features in academic ELF, using ELFA as the database. The data consists of spoken language, which provides direct access to the ways in which meanings are negotiated in ongoing discourse, and the speech events are typically polylogic. ELF discourse requires close cooperation from the participants, which is reflected in its enhanced explicitness among other things. The explicitation strategies speakers display facilitate mutual comprehensibility and contribute to social cohesion within the multi-participant groups. Such strategies also help overcome the potential problems participants might have in dealing with a variety of formal deviations from ordinary English as a native language (ENL). Most of the time ELF bears a very close resemblance to Standard English, but signs of incipient ELF-specific developments are also in evidence.

1 Introduction

English is used as a global lingua franca in an enormous range of domains, from international politics to entertainment, from air traffic to academia, trade, diplomacy and social media. Its non-native speakers outnumber its native speakers, and bend the language to their own purposes. Yet empirical research into English as a lingua franca (ELF) has been thin on the ground, apart from some pioneering studies on mainly a small scale. It is only very recently that ELF has begun to attract serious research, but the change has been fast: currently a surge of interest in lingua franca communication is sweeping the research community. It is nevertheless still true to say that ELF is better known as a topic of debate than empirical research.

During its short history, ELF research has been strongest in business and academia. Both are fields of considerable international weight. They are deeply and inherently international, with their activities spanning the globe, and they have adopted English as their lingua franca, independently of whether any native speakers are present. Both domains also rely heavily on good language and discourse skills. Argumentation and negotiation in demanding situations are prerequisites of success – while faithfulness to English as a native language (ENL) standards of form often takes the back seat.

This paper explores some central features in academic ELF as found in the ELFA corpus. ELFA is the largest existing project on academic ELF, and it has compiled the first large electronic corpus in the field. This paper argues that in order to understand
the use of English in present-day academic communities, it is vital to look at English as a lingua franca. This is the English that is most widely used for scientific and scholarly pursuits, and as it comprises the spoken mode, it is where the language can be expected to undergo particularly fast change. Keeping track of changing English requires observing its use as an international lingua franca. Along the same lines, for a socially oriented interest in the communities that constitute academia, we need to investigate their present reality where internationalism is omnipresent and intertwined with local pursuits in complex ways.

2 Academic ELF as an object of study

Academia is in many ways a typical ELF domain: it is international, mobile and its dependence on English has skyrocketed in the last few decades. Academia is thoroughly dependent on cooperation across national borders and internationally negotiated standards, especially in science, where cutting edge research teams operate in several countries and recruit from anywhere in the world. Lingua francas have characterised academia from the start. Since the Second World War the shift has been to English, which has strengthened its foothold ever since. In contrast to institutions established by international agreements such as the EU, the ASEAN or the UN, which have adopted their common languages by joint decision amidst competing political interests, academia is a ‘natural’, spontaneously arisen domain where the choice of language has taken shape in something like an evolutionary process.

2.1 Academia as an international community

The academic domain in itself is somewhat fuzzy-edged, even though we can point out institutional entities that are central to it: universities, scholarly and scientific publications together with research institutes and research teams undoubtedly constitute its core. Many countries have national bodies such as ministries, government departments, and research councils, which control resources and accreditation systems, but academic institutions tend to maintain a fair degree of autonomy and have not confined themselves to national boundaries. Even though we can point to a number of real social entities that maintain academia, “the scientific community” itself is not a tangible entity; individuals are not in direct contact with all other members of the community, nor is there a set of “members of a discipline” that might be enumerated. Academic communities bear a certain resemblance to what Anderson (1991) called “imagined communities”: members do not all know or meet each other personally, but people nevertheless see themselves as members of such groups, and identify with them.

Academic communities are multi-centred, without one hub to revolve around, but with many overlapping hierarchies, some regional (like EU research funding systems) or national, others discipline-based (many university departments) or research focus based (CERN, Max Planck Institute), some maintained by public resources (most universities in the world), others by private means (private universities and research centres). Criss-crossing over these systems are publishing channels with their own
hierarchies, competition, and enormous influence in the shaping of academia. Research output and publication profiles are in the process of shifting from Anglo-American or G8 nations dominance towards non-anglophone dominance: the language is English, but the research is not. For example, the recent report on research output in the so-called Bric countries (Brazil, Russia, India, and China) shows remarkably sharply rising curves in research output in these countries which, with the exception of Russia, are expected to be on a par or outperform most G8 nations in research output in the next few years (Thomson Reuters report 2010). In the field of Applied Linguistics, Hewings (2002) demonstrated an enormous rise in publications from other than English-speaking countries in English for Specific Purposes over the last couple of decades.

Academic mobility is hardly a new phenomenon either; its roots are in the medieval times when the first universities were founded around Europe (e.g. Koski 1993; Mäkinen 2003). It has seen quieter and livelier times, and over the last couple of decades or so it has undergone another upward swing, with a particularly swift rise in student mobility. The early internationalism of the small, scattered circles of learned men and later internationalism motivated by the nature of research have spread out to study programmes on a vast scale. Internationalism is now urged by national governments and other sponsors of research and higher education. Recent student mobility in Europe has taken the form of exchange programmes and joint degrees, both of which have expanded particularly in the last decade and spread the use of English as a lingua franca to campuses in non-English-speaking countries. Overall, the proportion of students pursuing studies in a foreign country has seen a fast increase, being currently roughly two thirds of those enrolled in international programmes (Wächter & Maiworm 2008).

2.2 Critical views of English in academia

The recent linguistic developments in the university world have not been without their critics: one line of opposition has worried about the fate of small indigenous languages, which may be pushed aside by the ubiquity of English (Phillipson 1992, 2003; Zegers & Wilkinson 2005), others about the standards of learning (Davidsen-Nielsen 2009; Jansson 2008a), and some about the linguistic inequality in the domain of international publication where English is the dominant language (Papers in Carli and Ammon 2007). Collections of papers in Denmark (Haberland et al. 2008; Harder 2009) and in Sweden (Jansson 2008b) have debated the status of local languages versus English, and many of these make policy recommendations towards maintaining a healthy balance among languages (e.g. Ammon 2007; Phillipson 2009; Preisler 2009). One of their major concerns has been that the adoption of English as the language of instruction in higher education may cause even nationally well-established majority languages to suffer domain loss (e.g. Preisler 2009; Kangas Christensen 2009; Davidsen-Nielsen 2009). However, some studies have produced evidence indicating that the use of the local language has not been reduced as a consequence of English-medium instruction (Madsen 2008). Interestingly, the strongest fears about the detrimental effects of ELF-medium study programmes on learning have abated in the course of the last few years (Wächter 2008, see also Maiworm & Wächter 2002), possibly with increasing experience of these programmes in action. Nevertheless, most
of the debate over English is about language policy, not about standards of English or English as a lingua franca – English is tacitly assumed to equal ENL. There are exceptions, though, so for example in the Netherlands intense debate seems to be going on about ENL versus Dutch-influenced academic English (Edwards 2010).

On the whole, while the debates continue in many places, the use of English as a lingua franca has become accepted as a fact of life in European higher education (see e.g. Jakobsen 2009). To quote Wächter:

> While the problem has not changed, Europe's higher education institutions have become accustomed to the communication situation in the international classroom. What once created frustrations is today viewed as a normal condition. (Wächter 2008: 3)

It is indeed important to note the growing sense of normalcy associated with studying in English in Europe. This is obviously reminiscent of the medieval situation where Latin was the lingua franca of learning at a time when it was not spoken as a first language any longer. The crucial point to keep in mind is that the lingua franca in both cases is used for achieving common goals in research and education, not used or learned for the purpose of linguistic or cultural identification with a community that uses it as a national language. It seems that much of the debate surrounding English as a lingua franca is confused about the difference between traditional goals of learning “general English”, still upheld in many places around the world especially in schools, and English as an academic lingua franca.

### 2.3 Investigating academic language

Research into academic English started out with a clear orientation to applications for teaching (see e.g. Swales 1985, Trimble 1985), even though it has moved towards basic research as well. It has served a number of applied domains, mostly teaching but also translation, interpreting, terminology work, and language editing. To be useful to applications, it has been necessary to describe the language of professional communities adequately. The field has long recognised the need to reach beyond the surface manifestations of technical language and to understand the social contexts where the language is used and the purposes to which it is put. This has led to important steps forward. One leap in awareness was made when speaking was taken up as a serious research topic; the crucial role that speech plays in academia began to gain ground more widely with the compilation of electronic corpora on academic speaking in the late 1990s.

English for academic purposes (EAP) research was for a long time very strongly oriented to investigating the written mode, largely because reading was perceived to be important for purposes of study, and writing for the achievement of qualifications and positions. Yet a glance at the multifarious environments of academic language suffices to show that both speaking and writing are at stake. Spoken EAP corpora from the late 1990s reflected a shift in awareness in the teaching and testing of academic language: assessment and appropriate support to large numbers of students was in need of research-based solutions in the domain of speaking just as in writing. The first major corpora, MICASE\(^2\) and T2K-SWAL\(^3\) were compiled in the US, both with a direct interest
in testing and teaching overseas students applying to study in American universities. The BASE corpus followed suit in the UK. The idea was, in line with the study and teaching of written skills, that close observation of what ENL speakers do would yield the best basis for teaching and assessing students with other first languages. In an environment where English is the main language of the university and the community at large, this is not an unreasonable point of departure. However, a more global look at English presents a very different picture. Not only have exchange programmes proliferated, but the language of conferencing is also English (see e.g. Ventola et al. 2002) – or more precisely ELF. The investigations into the language of conferences in Ventola et al. (2002) are an important opening, but it is based on the tacit assumption that despite the thoroughly international nature of academic conferences, the universal adoption of English implies the adoption of ENL models. Yet most of the uses that language is put to in academia are being carried out in ELF all over the world.

The interest value of spoken discourse does not stop at its practical usefulness any more than that of written text; both hold much promise for the scholar who seeks to understand discourse in academia. Speaking is at least as much a key to making sense of academic discourse as writing. It is a crucial ingredient in maintaining social structures, and academic institutions engage in constant talk: we hold lectures, seminars, and consultations as part of our daily routines, we organise conferences, panel discussions and public lectures, we make speeches at graduation ceremonies, and we talk through our endless meetings to maintain and negotiate our institutional relations at all administrative levels of our organisations. We can see this working at the macro-social level as repeated action which creates and maintains social structures, as depicted in Giddens’s Structuration Theory (1984).

At the more micro-social level of interaction, talk plays a crucial role in socialising new generations to academia and professions: we pass on explicit and tacit understanding of the norms of academic discourse, and of preferred ways of talking. The role of talk in academia can be likened to Gilbert & Mulkay’s (1984) ‘contingent’ repertoire that scientists engage in informally and behind the scenes, where mundane matters of serendipity and luck get talked about along with power, interpersonal relations, and struggles over position and financial resources (see also Mauranen 1999, 2001). This can be pitted against the ‘empiricist’ repertoire of written presentation, where reports of experiments, results, and their theoretical implications are presented in an impersonal and detached manner. There is more interpersonal engagement than first meets the eye in research articles, as research into EAP has been eager to show over the last fifteen years, but what gets written and published is still far from the intimacy and freedom of the spoken word.

ELF research has several reasons for its interest in academic language. Academic language exerts considerable normative influence on standard languages. We are used to thinking of the “educated native speaker” as the ideal speaker that language standards are modelled on, and university education is of course the institution that generates such speakers in real life. In view of the way English is developing in the world, the target speaker may not be a native speaker in the future, but probably educated all the same. The emphasis in higher education on English all over the world brings English into contact with a very large number of the world’s languages, as Thomason observed some years ago (2001). Academia becomes an important source of ELF features, because language contact is a major factor in bringing about linguistic change.
The academic communities that use English as their lingua franca vary widely in their objectives, duration and location, ranging from research teams to Master’s programmes and short exchange programmes. International research teams with a global reach may be funded for a few years at a time, and located in their respective institutes, in one location, or divided between different arrangements. Some research centres recruit internationally but are permanently located in one place (Max Planck Institute; CERN). Doctoral students and postdoctoral researchers may spend a few years in such teams, changing places as their careers take shape. International Master’s programmes have been mushrooming and shorter student exchange programmes at undergraduate or graduate level have become routine particularly in Europe. All this mobility and its associated multi-layered networks contribute to complexity in the relevant linguistic settings.

Such multi-layered mobile networks are typical examples of the kinds of links known as ‘weak ties’ in social network theory (Granovetter 1973). As Milroy and Milroy (1985) suggest, linguistic innovators are likely to be individuals who are in a position to contract many weak ties. Thus, networks of weak social ties tend to favour linguistic innovation and dialect levelling, in contrast to networks where strong ties dominate. To quote L. Milroy:

Networks constituted chiefly of strong (dense and multiplex) ties support localized linguistic norms, resisting pressures to adopt competing external norms. By the same token, if these ties weaken[,] conditions favourable to language change are produced. (L. Milroy 2002: 550)

Academics working and studying in ELF-medium environments are likely to have linguistic innovators among them, and their mobility and multiple networking provide good platforms for innovations to diffuse. Many of the repeated occasions of talk that maintain academic structures involve international speakers of high status, whose influence on what is regarded as appropriate academic language is likely to be strong. They may affect other ELF speakers’ notions more than those of ENL speakers, but given their relative numbers, the flow of new forms may be considerable.

ELF is always immersed in multilingual environments with plurilingual speakers. This is very clear in universities: the matrix culture is frequently one where English has no major status, and speakers invariably possess other language resources. The first language for most participants is other than English, and typically speakers have some knowledge of third, fourth, and nth languages as well. The presence of other languages may easily get underestimated because English is the working language. The emerging picture from studies that have charted the use of local languages in English-medium study programmes shows that local languages do play a role even in cases where the programmes are not of a long duration (Shaw et al 2009; Wächter & Maiworm 2008, Hynninen this volume, see also the DYLAN project5). These are highly complex language contact situations (see e.g. Mauranen 2005), in which it is not realistic any longer to hold on to linguistic notions that derive from the dominant ENL assumption of English-speaking countries.

Despite the straightforward aim of the first EAP corpora to analyse native speakers’ language to provide a model to non-native speakers, changes had already taken place in the conceptualisations of English by the time the corpora were completed. Budding awareness of English as an international lingua franca, rather than
as a language exclusively belonging to its native speakers in the ‘core’ cultures, was spreading as a result of contributions from scholars like Widdowson (1994), Jenkins (2000, 2007) and Seidlhofer (2001, 2004). A decade later there is a much wider awareness of the implications of the international character of the academic world. The first corpus of academic English spoken as a lingua franca, the ELFA corpus\(^6\) started in 2001, close on the heels of the first ENL speech corpora. Interestingly, ELF has reversed the usual progression of research from written to spoken language; although another ELF corpus, VOICE\(^7\), has been completed in Vienna and others are starting out, there is no written database of English as a lingua franca as yet.

Insofar as the general goal of research is to understand major socially relevant uses of English for academic purposes today, and to develop applications on the basis of this understanding, the issue of ENL is neither central nor indispensable.

3 The ELFA project: ELFA Corpus and SELF

English as a lingua franca was a virtually unexplored territory at the turn of the millennium, and academic ELF a completely white spot on the map. To make a start, the compilation of an academic ELF corpus was begun at the University of Tampere in 2001. Progress was slow until 2004, when the ELFA project\(^8\) received major funding, and soon afterwards moved to the University of Helsinki. The recordings were also extended to two technological universities. The ELFA corpus, completed in 2008, now consists of 1 million words of spoken English in university contexts. This section gives a very brief overview, but its more detailed description together with its compilation principles can be found in Mauranen and Ranta (2008) and Mauranen, Hynninen and Ranta (2010).

Compilation criteria in compiling ELFA have been generally ‘external’, that is, identification of the prominent genres of the discourse community on a social, not language-internal basis. The speech event types reflect the self-understanding of the discourse communities of their activities, thus resulting in ‘folk genres’ recognised in faculties, departments, or conference organisers, and event labels like ‘seminar’ and ‘lecture’ appeared across institutions.

The basic unit of sampling was the ‘speech event type’, a looser notion than ‘genre’, and perhaps more appropriate to folk genres. The requirement of using English set its limitations to event types; in an environment where English is not universally used, we cannot assume the same overall event type selection as in a single-university approach in an ENL context. ELFA included more than one university to get a better-balanced selection of disciplinary areas. A ‘right’ balance is unlikely to exist, and the dividing lines would be different if different universities were chosen (Mauranen (2006a).

It is important to make a clear distinction between an ELF corpus and a learner corpus such as the ICLE\(^9\). Obviously, a number of linguistic features are shared between learners and ELF speakers, and people can alternate in both roles even during the same day. There are still very strong reasons for keeping learner and speaker events apart, because social, cognitive, and interactive parameters shift in important ways when we move from one of these event types to another. If authentic language use is the target, the situational parameters must reflect this as closely as possible. A crucial principled difference between learner and ELF corpora is also that
learner corpora keep a close eye on the proficiency level of the learners in the corpus or any section of it. This makes sense in view of the questions asked of learner corpora, which often relate to L2 development. For an ELF corpus that would be counterproductive, because ELF is used between speakers of varying proficiencies. Attempts to control for proficiency would miss out on an important situational parameter, the natural asymmetries among speakers. The corpus thus consists of naturally occurring situations where English is the real lingua franca, where participants may have different proficiencies, do not share an L1, and where they are not in an English language teaching class.

Some language-internal criteria were also used: maximisation of language background variation and favouring dialogic events. Both were deliberate choices, made with the main research goals in mind.

In addition to the corpus, it was felt that a more qualitative and in-depth approach was also needed to glean data from participants themselves, and to supplement the research approaches that a corpus permits. The SELF project (Studying in English as a Lingua Franca) at the University of Helsinki was begun in 2007 to explore the speaker’s perspective with an ethnography-informed orientation. Together, ELFA and SELF represent a holistic approach to English for specific purposes (ESP). A microanalytic approach was adopted in SELF, collecting data from interrelated speech events, interviews, field observations, and written documents. Negotiations of meaning and practices of mediation in university discourse have emerged as one of the central issues (Hynninen in press and this volume), as are the realities of the multilingual context. Some of the work has extended to vocational schools (Lammi 2009 and this volume). The research group incorporates work in the Aalto University, exploring lectures in forest products technology, particularly their interactional features (Suvinitty 2008 and this volume) and sounding out lecturers’ thoughts and attitudes about English in higher education, including their own English (Pilkinton-Pihko this volume).

4 Some discourse features in academic ELF

A prominent discourse feature that comes up in ELF is speakers’ tendency to use communication strategies that are emphatically explicit (Mauranen 2007; Dewey 2007). This is not far from what Blum-Kulka (1986) termed “explicitation” in translators’ work; explicitation has since been established as one of the most firmly supported universal features of translations, independent of the languages involved (e.g. Mauranen 2000, 2007; Olohan 2004). It is therefore quite likely that explicitation is a natural response in individuals operating in a language contact situation. After all, translation, interpreting, and speaking a foreign language all involve simultaneous bilingual processing. Since this is such a prominent strategy, I shall illustrate it with some typical categories here: negotiating topics, metadiscourse, and self-rephrasing.

4.1 Negotiating topics

A common feature of spoken discourse is a pattern where the lexical subject of a clause is fronted so that it precedes the clause, and the subject is then represented by
a coreferential pronoun within the clause itself: **one of my friends she** tried to enter to the university. This pattern is known as ‘left dislocation’ in standard grammars, and it highlights the subject by bringing it forward. It is likely to improve clarity in spoken discourse in that it helps the interlocutor keep track of what the speaker is talking about. It is a common pattern in many vernaculars, a perfectly normal feature of spoken English, but normally regarded as nonstandard in written language. Its interactive role has been recognised in spoken language research. Carter and McCarthy (2006) call it ‘header’, Ford et al (2003) talk about ‘negotiating referent’, observing that this pattern is used to ensure that interlocutors have the same topic in mind before going on. I have chosen the term ‘negotiating topic’ (Mauranen 2007), because it is the topic referent that is at stake here. McCarthy (1998) suggests that using a ‘header’ is an act of consideration to the listener. If this is the role the pattern plays in ENL speech, it is easy to see how useful it can be in ELF discourse; it is likely to facilitate comprehension by helping ensure that speaker and hearer have the same topic in mind, and it serves an interactional function in indicating willingness to cooperate. Cooperativeness has been found to be highly characteristic of ELF (e.g. Karhukorpi 2006); just like in real-life L2 encounters (Kurhila 2006). This pattern is common enough in ELFA, with some examples below (Example 1):

(1)

\begin{itemize}
  \item the main point of these it was related to the local adaptation
  \item craftsmen who paid enough tax they were given the right to vote
  \item these lions they don’t eat each other
  \item because the russians they didn’t care er or because
\end{itemize}

The pattern was used for reference to individuals (**one of my friends she** tried...), but very often the referent was a group (**the Russians, craftsmen**), an institutional body (**the parties**), a physical object (**these kind of kidneys**), or an abstract object (**these different layers of identity**). It is a flexible strategy for introducing or maintaining topics in ongoing discourse, and a great deal of anecdotal evidence supports the interpretation that speakers using a foreign language make particularly good use of it. It is easy to appreciate its usefulness under conditions where sharedness of both linguistic and world knowledge cannot be taken for granted.

### 4.2 Metadiscourse

An important kind of explicitness is that which takes the form of discourse reflexivity, or metadiscourse (**what I would like to ask you**). I take metadiscourse to be discourse about the ongoing discourse, and such a fundamental feature of human language that it is a vital, perhaps universal part of all language (Mauranen 2010). I argued earlier (1993) that comparatively small and homogeneous cultures, such as the Finnish culture largely was twenty years ago, have less need for such explicitness than larger and more heterogeneous cultural formations, such as the large, varied and somewhat amorphous Anglo-American. Following the same logic, we could posit that in the heterogeneous and highly complex lingua-cultural encounters using ELF, metadiscourse is a serviceable means of coping with the exigencies of the situation. This is supported by our data: metadiscourse is ubiquitous in all genres in ELFA.
It is interesting to note that the overwhelming majority of research into metadiscourse (e.g. Crismore 1989; Mauranen 1993; Hyland 2005; Ädel 2006) has analysed written language, although it is characterised as projected dialogue between the writer and the reader. Some spoken uses are very similar to written ones, and readily found in monologues such as lectures and conference presentations (Examples 2 and 3):

(2)  
"oh here i have some sort of a summary of things going on within the project"

(3)  
"<S2> er erm so, er i'm talking about er (<NAME>) paper it's about democratisation in poor countries and er she talked just on er tanzania because she live in tanzania so. i will give you the introduction, er when she said…"

In dialogues, metadiscourse assumes a far broader range of roles, especially in discourse management. This is a crucial aspect of successful discourse, orienting either to organising text or organising the interaction (for this distinction, see Sinclair and Mauranen 2006). The examples below from multi-party discussions show that metadiscourse introduces topics (4), explicates the speaker’s relation to the contents of what they are saying (5), and indicates that speakers want to change the course of the discussion (6).

(4)  
"<S1> what about we've we've talked about groups before and for example in spanish…"

(5)  
"<S12> cause there's a difference betwe- in [the rules if it's] <SU> [yeah @@@] </SU> if it's like, you know i think i'm not too familiar with the differences but i think it refers to that i don't know what they're allowed to do but some things are not allowed…"

(6)  
"<S1> erOkay before we go to the next topic, i i think that, in a way the question <NAME> made what made you study or be- become interested on this issue it is a relevant question cause this your topic leads us a bit further to more general <S2> yeah </S2> discussion about human rights or in general whether we can…"

Example (6) is particularly interesting in that it highlights a characteristic of spoken discourse. Not only does it introduce a topic, one of the most widely recognised functions of metadiscourse, but it also changes the direction of the ongoing discourse as a response to something that another speaker has brought up during the discussion. There is of course nothing extraordinary about such ‘other-orientation’, but it is not covered in the metadiscourse literature, which has not looked much into spoken language at all, and even less into dialogue.
4.3. Rephrasing

Linguists have shown different, even polarised responses to repetition in speaking. On the one hand, it has been perceived as a deplorable defect, something that speakers do because the pressures of online speaking prevent them from producing completely smooth output that would look essentially like writing. This attitude lingers on even in respectable descriptive grammars, such as Biber et al. (1999):

> We may refer to the grammar of speech as ‘dynamic’, in the sense that it is constructed and interpreted under real-time pressure, and correction or reformulation is possible only through hesitations, false starts, and other dysfluencies. (Biber et al.1999: 1066)

The effect of formulations of this kind is to render spoken language somehow inferior to writing, even though this probably was not the writers' intention. Despite general agreement that the spoken mode is the primary mode of language, the idea that speech is inherently ‘dysfluent’ is surprisingly widespread, and it has affected the way second language speech is regarded. Yet the same kinds of repetitions and rephrasings are observable in L2 and L1 speech — even though in proportionally greater quantities in L2 (Mauranen forthcoming).

A positive view of repetition was promoted in Tannen’s (1989) influential book *Talking Voices*, and repetition for rhetorical effect and clarity has been much investigated in academic writing. The use of repetition and rephrasing in monologic presentations seems to play an important role in making lectures more comprehensible (Suviniiitty in preparation). However, it does not seem to be a function of greater language proficiency or dependent on the speaker’s L1 (Mauranen 2009).

By saying the same thing more than once, speakers can buy time (people um of of of higher incomes), do self-repairs (*and that takes place in two ways or has taken place*, and...), rephrase what was already said to make it clearer or more to the point, or perhaps more effective (*cultural approach to the history of technology will bring us new insight will will will enable us to understand the developments...*). Repetition and rephrasing is also used to organise discourse even in relatively short turns:

(7)

```markdown
<S6> yeah but these salaries are not public so you cannot compare <S1> yeah
</S1> you you cannot compare in the same place you have to to make a meeting and to compare each er salary sheet to to see whether the people don’t know what can they do it’s not public.
```

Most of these and other similar uses enhance clarity and explicitness in monologues and dialogues alike. Where dialogue makes a separate departure is in other-repetition (allo-repetition), which connects explicitness to cooperativeness. In the extract below (8), six different speakers contribute to working out what kinds of occupational diseases were found in a workplace if any, and the repetition of occupational disease(s) helps maintain coherence in the discourse, as do its superordinate (*occupational related disorders*) and a kind of subcategory (*injuries*) too.
In addition to discourse coherence and clarity, allo-repetition also helps construct a cooperative atmosphere with participants indicating that they are contributing to the same topic, making sense of it together. Considerable cooperativeness has often been attributed to ELF (e.g. Cogo and Dewey 2006, Hüttner 2009, Karhukorpi 2006, Klimpfinger 2009), and frequent other-repetition appears to be an important element in achieving it. Constructing and maintaining a cooperative atmosphere is important in dealing with argumentative situations where different views and disagreement are part of the normal state of affairs. In the following, rephrasing the same contents with roughly synonymous expressions (*sometimes, not always, in general, on average*) by both parties seems to ascertain an underlying goodwill despite incomplete agreement:

A final example (10) of collaborative rephrasing illustrates the negotiation of language that takes place in the co-construction of linguistic group norms. Such negotiation is undoubtedly a normal part of communities of practice, but notably prominent in ELF.

Repetition as well as the other types of explicitation illustrated above, topic negotiation and metadiscourse, are all important ingredients in achieving mutual
comprehensibility. They contribute to not just discourse coherence but also social cohesion, the achievement of which is of particular concern in situations where a variety of linguistic and cultural backgrounds come together.

5 Lexicogrammatical features of academic ELF

The area where ELF is in sharpest contrast with Standard English is in lexis and grammar. While it is traditionally accepted that the less speakers deviate from Standard English, the better their chances of successful communication are, we already saw in the previous section that communicative success may essentially be achieved through discourse features. ELF speakers seem relatively little bothered about grammatically non-standard forms, and a few types seem to be found in ELF talk repeatedly (see for example Seidlhofer 2004). Hülmbauer (2009) suggests that grammatical deviations from Standard English are not just something that can be overlooked, but may in fact be an integral part of communicative success; simplification of forms, or drawing on analogies from other shared languages can be an asset rather than a hindrance.

Among the commonest non-standard features are articles and prepositions. Articles are often not used where Standard English would have them (if you don’t get like white collar job; in Russia we have other system from here), inserted where Standard English would not have them (had a big importance) or just chosen differently (cause an deposition of iron). Prepositions (the iron may accumulate to other organs; you study about the name; insisted for full independence) also often deviate from their conventional ENL uses. Similar lack of accuracy can be seen in plural noun endings and third person singular verb endings (respect for human right; she live in Tanzania).

Certain tendencies to regularise morphology can be seen: irregular verbs get inflected according to the regular paradigm (i have just showed you), and turning uncountable nous into countables (offsprings; researches). The productive resources of morphology get drawn on to generate forms that go beyond the limits of conventional usage (youngness; relativeness; maximalise).

We cannot dismiss such repeated findings as arbitrary mistakes. Some of course are just passing slips of the tongue, which tend to be ignored by conversationalists just as they are in L1 communication. It is not the comprehensibility of the anomalies that requires an ELF-specific explanation, but their recurrent features. Morphological overproductivity, for instance, seems to ensue from the nature of morphology, which is highly prolific and held back by convention rather than rule. Academic language makes good use of this productivity in constantly creating new terminology. We tend to understand new formations by analogy – so for example femininish appeared in MICASE without any noticeable reaction from interlocutors, probably because -ish is highly productive.

Ranta (2006, 2009, and in preparation) has observed some recurrent features of verb syntax, for instance that the -ing form of the verb is used more liberally than in Standard English, and that agreement, hypothetical if-clauses, and word order also often show non-standard features. The features that she found are shared with vernacular varieties and postcolonial varieties of English. They are thus non-standard, but very English.
Phraseological units in ELFA are frequently non-standard, as illustrated by this list:

(11)

you have to work on it but it’s worth the while
the population failed in that time
factors that you need to take in account
there is quite much problems
i think that on the end somehow the law has had some results
clashes with soviet union was in the matter of few inches

Again, despite the approximate rather than accurate forms in standard language terms, comprehensibility is not adversely affected. Where comprehension mostly seems to break down is over pragmatic matters rather than lexicogrammatical accuracy (Mauranen 2006b). Nevertheless, phraseological units have been considered a final stumbling block for L2 learners, since the inaccuracy seems to persist even at very high levels of proficiency (e.g Nattinger and DeCarrico 1992; papers in Meunier and Granger 2008). Yet if we compare the five most frequent trigrams, or lexical bundles of three in ELFA and MICASE, the similarity is striking:

(12) Top 3-grams

<table>
<thead>
<tr>
<th>ELFA</th>
<th>MICASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I don’t know</td>
<td>1 I don’t know</td>
</tr>
<tr>
<td>2 I think that</td>
<td>2 a lot of</td>
</tr>
<tr>
<td>3 one of the</td>
<td>3 one of the</td>
</tr>
<tr>
<td>4 a lot of</td>
<td>4 a little bit</td>
</tr>
<tr>
<td>5 you have to</td>
<td>5 you have to</td>
</tr>
</tbody>
</table>

As we can see, even the rank order is nearly the same, with only two cases that differ: I think that, which is the second most common in ELFA, ranks 12th in MICASE, and is thus not far off. The odd one out is a little bit, which only takes the 43rd position in ELFA. Despite this similarity, there are differences as well, but it is easy to lose sight of the similarities that may not be particularly eye-catching, and therefore it is occasionally healthy to remember the limits of the variation.

To gauge the potential for variation and similarity in ENL and ELF, I take up an example from MICASE: let me say a few words about the solution. It is a common metadiscursive pattern in academic discourse: let me with a communicative verb. These were found in the first 600,000 words of ELFA (see Example 13):

(13)

on information and not on communication let me explain why. I do believe that very very troublesome. because. let me give you one example. the the results <S2> mhm <S1> mhm... yeah eh let me ask you first what would ens if you if you only talking about eh oh let me correct myself not only but

Instances of these were found in MICASE as well, with one exception: let me correct myself. One might be tempted to say that it is not native-like, even if not a
breach of any structural constraint that could be pointed out. Searching MICASE for verbs in the verb slot reveals that there is a set of clearly preferred communicative verbs: TELL, GIVE, ASK, SAY, LOOK, TAKE, and SHOW. However, in addition to the preferred verbs, a number of single occurrences also came up in MICASE:

(14) recounted in these tales. okay. at this point, uh let me clarify at this point most ... so let me just elaborate a little bit and then we'll get to false consciousness. said enough um let me make my comment which is that in in natural science imizes the total value. let me just stop you for a minute Rob mhm so, just okay. um did we talk about yesterday, um... let me backtrack here. i can see any other questions? if not let me leave this model, and i wanna start talking okay so let me just, let me be episodic. i me- w- one of the things that is pective, let me just give you a a forty-five second sketch of Einstein's early life.

Some, like CLARIFY or ELABORATE, would not intuitively strike one as in any way unusual. Others, like BACKTRACK, BE episodic, or LEAVE this model sound less routine-like, and the last example delightfully creative. Each of these nevertheless just occurred once. The point I am making is that although corpora yield interesting and useful insights into typical usage, they can equally well reveal variation beyond the frequent or preferred, and show unique solutions that speakers come up with. Thus, if the ELF speaker with let me correct myself hit upon a non-routine solution to an online communicative need, while at the same time fitting it neatly within a conventional frame, several ENL speakers did exactly the same.

To continue with this little example (let me say a few words about the solution), it harbours another common phraseological unit. A MICASE search for words about yields five instances (15).

(15) models, gonna say a few words about why this problem is difficult let me say a few words about the solution. um... in my to conclude with a, few words about Hilbert and his problems. Hilbert just a few words about the planning committee um who to solve, and a few words about the solution, and then towards

A few words about is a fixed cluster, with no variable parts. This cluster appeared in ELFA as well.

(16) can be delay also in that er then a few words about the restrictions, you have parts of some some places er first few words about the background er this

In addition, ELFA contained variants to the formula:

(17) thanks well can you say couple of words about the sources or the reference
But more interestingly, ELFA displayed an alternative, preferred pattern of its own, as shown here:

(18)

you have you have done, then some words about the thickness of the book then if we open book er say some words about the nomenclature that you you are (xx). (xx xx) okay er some words about the state of the art in huma verified with experiments and some words about er the aspects the contents influences of er the past, now some words about the problem (xx) preservation er i'm at last i'm going to, say some words about the library information so way yeah most important then some words about app- theoretical approach

In view of the last three examples (16–18), it looks more like a partly variable pattern, and intriguingly there is a new preference for a pattern that was not there in the ENL data. It seems, then, that ELF use also generates pattern preferences of its own, possibly of those kinds that people from a large number of L1 varieties feel comfortable with. These are likely to be of the same kind that you see in dialect contact – unmarked, simple, ‘natural’ (see e.g. Thomason 2001, Trudgill 1986).

6 Conclusion

Insofar as the ESP research community seeks to understand present-day academic speaking in English, it is vital to take on board English as a lingua franca. To capture its contextualised use in real life, ELF is a better representative than native English. Conversely, for a linguist interested in ELF, academia is an excellent vantage point for exploring developments in the language: it is one of those influential domains that have widely adopted English as their common language, and it is one where international communication characterises the domain across the board, as we have seen above. With the recent expansion of internationalism, the spread and influence of English used by second-language speakers can be expected to keep growing.

Methodologically, this paper like the projects it is part of suggest a holistic view of ESP, which includes both qualitative analyses of discourse in context and a large-scale view that corpus research makes possible. Both are crucial to understanding changing English in the fields we are investigating. In analysing spoken discourse in context, it is crucial to gather data from dialogic interaction. To capture linguistic self-regulation in social groupings, in this case communities of practice consisting of complex networks that involve mobile multilingual participants, interactional discourse is where language and norms can be seen in the making. Processes central in this, such as accommodation and the co-construction of intelligibility can only be seen in interaction. Participants in ELF encounters need to negotiate meaning and manage interaction like any speakers in conversation, but in addition, they must negotiate language to a far greater degree than speakers do in monolingual situations. Interaction also gives a close-up view on cultures in contact; the immediacy of interactional data is inaccessible to comparisons between cultural artefacts such as texts or culture-specific lecturing styles, which are necessarily further removed from the actual contact between individuals.
The perspective that a large corpus gives to the analysis of change and success in language opens vistas to relevant units of meaning and emerging ELF patterns that may well be new to ENL but work well in the international context. These pave the way to shifting views of ‘core’ phenomena in language, away from the traditional units of form and sense. They can really only be accessed by corpora large enough to show repeated patterning of elements.

Notes

1 For the ELFA Corpus of English as a Lingua Franca in Academic Settings, see http://www.helsinki.fi/elfa/elfacorpus.

2 For the MICASE corpus, see http://micase.elicorpora.info.

3 For the T2K-SWAL corpus, see http://www.ets.org/Media/Research/pdf/RM-04-03.pdf.

4 For the BASE corpus, see http://www2.warwick.ac.uk/fac/soc/al/research/collect/base.

5 For the DYLAN project, see http://www.dylan-project.org/Dylan_en/index.php.

6 For the ELFA corpus, see http://www.helsinki.fi/elfa/elfacorpus.

7 For the VOICE corpus, see http://www.univie.ac.at/voice.

8 For the English as a lingua Franca in Academic Settings project, see http://www.helsinki.fi/elfa.

9 For the ICLE corpus, see http://www.uclouvain.be/en-cecl-icle.

10 For the Studying in English as a Lingua Franca (SELF) project, see http://www.helsinki.fi/elfa/self.

11 I use SMALL CAPS to refer to a lemma, i.e. a lexical item and its associated inflected forms.

Transcription conventions

The transcription conventions found in this paper follow the ELFA Transcription Guide (http://www.helsinki.fi/elfa/ELFA transcription guide.pdf):

Speaker codes:

<S#> Identified speaker

<SU> Unidentified speaker

<BS#> Bilingual speaker of English
Symbols:

<S1> </S1> Utterance begins/ends
(text) Uncertain transcription
(xx) Unintelligible speech
te- Unfinished utterances
@@ Laughter
. Brief pause while speaking, 2-3 sec.
. Pause, 3-4 sec.
[text] Overlapping speech (approximate, shown to the nearest word, words not split by overlap tags)

<NAME> Names of participants
… Omitted text from transcription

Corpus references

BASE http://www2.warwick.ac.uk/fac/soc/al/research/collect/base
ELFA http://www.helsinki.fi/elfa/elfacorpus
MICASE http://micase.elicorpora.info
VOICE http://www.univie.ac.at/voice

References


Granovetter, Mark (1973) The strength of weak ties. *American Journal of Sociology*, 78 (6), 1360–1380.


Hülmbauer, Cornelia (2009) "We don’t take the right way. We just take the way we think you will understand" – The shifting relationship between correctness and effectiveness in ELF. In Anna Mauranen & Elina Ranta (Eds) *English as a Lingua Franca: Studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing. 323–347.


Hynninen, Niina (this volume) "We try to to to speak all the time in easy sentences" – Student conceptions of ELF interaction. *Helsinki English studies*, 6, 29–43.

Jakobsen, Anne Sofie (2009) "Ellers er det bare lige ud af landevejen" En intervewundersøgelse af ti underviserses holdninger til og erfaringer med
engelsksproget undervisning ved det Biovidenskabelige Fakultet (LIFE), Unpublished MA thesis, University of Copenhagen.


Lammi, Riikka-Liisa (this volume) Backchannels and repetition in ELF in a hairdressing setting. Helsinki English studies, 6, 118–131.


© A. Mauranen 2010


Preisler, Bent (2009) Complementary languages: The national language and English as working languages in European universities. In Peter Harder (Ed.) *English in


Suviniitty, Jaana (this volume) Lecturers’ questions and student perception of lecture comprehension. Helsinki English studies, 6, 44–57.


Anna Mauranen is Professor of English at the University of Helsinki. Her publications focus on English as a lingua franca, corpus linguistics, academic rhetoric, and modelling spoken language. She is currently running a research project on spoken academic English as a lingua franca (the ELFA project). Her major publications include English as a Lingua Franca: Studies and Findings (ed. with Ranta 2009), Linear Unit Grammar (with Sinclair 2006), Translation Universals – Do They Exist (2004), Cultural Differences in Academic Rhetoric (1993).