UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Multilingualism in English as a lingua franca : Flagging as an indicator of perceived acceptability and intelligibility

Hynninen, Niina Margareta; Pietikäinen, Kaisa Sofia; Vetchinnikova, Svetlana Nurmi, Arja; Rütten, Tanja; Pahta, Päivi

2017

http://hdl.handle.net/10138/305626

Hynninen, N M, Pietikäinen, K S & Vetchinnikova, S 2017, Multilingualism in English as a lingua franca : Flagging as an indicator of perceived acceptability and intelligibility. in A Nurmi, T Rütten & P Pahta (eds), Challenging the Myth of Monolingual Corpora. Language and Computers, no. 80, Brill , Leiden, pp. 95-126. https://doi.org/10.1163/9789004276697_007

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

Final draft of published paper:

Hynninen, Niina, Pietikäinen, Kaisa & Vetchinnikova, Svetlana (2017) Multilingualism in English as a lingua franca: Flagging as an indicator of perceived acceptability and intelligibility. In Arja Nurmi, Tanja Rütten & Päivi Pahta (eds) *Challenging the myth of monolingual corpora*. Leiden: Brill Rodopi. 193–212.

Multilingualism in English as a lingua franca: Flagging as an indicator of perceived acceptability and intelligibility

Niina Hynninen, Kaisa Pietikäinen and Svetlana Vetchinnikova

University of Helsinki

Abstract

Lingua franca interactions are inherently multilingual; speakers' first and other languages are always implicitly present in such interactions. However, the extent to which speakers resort to their multilingual resources depends on how acceptable and understandable they expect these resources to be with particular interlocutors or within particular speech situations. In this chapter, we draw on the multilingual practices observable in the Corpus of English as a Lingua Franca in Academic Settings (ELFA 2008) to generalise to other practices which can be treated as similar from the perspective of perceived acceptability and intelligibility. With this aim in mind, we focus on elements tagged as <FOREIGN> in the corpus – that is, elements recognised as code-switches by the corpus compilers – and investigate the lexico-grammatical regularities in their co-text. Using both corpus linguistic and micro analytic methods, we show that speakers use these regularities to flag potentially problematic items, including but not limited to code-switches. In this way we extend our scope beyond elements which have traditionally been regarded as multilingual. We expect this research to contribute to the development of corpus linguistic methods in studying elements perceived as "foreign" and thus potentially problematic in the discourse.

1. Introduction

The majority of classical sociolinguistic studies revolving around the phenomenon of code-switching have investigated language alternation in bilingual communities, where more often than not, the

speakers' linguistic backgrounds are relatively similar. Although this research has contributed tremendously to our understanding of languages in contact, until recent years, there seems to have been a bias towards whose code-switching is valid for scrutiny. For example, in 1995, Auer claimed that "new" European bilingual communities of migrant workers were problematic for the inquiry of language negotiation because they were "too young and culturally unstable to have developed shared norms of language choice" (Auer 1995: 127). Recently, however, we have seen a rise in the number of sociolinguistic analyses of code-switching in intercultural settings where the speakers resort to a lingua franca (e.g., on English, see Pölzl 2003; Pölzl and Seidlhofer 2006; Klimpfinger 2007, 2009; Cogo 2012; Pietikäinen 2012, 2014, in press; Turunen 2012; House 2014; Mauranen 2013). Such cross-linguistic environments are in fact especially fruitful for the research of language alternation, precisely because the speakers may have contrasting understandings of the sociolinguistic norms at play, and thus have to negotiate these situationally.

A possible explanation for the exclusion of speakers from differing language backgrounds from code-switching research in the past (with the exception of second language classroom interactions, e.g., Fröchlich, Spada and Allen 1985; Luk and Lin 2006) may be that researchers commonly regarded e.g., nonce borrowing (Poplack, Sankoff and Miller 1988), crossing (Rampton 1995), language negotiation and transfer (Auer 1995) and code-mixing (i.e., intra-sentential code-switching; Muysken 2000) as not belonging under what they defined as code-switching. However, there now seems to be a tendency in sociolinguistic inquiry to move away from this differentiation and to treat code-switching as an umbrella term for linguistically speaking all types of uses of two or more languages within a single interaction, or to refer to fluid mixing of languages as translanguaging (García and Li Wei 2014).

In this chapter, we investigate code-switching practices in English as a lingua franca (ELF) interaction by looking at the co-text of elements tagged as <FOREIGN> in the Corpus of English as a Lingua Franca in Academic Settings (ELFA 2008). We first review studies which have examined code-switching in ELF in section 2. In section 3, we explain our phraseological approach to the analysis of the common co-text of code-switching, and in sections 4 and 5, we describe the data and methods in more detail. In section 6, we analyse the co-text of code-switching using corpus linguistic methods and examine the regularities in its lexico-grammatical patterning. We suggest that the patterns identified are used by speakers in ELF interaction to flag items such as code-switches that are potentially problematic for the interaction. In order to illustrate this, in section 7, we provide a qualitative analysis on how and why the flagging is used, drawing on discourse and conversation analytic approaches. We argue that the extent and type of flagging around code-switched items is related to the acceptability and/or intelligibility of the code-switch anticipated by the speaker within the given communicative situation. Finally, in section 8, we use

the patterns identified in section 6 to search for other items which exhibit similar lexico-grammatical regularities in their co-text. The similarity in these patterns indicates similarity in the perceived acceptability and/or intelligibility of the items they flag. We expect that an exploration of such other items and their co-text will increase our understanding of the code-switching phenomenon. Importantly, it will also make it possible to extend the scope of studying multilingual practices in monolingual corpora beyond elements which have traditionally been considered multilingual.

2. Code-switching in ELF: from functions to perceived acceptability

Because of the diversity of speakers' linguistic backgrounds in ELF interaction, variability and recourse to other languages than English are typically seen as central characteristics of such interaction. Perhaps for this reason, code-switching in ELF has also been studied extensively, mostly at the discourse function level in academic interactions (e.g., Klimpfinger 2007, 2009; Turunen 2012; Mauranen 2013), casual conversations among academics (e.g., Pölzl 2003, Cogo 2009), business environments (Cogo 2012), and in intercultural love relationships (Pietikäinen 2012, 2014). Although these studies present the pragmatic discourse functions of code-switching in ELF slightly differently, we can see some main similarities between the functions. Perhaps the most predictable reason to switch to another language is word-search and appealing for assistance from co-interlocutors in linguistic difficulties (Klimpfinger 2007; Cogo 2009; Turunen 2012; Mauranen 2013; Pietikäinen 2014). The second and third categories are familiar from bilingual studies: specifying addressees (or situational negotiation) by code-switching into the shared L1 or the language more commonly spoken with that interlocutor (Klimpfinger 2007; Cogo 2009; Turunen 2012; Pietikäinen 2012; Mauranen 2013) and borrowing terminology from other languages (also called referential or topical code-switching), especially when the term does not translate to English directly (Turunen 2012; Pietikäinen 2012; Mauranen 2013). Code-switching is also used in signalling solidarity toward the other group members, or a membership of a certain cultural group (Pölzl 2003, Klimpfinger 2007, Cogo 2009, Mauranen 2013), in facilitating nuanced expression and emphasis (Cogo 2009, Pietikäinen 2012), in introducing a new idea (Klimpfinger 2007) or demonstrating language knowledge (Pietikäinen 2012). Sometimes ELF speakers switch "for no apparent reason" (Cogo 2012: 308), such as when speakers are under a lot of cognitive pressure and slip out something in another language (Turunen 2012, Mauranen 2013), but Pietikäinen (2014) also shows that in long-term ELF relationships where the couples generally regard code-switching permissible, over time the speakers start to switch more often without notable awareness of it. In fact, Pietikäinen (in press) concludes that ELF couples' code-switching practices resemble the translanguaging practices of bilingual families (cf. García and Li Wei 2014), where any available languages can be resorted to (as long as they are understood by the spouse).

Surprisingly, although the pragmatics of code-switching have been examined in such an extensive way, and although the functions of code-switching have been identified by examining the talk surrounding the code-switches, there have not yet been many studies that would scrutinise the co-text of code-switches in terms of the lexico-grammatical patterning it might be exhibiting, particularly from the viewpoint of acceptability/intelligibility. Pietikäinen (2014, in press) finds that in ELF couples' talk, some short words such as response tokens and question words as well as place names and some short nouns are switched without any kind of hesitations, pauses or other flagging surrounding the switch that would indicate that the speaker is consciously switching and/or trying to prepare the listener for the switch. These switches are also not reacted to by the listener, but rather, they are accepted in the interaction as such, without the need to negotiate interpretations. It is likely that the familiarity between ELF couples, particularly their knowledge of each other's linguistic resources and their habituality of talking to one another, reduces the need to negotiate the code-switches (cf. Poplack 1987). The seemingly automatic switching further suggests that what to an outsider appears to be a code-switch, may not be interpreted as that by the speakers.

In the academic speech events of the ELFA corpus, we can expect that the degree of familiarity between the speakers is lower than, for example, in ELF couples. It can thus be assumed that what Mauranen (2012, see also 2013) calls "matching speaker perspectives", or negotiating interpretations to achieve mutual intelligibility, is a more laborious task for participants in the ELFA corpus speech events than for speakers who know each other well. In fact, previous research into ELF interaction in more "transient" (Mortensen 2014) ELF settings has shown that strategies that enhance clarity and explicitness, such as rephrasing, metadiscourse and syntactic explicitation achieve particular prominence (e.g. Mauranen 2007, 2013). Such strategies also serve in preventing misunderstandings (see e.g., Mauranen 2006; Kaur 2009; Pietikäinen 2016). In terms of code-switching, the need to negotiate interpretations relates to what linguistic resources speakers can assume others to understand. Particularly in some bilingual communities, switching from one language to another may form a "part of an *overall* discourse strategy", and the speakers may switch in ways that minimize "the salience of the switch points" (Poplack 1987: 65, original italics), whereas such code-switching behaviour may not be applicable in multilingual settings where speakers cannot necessarily expect their interlocutors to share more than one language. In the light of this, we approach code-switching from the perspective of perceived acceptability and intelligibility of the code-switch, that is, we suggest that the amount of flagging around a code-switch indicates how acceptable or understandable the speaker thinks the code-switch is. Flagged code-switching, as opposed to smooth, unflagged code-switching, draws attention to itself by hesitations, false starts, lengthy pauses, retracting, explaining, metalinguistic commentary, and the like (Poplack 1987). Flagging thus suggests a perceived need to negotiate the acceptability or

appropriateness of the switch (Poplack 1987), whereas lack of flagging around the code-switch, and lack of negotiation of interpretations following the introduction of the code-switched element, implies that the speaker expects the listener(s) to understand the item and accept the code-switch. Acceptability is thus understood in this chapter in relation to perceived understandability, what a speaker seems to consider is required to allow the discourse to proceed undisturbed, and judged by how the discourse progresses (cf. Mauranen 2013; see also Hynninen 2016).

3. Taking a phraseological approach to the common co-text of code-switching

The recent three decades of descriptive linguistic research clearly show that language exhibits patterning at different levels of abstraction. For example, introductions in research articles are organized in rhetorical moves (Swales 1990), definitions follow their own local grammar (Barnbrook and Sinclair 1995, 2001) and phrases with a clear discourse function such as *marry me* do not tolerate much variation in their form (Pawley and Syder 1983). As discussed in the previous section, we know from previous studies that at least in certain situations code-switches tend to receive flags, that is, be surrounded by textual elements which serve to draw attention to the code-switched items. A question that raised our interest is whether the flagging of code-switches exhibits some type of patterning, too. However, to ensure objectivity of our analysis, we decided to describe the patterning of all co-text surrounding a code-switch, conventionally retrieving it in concordance lines, rather than any pre-identified flags.

We expect that in addition to serving merely a descriptive purpose, such analysis of the common co-text of code-switching can also have a practical application. As an example, Barnbrook and Sinclair (1995) developed their local grammar of definitions on the basis of dictionary definitions but later were able to apply the grammar to automatic identification of definitions in non-dictionary texts. Similarly, we hope that the regularities identified in the patterning of the co-text of code-switching can then be used in searching for other phenomena which are not tagged as <FOREIGN>, yet flagged in a similar way by the speakers. Finding such other phenomena would help us both better understand the code-switching phenomenon as well as extend our scope beyond elements which have traditionally been regarded as multilingual.

Another important perspective which informs our analysis is a linear, dynamic approach to text and speech in particular (Sinclair and Mauranen 2006; Mauranen forthcoming). According to this perspective, speech is produced and perceived in a linear temporal fashion along the dimension of time in chunks, one chunk at a time. The composition of a chunk is dynamic and depends on a number of factors including cognitive processing constraints, such as memory capacity, and principles of meaning

construction. We will not apply Sinclair and Mauranen's (2006) Linear Unit Grammar model to our analysis of the co-text of code-switching in a formal way, but adopting their perspective on the unfolding of text in discourse enables us to see flagging in a different light. Presumably, different flagging elements can be seen as units which speakers can supplement interactively one by one in any number as necessary. Thus, the number of such flagging elements used or the amount of flagging may be indicative of how acceptable and/or intelligible the speaker perceives the code-switch to be.

4. Data: ELFA corpus

The data for our analysis is drawn from the ELFA corpus (ELFA 2008), a one-million-word corpus of spoken ELF in university contexts (for detailed information, see e.g. Mauranen, Hynninen and Ranta 2010; Mauranen 2012). The corpus consists of transcribed recordings of authentic academic speech events: seminars, lectures and doctoral defences recorded in four Finnish universities, as well as presentations and discussions in international conferences. Both monologic and dialogic events are included, with speakers from 51 typologically diverse first languages (L1s).

According to the ELFA transcription guide[1], instances of "switching into a foreign language" are tagged with <FOREIGN>. We can thus treat such tags as an independent criterion of code-switches suitable for purposes of operationalisation. However, in principle, we are studying a phenomenon which the transcribers of the ELFA corpus *interpreted* as "switching into a foreign language". The code-switched elements were transcribed in cases where the transcribers were familiar with the language switched into, in other cases they were simply marked "(xx)". Longer passages and passages that took place simultaneously with but outside the main interaction were not transcribed. Typically the switches are intra-sentential and rather short, mostly single words (Turunen 2012). The most common language switched into is Finnish, followed by German, Russian, Swedish, and French (Turunen 2012). Notably, switching took place both to speakers' first and other languages (Turunen 2012; Mauranen 2013). Established terms adopted from, for instance, Latin into English are not regularly marked as <FOREIGN> in the data, and these instances, along with other similar cases of fuzzy boundaries, are left outside of the analysis unless otherwise noted.

5. Methods

We have used a combination of corpus, discourse and conversation analytic approaches. The first step in our analysis was to search the ELFA corpus for all the elements tagged as <FOREIGN>. We retrieved concordance lines of 651 code-switches occurring in 82 speech events. There were no tagged code-

[1] For the ELFA corpus transcription guide, go to http://www.helsinki.fi/elfa/elfacorpus#documents.

switches in 83 speech events. Each of us then made some initial observations about the kinds of co-text that occurred around the switches. At this stage, we also compared the language of the code-switch to the speakers' L1s. After reaching a shared understanding of the classification of the regularities we observed around the switches, we used two simultaneous strands of analysis drawing on the different approaches as discussed in more detail below.

First, we analysed a subset (n=120) of all the concordance lines retrieved, consulting the extended co-text where necessary, to identify possible regularities at any level of abstraction (section 6). With an aim to achieve as comprehensive a description as possible, we divided the co-text of code-switches in every concordance line into relevant chunks following the principle of assigning preliminary unit boundaries proposed in Linear Unit Grammar (Sinclair and Mauranen 2006), as in (5.1 a–c).

(5.1)

(a) [in Germany] [i think] [they're basically referred to as] [i think] [**ausländer**] [right] [i think] [so] [erm]
(b) [one of the biggest parties] [the **keskusta keskusta**] [or] [a centre party] [which has been…]
(c) [this block of erm mandatory subjects] , [which is called] [i think] [the **gumanitarni blok**] [er] [in Russian]

Then we classified the resulting chunks and arrived at a taxonomy of elements, most of which included some lexico-grammatical regularities. To extend the description from the subset of the concordance lines analysed to all the occurrences of <FOREIGN> tags, we then searched for co-occurrences of such lexico-grammatical regularities with a <FOREIGN> tag allowing for positional variation (AB/BA) and up to five intervening words between them. For this purpose, we used ConcGram (Greaves 2009).

The regularities we identified can be described as discoursal, structural and lexico-grammatical; in other words, they manifest themselves at different levels of language organisation. Thus, overlaps between levels can be expected. For example, more lexically specific patterns inevitably overlap with discourse acts. However, this does not contradict our theoretical framework, since we do not assume an inherent hierarchy in language organisation but rather, in line with usage-based theories, we see observable regularities as a frequency-based rather than a rule-based phenomenon (see Ellis 2002). This means that such regularities can be described but do not have to fall into watertight categories. According to usage-based theoretical thinking, redundancy is a natural property of language (see e.g. Ellis 2002; Bybee 2010).

There is also a practical reason why we would like to draw specific attention to lexico-grammatical, phraseological patterning where possible rather than focus only at a more abstract discourse level: due to common limitations of corpus-linguistic tools, such patterning is easier to search for, and therefore more useful for the final stage of our study, identification of phenomena similar to code-switching in its

co-text. These considerations explain the fact that sometimes the same concordance lines will be used to illustrate different types of patterns. Note also that one code-switch can be flagged with more than one type of pattern.

Having analysed the code-switching practices in the aggregate, using corpus-linguistic tools, we then proceeded to take a closer look at code-switching in its interactional context, using discourse and conversation analytic methods (section 7). The quantity and type of flagging were considered from the point of view of their acceptability and intelligibility in the interaction. Attention was paid not only to whether and how the speaker flagged the code-switch, but also whether interlocutors reacted to the switch in any way (e.g. by requesting for clarification). Repeated occurrences of the same code-switch in the same speech event were also considered, as were contextual factors drawn from the corpus metadata. To consider these issues in more detail, special attention was given to speech events with a high number of code-switches (15 out of the 82 speech events that included code-switching had 10 or more code-switches).

The last stage (section 8) of our analysis returns to corpus-analytic methods. It provides example searches of phenomena which seem to be flagged in a similar way to code-switches. We hypothesise that similarity in the co-text of two different phenomena may indicate similarity in how these two phenomena are perceived, in our case from the point of view of acceptability and/or intelligibility.

6. Common co-text of code-switching: What regularities can we find?

The regularities we identified in the co-text of code-switching can be roughly grouped by the discourse act they seem to be performing. These discourse acts are: explication, contextualisation, request for help and hedging using pragmatic markers (taken very broadly as anything which is used to downplay and mitigate the effect of a code-switch). We have organised our description of the co-text patterning using these groups.

Explication

Steps taken to explain meaning are very common in the co-text of code-switching in our data. This is probably related to the general feature commonly observed in ELF interactions, a tendency to enhanced explicitness (Cogo and Dewey 2006; Mauranen 2007; Mauranen 2012: Ch.6; Mauranen 2013; Pietikäinen 2016). As Mauranen (2012) writes, in lingua franca discourse "speakers raise the level of explicitness to help hearers keep track, to secure uptake of their intended meanings, and to show cooperativeness" (168). She also compares this feature of lingua franca discourse to a similar tendency

of translators to "add clarifying matter to the text as they translate" termed 'explicitation' by Blum-Kulka (1986) (Mauranen 2012: 167–168).

In the co-text of code-switching, the most common and structurally regular type of explication we have observed is glossing, that is providing a succinct interpretation or paraphrase of a word or phrase perceived as potentially problematic. Very often code-switch itself is used as such a paraphrase (6.1) and vice versa, code-switches can also be glossed (6.2).

(6.1)

it's a big player er minis- ministry of natural resources **ministerstvo prirodnykh resursov** er issuing licences[2]

also er lineal drawing and er **dibujo lineal**

we were regarded to be among the , super rich nations the **superreichen** together with switzerland japan

his idea of the whole or wholeness **tselostnost' ganzheit**[3]

then she's called mrs spokesman **rouva puhemies**

(6.2)

that are building these **kunskapsanläggning** these knowledge complexes

one of the best museums in in the country right now called **bild museet** er picture museum

they put a bottle of **absolut** er swedish vodka

you mentioned here this er **laatua verkkoon** high-quality web services

The most regular feature of glossing seems to be structural: noun phrase (NP) serving as a gloss is usually an appositive, that is, it is added to the anchor NP in apposition and is not integrated into the syntactic structure, such as in *"Bizet's most popular opera, 'Carmen', was first produced in 1875"* (Huddleston et al 2002: 1353). This structural syntactic feature can be observed in all concordance lines in (6.1) and (6.2).

Another way of providing a quick gloss is to link the code-switch and what often seems to be its translation into English by a coordinating conjunction *or* (6.3). Again, the code-switch and its translation seem to be interchangeable in the order of their appearance, which means that both the code-switch itself and the translation can serve as a gloss.

(6.3)

or folk art which is called postmilitary album **or** in russian **(dembinski) album**

[2] Backchannels and overlapping speech are usually removed from the concordance lines for clarity. **Bolding** is used for code-switched elements, underlining for the elements of co-text in focus. Only a few concordance lines are presented in each case.

[3] It is interesting that here equivalents in two languages are given.

the name of helsinki er hels- er **teknillinen korkeakoulu or or** helsinki university of technology

In (6.4), we see both syntactic strategies, apposition and coordination with *or* used at the same time.

(6.4)

and it it's belongs to the **muoto** 2005 programme design **or** shape 2005 actually
one of the biggest parties the **keskusta keskusta or** a centre party which has been the agrarian party

In addition to glossing, the meaning of a code-switch is often explained or hinted at by other more variable means, usually done after the potentially problematic item. These types of clarifications can be seen in (6.5).

(6.5)

for black people which is in finnish **neekeri** it's erm same as negro and it's
they say in germany **molch** it's a er kind of creature that
institute for minority rights **vähemmistöoikeuksien instituutti** it's a kind of there are le- lawyers mainly
<S4> what's **sauva** </S4> <S1> ah yeah this is the finnish version okay so leukocytes you understand

Very often the discourse act of clarification is realised through the use of the lemma MEAN (6.6).

(6.6)

which is essential for the russian word **tila** means only space , still in some
naistutkimus then . so **nais** meaning a woman and
italian **pericoloso** he said **pericoloso** it means oh danger danger
choose er **erwachsene preisstufe drei** it means er er adults grade price category three

Again, we also see symmetric cases where the meaning of a term is explained by providing an equivalent in a different language, i.e. by code-switching (6.7).

(6.7)

<S13> [everyone] knows what collateral means **lainan vakuus** < S13> <S9> it's the
LEF magazine quite popular in 1920s LEF means er **levyi front** the er left front , and
so-called research libraries the meaning of german words as er **forschungsbibliothek** er modern time

One other strategy that can be interpreted as explication is to precede a code-switch with a label which defines the category of entities it falls into, e.g. ‘a film’, ‘a word’, ‘a concept’ (6.8).

(6.8)

i've er started using the the concept er **kunskapsanläggning**

greek words **demos** and **kratos**
so much surprised after the presentation of the film **juurakon hulda**

Providing such labels in advance works both to warn the listener of a potentially problematic or unfamiliar item coming and to pre-define it.

Request for help

As has been already discussed in ELF literature, in the context of code-switching it is common to explicitly ask for help when searching for a suitable word (Turunen 2012; Mauranen 2013). Sometimes such requests for help are responded to and a possibly suitable word or term is provided by the interlocutors (6.9).

(6.9)

but not so much those er what is **valjaat** in english [er] </S8><S10> [lifts] </S10><S8> those lifting system er they
it says here that extra-hepatic er jaundice is contra- **vasta- vai mikä se oli** </S2><S6> [contraindication] </S6>
some sort of erm how do you say er **harjoitustyö** in sort of erm </S1> <S18> research exercises

However, very often such requests for help do not seem to be designed to receive a response, but rather to win time to search for a suitable word and as a hedge (6.10).

(6.10)

who were working as er , what are those **tuntiopettajia** teachers </S3>
i would say **mikä on manner** er the continental europe would be the centre
here were like er er truck drivers what is is it truck drivers i don't know **trukkikuski**

In both uses, a co-occurrence of *how/what* with *say/call* seems to be common for requests for help as a lexico-grammatical feature. We will use this feature later in the final stage of our analysis.

Contextualisation

We have noticed that to anchor the code-switch and in a way to justify it, ELF speakers often use the strategy of contextualising the code-switch by saying that this is a phrase we use, or they use, or this is a phrase used in X country/language. Interestingly, some of these acts of contextualisation seem to exhibit lexico-grammatical regularities. For example, constructions involving the verbs *call* and *say* are very frequent, with the most fixed phrases being *as we/they say*, *what we/they call*, *so-called, called* (6.11).

(6.11)

<S13> **khorp koon ka** as you say
main page of this people first network website **pipolfastaem** as they say in solomon islands pidgin
just give when when i introduce ourself we say **afro-suomalaiset** i don't know whether
the city really er er they say in germany **molch** it's a er kind of creature that
abolished and pol- and er er and and 49 regions **wojvodships** we we call in it in poland
er which stems from i think what we swedish call **läser traditionen** the r- reading
official struggle er against so-called er **kosmopolit** cosmo- cosmopolitans with

It is also common to mention the country or language of the code-switch (6.12).

(6.12)

this was a project called affectionately called **mutu** in finnish it has two names it's
like er general theory of knowledge but **épistémologie** in french is rather the
partly can use its own incomes < S7> <S13> like **metsähallitus** in finland <S7> yeah <

Sometimes more specific geographic locations are provided as in (6.13) or linguistic categorisations as in (6.14).

(6.13)

really amazed , erm first er by these kind of **nettinysse** in [tampere]
not a company it is a budgetary enterprise erm **(xx)** er in bialystok and city sport
you were saying may have been been said by some **(twa)** in the in the congo forest you

(6.14)

reason the reasons of this dominance of the word **ylioppilas** in the language of translations
particles **jah jaa** and **jajaa** in telemark- marketing calls <
when in in er russian scientific text you say **A i B neotryvno svyazany** A and B are

Hedging using pragmatic markers

Use of metalinguistic commentary for flagging code-switches is mentioned already in Poplack (1987). In a similar way, we find a number of discourse phenomena occurring around code-switches which mostly fall under the category of pragmatic markers. These markers seem to be used as hedges or to be more precise as some kind of 'speed bumps'. Despite subtle differences in the specifics of their pragmatic function, all of them seem to serve as inserts to slow down the discourse and warn the listener that there is something which requires attention. We find conversational hedges *sort of*, *kind of* and *like* downplaying the code-switch (6.15), general extenders *or something like that*, *or whatever*, *and so on*,

and all that adding imprecision and vagueness (6.16), discourse markers *I mean, you know* and, presumably, its variant *you understand,* appealing for the listener to activate previous or shared knowledge (6.17) as well as frequent instances of hesitation (6.15–6.17).

(6.15)

actor or an actress we have this same sort of er **näyttelijä** and **näyttelijätär**
really amazed , erm first er by these kind of **nettinysse** in [tampere]
this kind of german romanticist idea of er the **kulturnation** [and er and]
we get er equivalents like er **tila väljä** and so on

(6.16)

points from this by doing a **luentopäiväkirja** or something something like that
camera angles [and] point of of views and cuts [and] **mise en scène** and all that
you have words like **båt** and **boot** or whatever and things like that
we get er equivalents like er **tila väljä** and so on

(6.17)

their funds from er from the state anyway i mean **mellemfolkeligt samvirke** which is on
so er **afro-suomalaiset** you know that's the picture we just give when when i introduce ourself
you can speak with me in japanese go ahead **arigato** you understand <S9> @yeah@ <

We also see longer utterances which seem to be serving the same function of slowing down the progression of discourse and mitigating the code-switches (6.18).

(6.18)

a complete **décalage** well i don't know if you can use that word in english but
we say **afro-suomalaiset** i don't know whether my finnish colleagues will agree @with me@
er truck drivers what is is it truck drivers i don't know **trukkikuski** <S2> uh-huh <
you write about the er don't know how to pronounce it but the hermeneutical **hermeneuttinen** research method
using the the concept er **kunskapsanläggning** i haven't really thought about any good word for in in English

In addition to pragmatic markers, we also find frequent use of the definite article *the* and demonstratives *this, these* and *those* in association with a code-switch (6.19). In fact in our data, the definite article alone immediately precedes a code-switch in 51 cases (for examples, see the five last concordance lines in example 6.19).

(6.19)

or thought of as positive , for example there's this **hyvä jätkä** would you like to

could <S18> mhm < S18> suspect that this project this **laatua verkkoon** would in its
as we <S2> mhm < S2> do today and i was teaching this **variabiliteettilaskenta** we
certainly among those that are building these **kunskapsanläggning** these
central europe who were working as er , what are those **tuntiopettajia** teachers <
part here looking very similar to for example the **kungliga tekniska högskolan**
would be extremely unhealthy and it's all for the **lapsen etu** the er < S5> <S1>
regarded to be among the , super rich nations the **superreichen** together with
erm mandatory subjects , which is called i think the **gumanitarni blok** <S1> mhm <
and want them to embrace this you know the **eduskunta** the or whatever

We suggest that the demonstratives, and in some cases also the definite article, have a focusing function and therefore can also be considered instances of flagging. This conclusion is in line with Poplack, Wheeler and Westwood (1989) who consider such usage as a sign of flagging. They found that in conversations among Finnish-English bilinguals, where Finnish was the matrix language, what they call Finnish 'determiners', such as *se* 'it/this', *tämä* 'this', *tuo* 'that', *siellä*, *tuolla* 'there' and *täällä* 'here', were "particularly frequent in the environment of English nouns" serving a "switch-signalling" function (Poplack et al. 1989: 401-402).

Probably the most interesting observation we were able to make is that different patterns we identified seem to form some kind of pool of resources which are used to facilitate a code-switch and are often combined in various ways without any visible constraints in the co-text of one code-switch. In (6.20), we provide some examples of such combinations.

(6.20)

(a) from the main page of this people first network website **pipolfastaem** as they say in solomon islands pidgin

(b) we have a sort of mhm mhm mhm well what we call er **sektorsforskningens avveckling** the sort of taking away of sectorial research and the academisation of that research

(c) want them to embrace this you know the **eduskunta** the or whatever the government

In (a), there is a demonstrative *this*, glossing by using apposition and two very frequent phrases, *as they say* and *in Solomon islands pidgin,* which contextualise the code-switch. In (b), there is a hedge *sort of,* hesitation, pragmatic marker *well,* a construction with the verb *call* (*what we call*) and explication by apposition. In (c), we see a demonstrative, a pragmatic marker *you know*, inserts of the definite article, a general extender and a gloss. In the next section, we will approach such combinations of patterns from the perspective of perceived acceptability and intelligibility.

7. What may explain the quantity and type of flagging around a code-switch?

As was established in the analysis of the co-text of code-switching, the quantity and kind of flagging around a code-switch varies. Below, we take a closer look at examples of code-switching to illustrate different factors that may explain this, paying particular attention to how flagging can be seen to relate to acceptability and understandability.

Instances of least or no flagging

Some of the code-switches in the ELFA corpus had little or no flagging around the switch. Below, we take a look at such least-flagged examples, where, as far as the transcriptions reveal[4], the switch is embedded in the matrix-language syntax so that the switch does not seem to interrupt the speech flow, and where none of the interlocutors intervene in the switch.

The first example (7.1) illustrates that the"locality" of the code-switched element seems to be a factor that reduces the need to negotiate the acceptability of the code-switch.

(7.1)[5]

(a) <S2> mr ***kustos*** er mr candidate ladies and gentleman the dissertation of
(b) <S1> as a ***kustos*** appointed from the by the faculty of social sciences
(c) <S2> distinct ***kustos*** distinct opponents ladies and gentlemen er it has been
(d) who would like to ask further questions please ask the ***kustos*** for the floor </S1>
(e) <S4> as the ***kustos*** of this dissertation i declare the discussion concluded

In all, the ELFA corpus included 13 instances (not counting reading aloud passages) of the recognisably Finnish noun *kustos*[6], which refers to the faculty-appointed chair of the public examination of doctoral dissertations in Finland. The corcordance lines in (7.1) illustrate typical occurrences of the noun, all found in speech events related to the public examination of dissertations. All 13 instances found were embedded in the matrix language syntax with little or no flagging of the element as a code-switch. Rather, as shown in (7.1), it seems that *kustos* is treated similarly to other titles (e.g. *mr kustos er mr candidate*, *distinct kustos distinct opponents*), and that the code-switch is treated by the speakers as any

[4] It should be noted that, for instance, intonation and word stress are not marked in the ELFA corpus transcriptions.
[5] Please note that in this section, ***bolding and italics*** are used for code-switched elements, **bolding** for the elements of co-text in focus. The transcriptions follow the ELFA transcription conventions, but rising intonation ("?") and word stress (underlining) have been marked in examples 7.9 and 7.10. The symbol "@@"refers to laughter.
[6] The latinate spelling *custos* is sometimes used in materials written in English.

English element. In our quantitative analysis, the definite article "the" is listed along with demonstrative pronouns as a flag that seems to separate out the code-switch. While "the", which is also used in connection with *kustos*, may function as a flag the same way as demonstratives do[7], it seems that this is not always the case. In (7.1), when the definite article occurs before the noun *kustos*, no other flags are used, and the embedded switches do not seem to slow down the flow of interaction by attracting attention to themselves; in fact, the use of *the* follows Standard English grammar rules. The code-switches are from different L1 speakers and from different speech events in the corpus, and in addition, the switches occur in formulaic sequences that the speakers may have encountered in similar events before or read from instructions available for those participating in the public examination of dissertations. All this suggests that the noun has been widely accepted, and in fact, it seems that *kustos* could be seen to represent nonce borrowing (following Poplack et al. 1989: 403) rather than code-switching.

Another example (7.2) where the code-switches are surrounded with little or no flagging and where the switches do not seem to slow down the flow of interaction is provided below. The example is from a social sciences seminar, where S10, an L1 speaker of Swahili, code-switches into Finnish. The speaker is talking about changes in Finnish social security provisions. It may thus be that not only the mainly Finnish-speaking environment but also the topic influenced the choice of the speaker to code-switch into Finnish.

(7.2)

S10: ((…)) but now foreigners are not included and it's only maybe finns who get this free education but then they also have to pay through ***kela*** but before it wasn't like that it was totally free it was only in 1990- if I'm not mistaken 96 97 which everything started , of course also in the case of health care and we now have to pay everything if you go to buy this medicine at **the *apteekki*** no 10 per cent (per cent)

S15: yeah

S10: for students anymore so this is the change and er ,

[7] The following example illustrates how the definite article "the" can function as a flag in a similar way to demonstrative pronouns:

S1: …but if if the spokesman is a woman then @@ then she's called **mrs spokesman *rouva puhemies*** that's actually a bit funny…
<TURNS OMITTED>
S12: [i think it was] , i think it was the same with **the finnish title that you introduced the *rouva puhemies***

In the above example, "the" is used to refer back to an expression that has been used before, which highlights the specificity of the element. Notably, the code-switched element is also flagged by other means that single out the code-switch.

The first code-switch (*kela*)[8] is a reference to The Social Insurance Institution of Finland, and the second one (*apteekki*) is the Finnish word for pharmacy. The first switch is unflagged. The second one is preceded by the definite article *the*, and the demonstrative pronoun *this* suggests uncertainty in the choice of the noun *medicine*. However, the switch is embedded in the matrix language syntax, and the interlocutors do not react to or intervene in any way in the use of the switch, but rather provide minimal feedback (L1 Finnish speaker S15's *yeah*) that confirms interlocutor understanding and thus suggests that it was acceptable to use the code-switched elements in the speech event.

Embedded code-switches as illustrated in (7.1) and (7.2) suggest that elements that have been tagged as a code-switch in the ELFA corpus, because they have been recognised as "not English" by the corpus compilers, are not necessarily flagged as "non-English" elements in the speech events. Rather, they may be treated by the speakers as expected in the discourse and therefore not requiring any additional preparatory work or clarification. No or little flagging of a code-switch thus implies that the code-switched element is rendered acceptable by the speaker. Also, since the interlocutors do not problematise the switch either, it seems that there is no need to negotiate interpretations; the code-switch is accepted and understood as it is. In the case of unflagged code-switches it is also possible that speakers are not always fully aware of the alternation between languages, which would suggest what Pietikäinen (2012, 2014) calls "automatic code-switching".

Flagging for the sake of matching speaker perspectives

Flagging the code-switch singles out the code-switch as something special (see Poplack 1987): it indicates that code-switching needs to be justified somehow and that the acceptability of the switch is negotiable. Flagging is thus an important means to match speaker perspectives (see Mauranen 2012). This is illustrated in (7.3), where the speaker not only provides a gloss for the code-switch (*people first*), but also contexualises the switch (*as they say in solomon islands pidgin*).

(7.3)

S7: ((…)) here we have a statement er which I'm not going to read aloud it's taken from the main page of this **people first network website *pipolfastaem* as they say in solomon islands pidgin** and er er this er website and the whole people first network project began in the year 2001

The example shows extensive flagging around the code-switch, which indicates that the speaker could not expect his or her interlocutors to understand the code-switch without explanation. Rather, by

[8] Interestingly, this element was actually not tagged as <FOREIGN> in the ELFA corpus. Since our assumption is that a reader unfamiliar with Finnish would not be able to decipher the meaning of the element, we have singled it out in the analysis.

flagging, the speaker tries to make sure that the intended meaning is understood. Singling out the code-switch in this way makes it more acceptable in the discourse (cf. Poplack 1987).

A code-switch can occur only once or it may be used repeatedly in the same interaction. In general, it seems that a code-switch that only occurs once is more likely to be flagged than a repeated code-switch. One could assume that once the code-switch has been introduced in the interaction, it no longer needs to be flagged because it has been established as an accepted form, shared by the participants, and no longer needs to be viewed as a potentially problematic element. Example (7.4) from a lecture in Russian studies illustrates how the same speaker flags the code-switch when he introduces it in the discourse, but no longer flags the switch later in the interaction.

(7.4)

S1: ((...)) and er we had at that time during second world war very interesting publications , and er papers during different conference and this situation was er possible er till 1949 after that we had i won't go to (more) this term and er we had er **struggle official struggle er against so-called er *kosmopolit* cosmo- cosmopolitans** with <NAME> er the most best scientist and er he taught with wonderful education they had no opportunity to continue their researches and so on

((…)) and er other things it was very good er opportunity to for some libraries to be in the shadow of such er , erm so-called , er . ideological struggle and persons who had no opportunity after this year of **the struggle against *kosmopolits*** had no possibility to (xx) they had opportunity to be to become librarians , and to work during some many years in research libraries as bibliographers

What is interesting is that although the second instance of the code-switch is not flagged by hesitation or glossing, it is framed in a similar way with reference to the *struggle against*. It is likely that this mirroring of the structure is used purposely to remind the interlocutors of the meaning of the code-switched element.

Another example where the flagging around the code-switch reduces is given below.

(7.5)

S5: ((...)) i- in in in er **in swedish** er i I've er started using **the the concept er *kunskapsanläggning* i haven't really thought about any good word for in in english** er to er to er to to sort of underline that it's not just an institution it's not just a building er but it's rather **complex set of theories** er

((...)) er there is a growing awareness certainly among those that are building **these *kunskapsanläggning* these knowledge complexes** er that they are important

In the example, the code-switch is first flagged by contextualising, explicating and explaining the code-switch, whereas the pronoun *these* introducing the second instance of the code-switch serves a focusing function (see section 6) and suggests that the expression is treated as an established expression, even if the code-switch still receives a gloss.

In some cases, repeated code-switching of the same element was flagged similarly every time the switch occurred in the speech event. This is exemplified in example 7.6, where the code-switch *powiat* is always preceded by a gloss.

(7.6)

(a) we had communes *(xx)* we had **counties *powiat powiats*** and we had regions *wojvodships*
(b) by local government before the by **co- county *po- powiats*** er before the second world war
(c) of course county remain **county** remain ***powiat*** er remained as er administrative
(d) 60 regions , and 350 **counties *powiati powiats(xx)*** and we had
(e) i- in each **county** in in each in each ***powiat*** has w- er has to be one

In the example, the code-switching is done into Polish, by an L1 speaker of Polish, in a speech event where all other speakers (but not all participants) were L1 Polish speakers. Against this background, it is perhaps surprising that the speaker continues to gloss the code-switch throughout the interaction. In general, it seems that the presence of several speakers of the same L1[9] seems to have increased the amount of code-switching into that L1, but it does not seem to have reduced the amount of flagging, as also illustrated in (7.6). Rather, it seems that participants with other L1s were still a concern. However, it seems that the speaker's knowledge (or expectations) of participants' shared linguistic resources influences the kind of code-switching the speaker resorts to. In general, the language switched into was not related to a specific function of the code-switching (Turunen 2012; Mauranen 2013), but code-switching related to word searches was almost always done into the speaker's L1, and nearly all such examples in the ELFA corpus were from code-switching into Finnish, the main language of the environment. This suggests that the speakers assumed other Finnish speakers to be present and available for help. While contextual factors may thus not influence the amount of flagging in any systematic way, they may influence the acceptability of which language to switch into.

[9] Only the L1s of the speakers are documented in the ELFA corpus, which means that we do not know (a) what the L1s of the participants who did not speak in the events were and (b) what possible other languages the speakers spoke in addition to their L1(s) and English.

In all, we can say that a switch that a speaker can assume others to understand seems to receive less focus. As Firth (1996: 247, original italics) demonstrates in relation to the use of non-standard resources in ELF interaction, "participants can learn and use known (and also nonstandard) resources *as they become known-in-common* during the talk itself". There is no reason to expect that code-switched elements could not similarly "become known-in-common". In fact, a case where the flagging reduces has been reported in similar ELF data in Hynninen (2016: 161–162); and while in (7.4)–(7.6), the code-switch is not adopted by other speakers in the same speech event, Hynninen's (2016: 161–162) analysis demonstrates that code-switched elements may well be adopted as accepted forms in ELF interaction. Her example shows that after a particular code-switch was introduced in the interaction by explaining its meaning in English, the same code-switch was used by other speakers without flagging.

Sometimes, though, participants in an interaction can have different ideas about the acceptability of a code-switch. This is illustrated in the following analysis of a speech event with the highest number of code-switches in the corpus (103 code-switches in little over 1.5 hours), despite at least six different L1s. Recurrent code-switching to a specific language or languages in a particular speech event in the ELFA corpus can most often be explained by the overall topic of the event (e.g. translation studies[10], slavonic philology). In the case of our example event, the high number of switches relates to a concept in competence studies discussed in the event. When analysing the recording of the speech event using a conversation analytic approach, it quickly became apparent that the switches are mostly produced in German by an L1 German speaker (S1), the presenter at the seminar. Her presentation deals with, as she puts it, the "very traditional German term" *Bildung* which, as she explains, is translated "into English with the word education but it's quite something different". Of all the switches, 91 comprise of or include the noun "bildung" or the conjugated verb forms "gebildet", "sich bilden" etc., and 68 of these are uttered by S1. S1 very clearly articulates the first instances of "gebildet" and "bildung" but does not flag them in any other way, see (7.7). Apparently, she also uses a screen presentation from which the participants can read these terms.

(7.7)

S1: this our our view on if a person er is good he has to be ***gebildet*** he has to have ***bildung***

[10] In the case of translation studies, although the speakers may assume that the other participants share the linguistic resources switched into when discussing the translation of a text from one language to another, code-switches typically seem to be flagged. This may be explained by the need to separate different codes from each other in order to compare the codes, and thus the flagging seems to have less to do with the acceptability of the code-switch from an understandability perspective than in other kinds of speech events.

Notably, although S1 uses "bildung" and its variations without flagging in her speech throughout, other non-German speakers flag the term upon the first mention (7.8), or ask for a translation (7.9), which indicates that although S1 repeats the term constantly, its meaning is not entirely clear to everyone.

(7.8)

S4: ***bildung* as you define it** , can be seen and is always linked with

(7.9)

S6: c- can ***bildung*** , what is an- any other language? , or just , impossible to translate (xx)

Even other L1 German speakers flag "bildung" and "gebildet", although this is not always visible in the original transcripts as some laughter, hesitations, pauses and intonation which can be noted in the recordings have been omitted from the ELFA transcriptions.

It is evident that although S1 uses the German term without flagging and tries to explain the use of the term in spite of the lack of an English equivalent, the fact that she uses German in her explanations (as in (7.10) and (7.11)) and focuses on the grammatical use rather than the meaning of the word, makes the term even more difficult for the non-German speaking listeners to understand.

(7.10)

S1: wilhelm humboldt erm and he said erm when there was also in germany the beginning of educational sciences in er university and he said a person who who is ***gebildet* you don't have *bildung* in german you are *gebildet*** that's

SU: @@

S1: yes that's @@ er you ***man hat nicht bildung sondern man ist gebildet***

(7.11)

S1: a person who has ***bild***- w- who ***ist gebildet*** er **is somebody who is all around er competent not competent but *bil- gebildet***

The result of S1's switching practice is that she loses some of her epistemic status in front of the audience; the audience starts to laugh whenever she switches into German, and other speakers start overlapping her and producing both parallel discussions and straightforward criticism, as in (7.12) and (7.13).

(7.12)

S5: maybe it is ***bildung*** the ratio or whatever is something about <COUGH> **the german problem the germans want to have such such words**

(7.13)

NS7: **i don't mean to undermine it but** from an american stand point and i'll speak for all americans now , **who cares**

In sum, although code-switching has been found to facilitate understanding in some earlier ELF studies (e.g., Mauranen 2013; Pietikäinen 2016), upon closer look we are able to observe that this is not always the case. Even though some speakers may view embedding terminology from other languages within their English completely normal and unproblematic, their listeners may not be as supportive of this practice, as it might be interpreted as overcomplicating things. The analysis thus illustrates how important "matching speaker perspectives" (Mauranen 2012) is for achieving mutual understanding. Upon introducing the code-switched element, the German speaker did not flag the switch in any particular way other than clear articulation, which suggests that she treated the code-switch as acceptable. Her interlocutors, on the other hand, often flagged the switches, which suggests that not all of them shared the speaker's perspective about the acceptability of the code-switch. There was thus a clear discrepancy between the interlocutors' perspectives on the understandability/acceptability of the term. The analysis also shows that a high number of code-switches in a speech event does not necessarily indicate that code-switching is perceived any more acceptable. Particularly if interlocutors continue to flag an element throughout the speech event, this suggests that the element does not become a fully accepted form in the event.

Acceptability of cognitive processing slips

Some code-switches with least flagging (and typically in the form of hesitation) are best explained as cognitive processing slips (cf. Mauranen 2013). Often there was hesitation around L1 pragmatic particles that tended to occur in places where the speaker was processing his or her thoughts. A good example is the following where the processing is verbalised particularly clearly:

(7.14)

S6: … there's always some female words for him maybe **i i can't explain it in *also* i can't translate something like that it's er maybe** *(xx)*

In the example, the German adverb *also* (English "that is" or "I mean") occurs in a situation where the speaker is searching for ways to express herself. As Mauranen (2013: 238) points out, "[c]ognitive

processing slips seem also implicated in using L1 pragmatic particles where cognitive load tends to be particularly high, like utterance beginnings or other junctures, and they are often accompanied by hesitation markers". Our point is that these code-switches are not really code-switches in the sense that the speakers would consciously use them, but rather fall in the category of automatic code-switching (Pietikäinen 2014). Often also interlocutors do not react to these switches. However, interestingly from an acceptability perspective, speakers sometimes self-corrected their slips, which suggests that they considered such (seemingly unintentional) code-switching repairable. This is illustrated in (7.15) and (7.16). The first slip is into Danish *men* (English "but") by a Danish L1-speaker, and the second to Finnish *eli* (English "in other words") by a Finnish L1-speaker.

(7.15)

S2: the reason might be we don't know it exactly ***men*** the reason **but** the reason might be that in denmark

(7.16)

S7: it was a good example of one organisational model of how to improve PFM ***eli*** er **in other words** participatory forest management

Summary of flagging

In all, it seems that the presence or absence of flagging reflects the speaker's understanding of how familiar the code-switched element is among the participants. Code-switches that receive little or no flagging are not problematised by the speaker, but rather are used as a normal part of interaction. Unflagged code-switches may also occur unexpectedly in a way that suggests automatic code-switching (cf. Pietikäinen 2014). Sometimes, though, the interlocutors' perspectives about the understandability and/or acceptability of the item do not match those of the speaker (as in the *bildung* example), in which case a negotiation of interpretations is needed, and this may lead to flagging of the code-switch in subsequent turns. In general, flagging the code-switch singles out the code-switched element as needing justification, and flagging can thus be seen as a means to match speaker perspectives, and negotiate acceptability.

8. Similar flagging around elements not tagged as code-switches

In the previous sections, we have suggested that the extent and type of flagging around an instance of language use is related to perceived acceptability of this language use. We have analysed the flagging patterns around code-switches which are traditionally considered to represent multilingual elements. In this section we will extend the analysis to other phenomena that seem to be flagged in a similar way to

code-switches. We suggest that similarity in the flagging patterns is likely to indicate similarity in perceived acceptability. As such, other language uses which we might be able to find using this technique can both inform our understanding of code-switching practices as well as uncover other cases where ELF speakers prefer to flag items they perceive as potentially problematic from the point of view of acceptability and/or intelligibility. Given the inherently multilingual nature of ELF interaction, such other cases might also be related to multilingual practices which were not tagged as such in the corpus and therefore are difficult to identify (cf. Rütten in this volume).

In the corpus-linguistic analysis of the co-text of code-switching, we found several lexico-grammatical patterns, such as constructions involving the verbs *call* and *say*, phrases and frames such as *we/they say, what we/they call, in* 'country/language', and the phrase *what is it* or a co-occurrence of *how/what* with *say/call* in requests for help. At this stage of the analysis, we searched for these lexico-grammatical patterns occurring without <FOREIGN> tags in their immediate surroundings using AntConc (Anthony 2014) and ConcGram (Greaves 2009). Below we provide a few examples of what we were able to find.

Searching for co-occurrences of AS/SAY allowing for a few intervening words returns references to other individual people (*as you say*), people of other nationalities (*as Americans say*) and other professions (*as we say in journalism, as economists say).* In (8.1), the speaker seems to assume that the phrase '*inputs and skills'* might not necessarily be familiar to his/her interlocutors since it is a phrase commonly used by people of a different professional community. For this reason, presumably, s/he precedes it with a contextualising token, *as economists say*, and explains the meaning by adding a clarification remark.

(8.1)

> mostly the economic growth we have enjoyed , er has come from the rise of productivity in the manufacturing sector . and that is a result of as economists say the result of of **inputs plus skills** . that is , that we use the resources that we have more effectively than before

In (8.2), a Finnish speaker invites the audience to be '*active and aggressive'* but apparently assumes that this is an unexpected turn of phrase in the current context and prefers to contextualise it by marking it as a borrowing from American English (AmE)[11]. Thus, the phrase used in AmE does not seem to be automatically perceived as acceptable and intelligible in this ELF interaction.

(8.2)

[11] In fact, searching for the phrase *active and aggressive* in the Corpus of Contemporary American English (COCA) returns 18 hits, and in the British National Corpus (BNC) zero hits.

you are welcome to interrupt er at any point and er ask for clarification or put some remarks so please be **active and aggressive** as americans say

Queries for the co-occurrences found in requests for help return cases of open word search and uncertainty of a language use or form (8.3) as well as term negotiation (8.4). Note that in all of the cases the problem item is surrounded by other flags commonly found in the co-text of code-switches. These other flags are also underlined.

(8.3)

(a) these tables at the back of the book of the you know **appendices** how do you pronounce it

(b) about the intranjudi- intransu- w- what is it **unchangeability** forget it <S1> [mhm] </S1> [er] the unchangeability of these er categories

(c) the communities are also trying to be er kind of er er er er what is it er yeah **watchdog** or something like that

(8.4)

(a) and women have you know they're wearing the what is it called **hajab** </S1><NS3> hijab </NS3><S6> yeah [hijab]

(b) it's the one that was put on the the old erm erm how do you call this this **ice-curtain iron-curtain** on the old <NS4> [oh okay yeah] </NS4> [iron-curtain] that's where they put the green-belt

(c) <S2> [yeah yeah] because there are lot of erm how you call like economy on their illegal <S1> yes </S1> works <S1> yes yes </S1> so it is </S2><S1> it's underground <S2> [yeah] </S2> [economy] (**shadow) economy** </S1>

(d) it's like a a s- signature of a real artist like er er to you know er er how how do you call it not blink but er to , what is it called in english i'm not a native speaker but er i guess then it is not (xx) for native speakers er **to squint** no not to squint er <SU> squint </SU> oh there is <SU> squint why not mhm </SU>

On the basis of these examples, we suggest that the methodology we have adopted can facilitate identification of other items ELF speakers perceive as potentially problematic, including cases where they draw on their multilingual resources. In the light of these other instances of language use identified, we suggest that code-switches constitute just one type of items which can be perceived by speakers as potentially problematic and therefore requiring flagging. In this code-switches are similar to for example (a) using possibly unconventional language forms, or *approximations* (Mauranen 2012), (b) using borrowings from other domains of language or other discourse communities, and (c) using terminology which is likely to be unfamiliar to the interlocutors. The common patterning we have identified in the co-text of code-switching can thus be used to search for all these other phenomena.

9. Conclusions

In today's globalising world, where speakers more often than not have a command of more than one language, most corpora to be compiled are likely to contain multilingual practices. At the very least, some of the communicative events recorded in a corpus are likely to include bi-/multilingual participants. These participants will engage in a continuous reflection of whether or not to make use of their multilingual repertoire in a given situation with given interlocutors, and if yes in which way. This on-going background reflection will involve estimation of shared ground as well as linguistic resources available to the interlocutors. This in itself can already be considered a multilingual practice even if specific multilingual elements do not eventually surface, while they are certainly very much likely to, given the factors at play.

In this chapter, we analysed the co-text of elements tagged as <FOREIGN> in the ELFA corpus and identified its common lexico-grammatical regularities. We showed that speakers in ELF interaction use these regularities to flag potentially problematic items, including but not limited to code-switches. The extent and type of flagging used seems to be determined by the perceived acceptability and/or intelligibility of the item. By using the lexico-grammatical regularities identified as search items, we have extended our analysis from code-switches to other items evidently perceived as potentially problematic by speakers in ELF interaction.

The analysis we have presented entails two important conclusions. First, code-switches do not necessarily have to be treated as potentially problematic items. If the speaker estimates the code-switched item as acceptable and intelligible in a given context with the given interlocutors taking into account the linguistic resources available to them, s/he may not consider it necessary to draw attention to it with numerous flags of different types, if at all. But, as we have seen from the 'Bildung' case in Section 7, such estimation can also be incorrect and have unfortunate consequences. Second, code-switches are not the only items flagged as potentially problematic in (ELF) interactions. In Section 8, we provided several examples of items the speakers thought necessary to flag in a similar way to code-switches. Thus, the lexico-grammatical patterning of code-switches we described in this paper can be used to search for any items which are perceived by speakers as potentially problematic from the point of view of acceptability and/or intelligibility. In a multilingual setting, these items are often related to multilingual practices since it is the matching or non-matching of the linguistic repertoires, or, speaker perspectives, of the participants of communication which can invite negotiation. A tendency to enhanced discourse explicitness (Mauranen 2012) which is found to be characteristic of ELF interaction points in the same direction.

The global trend of using English as the language of communication in intercultural settings and the associated multilingualism of the speakers will most likely not change the essence of language communication – our processing of language is limited by the same cognitive constraints we have always had. As it is widely accepted, a listener's online language processing relies on anticipation of what is going to come next in the unfolding text, whether spoken or written (Christiansen and Chater 2016; Sinclair and Mauranen 2006). If the expectation formed is not confirmed, that is, the discourse does not proceed in the way the listener anticipated, there is a risk of communicative turbulence, even misunderstandings at worst. Fortunately, humans as an animal species, have an exceptional ability to read the expectations of their interlocutors (see, e.g., Tomasello 2008, 2014 on joint attention) and thus have the cognitive means to monitor them and interactively adjust their language use if needed. We suggest that it is such adjustment of language use in view of potentially problematic items coming that we have observed in the flagging patterns of code-switches in this chapter.

Acknowledgements

For Niina Hynninen's part, this study has been written in the *Language Regulation in Academia* research project (see http://www.helsinki.fi/project/lara), funded by the Kone Foundation [grant number 088787].

References

Anthony, Laurence. 2014. AntConc (Version 3.4.3w) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/.

Auer, Peter. 1995. The pragmatics of code-switching: a sequential approach. In L. Milroy and P. Muysken (eds.) *One Speaker Two Languages.* Cambridge: Cambridge University Press. 115–135.

Barnbrook, Geoff and John M. Sinclair. 1995. Parsing COBUILD entries. In J. M. Sinclair, M. Hoelter and C. Peters (eds.), *Studies in machine translation and natural language processing*, 13–58. Luxembourg: European Commission.

Barnbrook Geoff and John M. Sinclair. 2001. Specialised corpus, local and functional grammar. In M. Ghadessy, A. Henry and R. L. Roseberry (eds.), *Small corpus studies and ELT: Theory and practice*, 237–276. Amsterdam: John Benjamins.

Blum-Kulka, Shoshana. 1986. Shifts of cohesion and coherence in tanslation. In J. House & S. Blum-Kulka (eds.), *Interlingual and intercultural communication*, 17–35. Tübingen: Gunter Narr.

BNC: The British National Corpus. http://bncweb.lancs.ac.uk/.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Christiansen, Morten H. & Nick Chater. 2016. The Now-or-Never Bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences* FirstView. 1–52.

COCA: The Corpus of Contemporary American English. http://corpus.byu.edu/coca/.

Cogo, Alessia. 2009. Accommodating difference in ELF conversations: A study of pragmatic strategies. In A. Mauranen & E. Ranta (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars. 254–273.

Cogo, Alessia. 2012. ELF and super-diversity: A case study of ELF multilingual practices from a business context. *Journal of English as a Lingua Franca* 1(2): 287–313.

Cogo, Alessia and Martin Dewey. 2006. Efficiency in ELF communication: From pragmatic motives to lexico-grammatical innovation. *Nordic Journal of English Studies* 5(2). 59–93. [Special Issue: English as a Lingua Franca, ed. by A. Mauranen & M. Metsä-Ketelä].

ELFA 2008. The Corpus of English as a Lingua Franca in Academic Settings. Director: Anna Mauranen. http://www.helsinki.fi/elfa/elfacorpus. (30 May 2016).

Ellis, Nick. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 143–188.

Firth, Alan. 1996. The discursive accomplishment of normality: On "lingua franca" English and conversation analysis. *Journal of Pragmatics* 26: 237–259.

Fröchlich, Maria, Nina Spada and Patrick Allen (1985) Differences in the communicative orientation of L2 classrooms. *TESOL Quarterly* 19(1): 27–57.

García, Ofelia and Li Wei (2014) *Translanguaging: Language, Bilingualism and Education*. New York: Palgrave Macmillan.

Greaves, Chris. 2009 *ConcGram 1.0: A phraseological search engine*. [Software]. Amsterdam: John Benjamins.

House, Juliane. 2014. Managing academic institutional discourse in English as a lingua franca. *Functions of Language* 21(1): 50–66.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.

Hynninen, Niina. 2016. *Language regulation in English as a lingua franca. Focus on academic spoken discourse*. Berlin: De Gruyter Mouton.

Jenkins, Jennifer. 2015. Repositioning English and multilingualism in English as a lingua franca. *Englishes in Practice* 2(3): 49–85.

Kaur, Jagdish. 2009. Pre-empting problems of understanding in English as a lingua franca. In A. Mauranen & E. Ranta (eds.), *English as a lingua franca: Studies and findings*, 107–123. Newcastle upon Tyne: Cambridge Scholars Publishing.

Klimpfinger, Theresa. 2007. 'Mind you, sometimes you have to mix' – The role of code-switching in English as a lingua franca. *Vienna English Working PaperS*, 16(2): 36–61.

Klimpfinger, Theresa 2009. "She's mixing the two languages together" – Forms and functions of code-switching in English as a lingua franca. In A. Mauranen & E. Ranta (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars. pp. 348–371.

Luk, Jasmine C. M. and Angel M. Y. Lin. 2006. *Classroom Interactions as cross-cultural encounters: Native speakers in EFL lessons.* Mahwah, NJ: Lawrence Erlbaum.

Mauranen, Anna. 2006. Signaling and preventing misunderstanding in English as lingua franca communication. *International Journal of the Sociology of Language* 177: 123–150.

Mauranen, Anna. 2007. Hybrid voices: English as the lingua franca of academics. In K. Fløttum (ed.), *Language and discipline perspectives on academic discourse*, 243–259. Newcastle: Cambridge Scholars Publishing.

Mauranen, Anna. 2012. *Exploring ELF. Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.

Mauranen, Anna. 2013. Lingua franca discourse in academic contexts: Shaped by complexity. In John Flowerdew (ed.), *Discourse in context*, 225–245. London & New York: Bloomsbury.

Mauranen, Anna. forthcoming. Temporality in speech – The Linear Unit Grammar.

Mauranen, Anna, Niina Hynninen & Elina Ranta. 2010. English as an academic lingua franca: The ELFA project. *English for Specific Purposes* 29: 183–190.

Mortensen, Janus. 2014. Language policy from below: Language choice in student project groups in a multilingual university setting. *Journal of Multilingual and Multicultural Development* 35(4): 425–442. [Special issue: Language policies and practices in the internationalisation of higher education on the European margins, ed. by Josep M. Cots, Enric Llurda & Peter Garrett.]

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing.* Cambridge: Cambridge University Press.

Pawley, Andrew and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards and R. W. Schmidt (eds.), *Language and communication*, 191–227. London: Longman.

Pietikäinen, Kaisa S. 2012. *English as a lingua franca couples in interview: Code-switching stimuli.* MA thesis. Newcastle University.

Pietikäinen, Kaisa S. 2014. ELF couples and automatic code-switching. *Journal of English as a Lingua Franca* 3(1): 1–26.

Pietikäinen, Kaisa S. 2016. Misunderstandings and ensuring understanding in private ELF talk. *Applied Linguistics*. DOI: 10.1093/applin/amw005.

Pietikäinen, Kaisa S. in press. ELF in social contexts. In: Jennifer Jenkins, Will Baker and Martin Dewey (eds.), *The Routledge handbook on English as a lingua franca*. Routledge.

Pölzl, Ulrike 2003. Signalling cultural identity: The use of L1/Ln in ELF. *Vienna English Working PaperS* 12(2): 3–23.

Pölzl, Ulrike and Barbara Seidlhofer. 2006. In and on their own terms: The "habitat factor" in English as a lingua franca interactions. *International Journal of the Sociology of Language* 177: 151–176.

Poplack, Shana. 1987 [1985]. Contrasting patterns of code-switching in two communities. In E. Wande, J. Anward, B. Nordberg, L. Steensland and M. Thelander (eds.), *Aspects of multilingualism*, 51–77. Uppsala: Borgströms.

Poplack, Shana, David Sankoff and Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26: 47–104.

Poplack, Shana, Susan Wheeler and Anneli Westwood. 1989. Distinguishing language contact phenomena: Evidence from Finnish–English bilingualism. *World Englishes* 8 (3): 389–406.

Rampton, Ben. 1995. *Crossing: Language and ethnicity among adolescents.* London: Longman.

Romaine, Suzanne. 2000. *Language in society: An introduction to sociolinguistics*. 2nd edn. Oxford: Oxford University Press.

Sinclair John M. and Anna Mauranen. 2006. *Linear unit grammar: Integrating speech and writing*. Amsterdam: John Benjamins.

Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tomasello, Michael. 2008. *Origins of human communication*. Cambridge: MIT Press.

Tomasello, Michael. 2014. The ultra-social animal: Ultra-social animal. *European Journal of Social Psychology* 44(3). 187–194.

Turunen, Kirsi. 2012. *A study on code-switching in the ELFA corpus*. MA thesis. University of Helsinki.