Jan Kochanowski University Press

"The aim of this paper is…": Frame markers in English as a lingua franca of academic writing

Tereza Guziurová

University of Ostrava

ABSTRACT

This study explores variation in metadiscourse patterns in English as a lingua franca academic writing. The paper aims to investigate discourse reflexivity in English-medium research articles written by non-native speakers from ten different L1 backgrounds included in the SciELF corpus. Specifically, the paper focuses on one reflexive category, frame markers, which signal text boundaries, announce discourse goals, and label text stages (Hyland 2005), thus making the discourse organisation more explicit. The corpus comprises 72 articles from the field of social sciences and humanities, totalling over 432,000 words. The findings are compared with a specialized corpus of 72 published research articles written by Anglophone authors (approximately 621,000 words), which has been designed as a corpus comparable to the SciELF. The results indicate differences in the forms and functions of certain frame markers in the two corpora, suggesting that this type of discourse reflexivity shows language and culture-specific diversity.

Keywords: metadiscourse, frame markers, English as a lingua franca, research articles, SciELF corpus.

1. Introduction

English as a lingua franca is now a relatively established research field. Most of ELF research has focused on spoken interactions (Cogo – Dewey 2012), which were analysed using spoken ELF corpora, such as the Vienna-Oxford International Corpus of English (VOICE) launched in 2001, or the corpus of English as a Lingua Franca in Academic Settings (ELFA) completed in 2008 (ELFA 2008). Studies on the ELFA corpus yielded interesting results

on the nature of spoken ELF. Mauranen (2017), for example, has identified several processes typical of ELF speech: on the one hand it is structural simplification, which can be manifested in the regularization of irregular forms, or lexical simplification, i.e. a tendency to higher representation of the most frequent lexical items in ELFA (in comparison with L1 English). On the other hand, there are complexifying tendencies such as approximation, when speakers use approximate equivalents of target expressions whose meaning is easily recognizable (e.g. *on the other side*). This process is not seen as an individual idiosyncrasy (such as omitting articles by a given speaker), but as a more general tendency of speakers across different speech events and lingua-cultural backgrounds, that is a property of ELF (Mauranen 2017).

One of the observed tendencies of ELF speech at discourse level is enhanced explicitness (Mauranen 2017). Since speakers from different linguistic and/or cultural backgrounds cannot often rely on the shared cultural context and their interactions might be associated with uncertainty or unpredictability, they tend to use clearer and more explicit expressions. Moreover, clarifying and explicating strategies are cooperative (Mauranen 2012: 167) as they increase the chances of getting the message across. Enhanced explicitness takes many forms (e.g. rephrasing in speech), but a common manifestation of explicitation is discourse reflexivity (Mauranen 2017: 246), which makes discourse organisation more visible and enhances clarity by guiding readers through the discourse. Mauranen et al. (2016: 4) argue that "discourse reflexivity is central to academic discourse, and particularly relevant for academic ELF, where it can help increase clarity and explicitness among speakers from different linguistic and cultural backgrounds" (Mauranen et al. 2016: 46). Moreover, discourse reflexivity in ELF was found to have similar functions, forms and distributions across different L1 speakers (ibid.).

What has received less attention so far is written academic ELF. Among the research topics that have already been addressed are, for example, phraseological approximations in spoken and written academic ELF (Carey 2013), hedging modal verbs in ELF research articles (Mur Dueñas 2016), the rhetorical structure of research article abstracts (Lorés-Sanz 2016), science blogs (Mauranen 2013), and multi-word units of meaning in L2 Master's theses (Vetchinnikova 2014). It should be mentioned that there exists a substantial body of research on academic discourse in English written by non-native speakers. Drawing on the tradition of contrastive rhetoric (Connor 1996), researchers examined grammatical, lexical and discursive

features of academic texts written by L2 English speakers and contrasted them with L1 English texts. The issue with this kind of research was that the writing produced by native speakers was perceived as an ideal, a norm that the others should aim to meet (Mauranen – Hynninen – Ranta 2016). Nevertheless, the majority of writers and readers of academic texts nowadays are likely to be academics with other than L1 English background because English has gradually become the lingua franca of academia (ibid.).

In 2015, the first large database of written academic ELF was compiled – the WrELFA corpus (Written English as a Lingua Franca in Academic Settings). It consists of 1.5 million words and has three parts: PhD examiners' reports, research blogs, and unedited research articles (the SciELF corpus). The last part contains 150 research articles written by L2 users of English from 10 different L1 backgrounds, which have not undergone professional proofreading or checking by English native speakers. The SciELF corpus has already been examined from several linguistic viewpoints (see e.g. Rowley-Jolivet 2017; Mur Dueñas 2018; Murillo 2018; Wu et al. 2020; Shchemeleva 2022). For example, Mur Dueñas (2018) has compared the use of a lexico-grammatical structure, the anticipatory *it* pattern, in ELF research articles and ENL research articles from similar disciplines. She has found that the *it* pattern most commonly expresses attitudinal meanings in both corpora, but generally it is used more frequently by ELF authors. Moreover, ELF scholars use some chunks which are absent in ENL articles, for example a wider range of discourse verbs in the pattern *It V-link ADJ to*, which suggests possible innovations or creativity when expressing interpersonal meanings in ELF (Mur Dueñas 2018).

Another study by Bondi and Borelli (2018) is particularly relevant for this paper because it investigates how non-native English speakers use metadiscursive resources in the economics component of the SciELF. Analysing positive (overused) and negative (underused) keywords in ELF in comparison with the reference corpus of published research articles, the study shows that the SciELF corpus is characterized by some prototypical metadiscursive elements, such as neutral evidentials (*according to*), general labelling nouns identifying elements of cognition (e.g. *characteristics*) and text (e.g. *paper*) rather than event (e.g. *change*) and discourse (e.g. *claim*), and also prototypical frame markers pointing to topic and focus (e.g. *consider*) rather than arguments (e.g. *show*). The ELF articles also show a marked underuse of textual and personal deixis (*I, this, these*) compared to published articles which are characterized by a greater authorial presence in the text (Bondi – Borelli 2018).

Generally, discourse reflexivity or metadiscourse is an interesting research area which can help reveal some distinctive features of written academic ELF. Compared with other writing cultures, Anglo-American academic English tends to be described as "reader-oriented" as it is associated with a higher level of interactivity and it puts the responsibility for clarity and understanding on the writer rather than the reader (Hyland 2005; Čmejrková 1996). In terms of discourse organization, it has been characterized as "more explicit about its structure and purposes", containing a noticeable amount of metadiscourse (Hyland 2005: 117; Swales – Feak 2012). On the other hand, as mentioned above discourse reflexivity proved to be particularly relevant for academic ELF, where mutual understanding and explicitness are important (and maybe more important than correctness) (Bondi – Borelli 2018). The aim of the present study is to investigate discourse reflexivity in written academic ELF, represented by the SciELF corpus, and to find out whether a higher level of explicitness typical of spoken ELF also characterizes ELF academic writing. The study focuses on one reflexive category, frame markers, which signal text boundaries or elements of text structure (Hyland 2005). They can sequence parts of the text, announce goals or label text stages and are signalled by such expressions as *first*, *next*, *my purpose is*, *to conclude* etc. Frame markers refer to discourse acts or stages, making discourse organization more explicit and accessible to readers.

Specifically, the present study addresses the following research questions:

1. What are the forms and functions of frame markers in ELF research articles?
2. Are there any differences in the use of frame markers between ELF research articles and published research articles written by Anglophone authors?

2. Frame markers in academic writing

Frame markers are part of Hyland's (2005) metadiscourse model that defines metadiscourse as "self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community" (Hyland 2005: 37). Understanding metadiscourse primarily as writer-reader interaction, Hyland draws a distinction between two types of metadiscourse – interactive and interactional. Interactive resources are used to organize

a text taking account of the reader's knowledge and processing needs. They comprise transition markers (conjunctions and adverbial phrases signalling relations in the text); endophoric markers which refer to other parts of the text (e.g. *noted above*); evidentials, i.e. citation practices; code glosses which help readers understand the text by rephrasing or explaining what has been said (*in other words*); and frame markers which signal text boundaries (*in conclusion*).

Interactional resources concern ways writers comment on and evaluate the content of propositions, expressing epistemic and affective stance. They also express writer-reader interaction as they aim to involve readers in the arguments. While interactional resources, including hedges, boosters and engagement markers, have been extensively researched in academic writing (e.g. McGrath – Kuteeva 2012; Hyland 2004; Hu – Cao 2015; Hyland – Jiang 2016; Dontcheva-Navratilova 2021, etc.), interactive resources have attracted less attention (probably with the exception of transition markers and reformulation markers).

Linguistic expressions referred to as frame markers are present in most conceptions of metadiscourse. For instance, one of the early taxonomies developed by Crismore et al. (1993) contains several categories that overlap with frame markers: sequencers which indicate ordering of material (*first*, *next*), topicalizers indicating a shift in topic (*now*; *in regard to*) and partly illocution markers which name the act performed (*to sum up*). Hyland (2005) reorganized Crismore et al.'s (1993) categories and introduced the notion of frame markers to denote rhetorical units which mark elements of text structure and help readers follow the development of the discourse. In his metadiscourse model, Hyland characterizes frame markers as follows:

> *Frame markers* signal text boundaries or elements of schematic text structure. […] Items included here function to sequence, label, predict and shift arguments, making the discourse clear to readers or listeners. (2005: 51)

Hyland does not elaborate the definition much further in his original framework, but he provides a more detailed account of individual categories in his later study (Hyland – Zou 2020). Specifically, frame markers can be divided into:

- **sequencers** used to sequence parts of the text or to internally order an argument (*first*, *next*);

- **labellers** which explicitly label text stages (*to summarize*, *in conclusion*); they often indicate the speech act that will be performed;
- **goal announcers** which state the author's purpose in the text (*my purpose is*, *the paper proposes*);
- **topic shifters** which indicate a shift in the direction of the text (*well*, *now*, *let us return to*) (Hyland – Zou 2020).

Frame markers thus seem to be salient in the overall organization of texts, especially of research articles, which need to be carefully structured to convey the author's intended message and also to conform to conventions of academic writing and genre expectations.

A number of studies have investigated frame markers, usually as part of the whole metadiscourse framework, across different genres, languages and disciplines. A detailed analysis of frame markers is provided by Hyland and Zou (2020) who explored how academics recontextualise their scientific findings from journal articles to academic blogs. Analysing 50 blogs and 50 articles with the same authors and topics, they have found that bloggers generally use more frame markers to present complex research material to lay audience. They have also shown that labellers and topic shifters occur more frequently in research articles to assist specialists follow longer and more complex arguments, while sequencers (especially listing sequencers) are more numerous in blogs to help general readers understand connections in the text, facilitating processing of the material (Hyland – Zou 2020).

Cross-linguistic studies of metadiscourse have confirmed that its use in academic writing is influenced by writers' linguistic and cultural backgrounds (Mauranen 1993; Peterlin 2005; Mur Dueñas 2011; Mu et al. 2015). For example, Mauranen (1993) in her seminal study of text reflexivity finds that Anglo-American writers use more reflexive text than Finnish writers, which indicates that Anglo-American writers are more concerned with guiding the reader in the discourse, showing more explicit presence in the text. A comparison of the use of metadiscourse in English and Chinese research articles from applied linguistics has revealed that English articles contain more interactive metadiscourse features, including frame markers, which is attributed to different rhetorical strategies between the languages and the fact that Chinese writers share more background knowledge as they address the local discourse community (Mu et al. 2015). Dahl (2004) investigated the relation between language and discipline as two important variables influencing the use of metadiscourse. She examined textual metadiscourse in research articles across three languages (English, French

and Norwegian) and three disciplines (economics, linguistics and medicine). Her findings suggest that the language variable is key within economics and linguistics, where English and Norwegian display similar patterns, using frame markers much more than French. The situation is different in medical texts, which show almost identical frequencies of frame markers in all three languages, pointing to stable disciplinary practices in medicine (Dahl 2004: 1822).

To summarise, previous studies have shown that frame markers as part of interactive metadiscourse are important tools of organising academic texts and they are sensitive to their social and rhetorical context (Hyland – Zou 2020). Since this study focuses on unedited research articles which have not been shaped by language professionals, reviewers, editors and other "literacy brokers" (Lillis – Curry 2010), the results might shed light on how L2 English scholars structure their scientific texts and what their rhetorical preferences are.

3. Data and methodology

In order to investigate how L2 writers use frame markers in English as a lingua franca academic writing, two corpora have been compared. The first one is composed of articles from the SciELF corpus, which comprises 150 unedited research articles written by L2 users of English, totalling 759,300 words. The papers have not undergone professional proofreading or checking by English native speakers, and most of them are final drafts of unpublished manuscripts. The corpus is divided into two broad disciplinary domains – sciences (labelled 'Sci') and social sciences and humanities (labelled 'SSH'). The Sci part contains 78 articles (326,463 words), which are drawn from natural sciences (79%), medicine (18%) and agriculture and forestry (3%). The SSH part contains 72 articles (432,837 words) and it includes texts from social sciences (45%), humanities (34%) and behavioural sciences (21%). Since disciplinary variation in the use of metadiscourse is high (Hyland 2005), I have limited my focus to social sciences and humanities articles (the SciELF-SSH subcorpus) in this study.

The authors of the papers come from ten different L1 backgrounds (see Table 1). The number of articles in each language group varies, but the goal of ELF research is not to make L1-based comparisons, but to examine how people from different language backgrounds use English as a lingua franca. Specific disciplines represented in the corpus are psychology, sociology,

educational sciences, economics, classical philology, linguistics, philosophy, law, theatre studies, anthropology, history, urban design, literary studies, information sciences, social policy, and art history (SciELF 2015).

Table 1. Distribution of the L1 language groups in the SciELF-SSH corpus

First author's L1	No. of articles	No. of words	% of total words
Chinese	10	44,196	10%
Czech	10	59,569	14%
Finnish	10	59,118	14%
French	8	48,373	11%
Italian	5	31,249	7%
Portuguese (Brazil)	6	39,223	9%
Romanian	4	25,197	6%
Russian	6	38,834	9%
Spanish	7	51,383	12%
Swedish	6	35,695	8%
Total	**72**	**432,837**	**100%**

The SciELF–SSH corpus is contrasted with a reference corpus, which was compiled in 2019 (labelled 'CSSH'). The reference corpus is composed of 72 English-medium articles, which were published in prestigious academic journals. The CSSH corpus is comparable in terms of disciplines so the number of articles in each discipline corresponds to those in the SciELF-SSH corpus. For the purposes of corpus compilation, some specific disciplines were subsumed into a more general category; for instance, an article from corpus linguistics was included in the 'linguistics' category. The articles were drawn from the Web of Science and Scopus databases. I selected only well-established journals based on their impact factor (Web of Science data 2018) or SCImago rankings (Scopus 2018); only journals in Quartile 1 were included. If one discipline was represented by numerous articles in the SciELF (e.g. economics), papers from several journals were selected; therefore, the CSSH corpus is comprised of research articles from 41 different journals altogether.

With regard to the language background of the authors, the intention was to select articles written by Anglophone writers. The SciELF corpus includes 45 single-authored articles and 27 multi-authored ones, so the reference corpus has been compiled accordingly. In the case of single-authored articles, it was possible to assume that the author's L1 was English

based on their names, affiliations and professional CVs (if available). The situation was more complicated with the multi-authored ones, since it was sometimes difficult to find texts where all the authors were native speakers. Therefore, in the case of multi-authored articles, the first author is always a native speaker of English, but in 10% of papers not all the other authors are L1 English speakers. However, since all the papers have been published in highly ranked journals, it is safe to assume that the texts have undergone editorial changes and/or have been proofread.

The CSSH corpus comprises the same number of articles as the SciELF-SSH, i.e. 72, totalling 621,267 words, with 8,629 average word count per article. The articles are on average longer than those in the SciELF, which is mainly caused by the length of economic papers published in high-impact journals; the average word count in economic papers is 13,112. All the papers were published between 2016-2018. Table 2 shows the composition of both corpora.

Table 2. Composition of the SciELF-SSH corpus and CSSH corpus

	SciELF-SSH corpus	CSSH corpus
No. of articles	72	72
No. of words	432,837	621,267
Average words /article	6,012	8,629
Texts	final drafts of unpublished RAs	published RAs
Disciplines (no. of articles)	educational sciences (15), economics (15), linguistics (13), sociology (6), psychology (3), history (3), social policy (3), classical philology (2), philosophy (2), law (2), theatre studies (2), anthropology (2), urban design (1), literary studies (1), information sciences (1), art history (1)	educational sciences (15), economics (15), linguistics (13), sociology (6), psychology (3), history (3), social policy (3), classical philology (2), philosophy (2), law (2), theatre studies (2), anthropology (2), urban design (1), literary studies (1), information sciences (1), art history (1)

Prior to the analysis, the texts in the reference corpus were processed. I followed the *SciELF Corpus Manual* (Carey 2015) to make sure that the corpora were comparable, so bibliographic references, block quotes and long stretches of foreign text were omitted in the plain text files (used for concordances), while abstracts were kept.

First, the corpora were searched for specific features which could potentially act as frame markers using *AntConc* (Anthony 2019). The selection of frame markers was based on the list from Hyland and Zou's study (2020) and my previous research on metadiscourse (Guziurová 2018), including the subcategories of sequencers, labellers, goal announcers and topic shifters. Then all retrieved items were examined in context to ensure they functioned as frame markers. Since one of Hyland's (2005) principles of metadiscourse is that it is distinct from propositional aspects of discourse (see also Mauranen 1993), all cases had to be manually checked to ensure they were reflexive. For example, when the word *now* refers to the extralinguistic context functioning as a typical time adverbial, it is not regarded as reflexive (1). However, when it refers to the current text, organizing writer's arguments, it is classified as a frame marker (2).

(1)   World Travel & Tourism Council estimates that from direct and indirect activities, the tourism sector *now* provides a remarkable 9.2% of world GDP. (SSH51)

(2)   Let us *now* proceed to a systematic overview, and a further characterisation, of the four varieties. (SSH11)

Since the two corpora are not equal in size, all the results have been normalized to 100,000 words. Then log-likelihood tests were performed to determine differences of statistical significance. All statistical tests were performed on raw data using the online calculator developed by Paul Rayson (Rayson 2008). If the p-value was <0.05 (the threshold level usually set in linguistics), the results were regarded as statistically significant. The results will be discussed in the following section.

4. Results and discussion

4.1 An overview of frame markers in the two corpora

Overall, the results of the quantitative analysis show that there are no significant differences in the frequency of frame markers between the two corpora. As Table 3 demonstrates, frame markers are more numerous in published articles written by Anglophone authors, but the difference is not statistically significant ($G^2 = 0.13$, $p > 0.05$). However, the distribution of individual types of frame markers varies in the two corpora. In the ELF

corpus, the most prominent are goal announcers, accounting for almost 40% of all frame markers. In the CSSH corpus, sequencers are the most frequent category (40.9% of frame markers), with goal announcers being the second most frequent type, accounting for 32% of the total. In both corpora, topic shifters are the least numerous; however, ELF writers use almost twice as many shifters as Anglophone writers and the difference is statistically significant ($G^2 = 10.83$, p<0.01).

Table 3. Overall incidence of frame markers in the two corpora

Frame markers	SciELF-SSH corpus			CSSH corpus		
	Total no.	Freq. per 100,000 w	%	Total no.	Freq. per 100,000 w	%
Sequencers	195	45.0	27.4	425	68.4	40.9
Goal announcers	281	64.9	39.5	335	53.9	32.2
Labellers	163	37.7	22.9	221	35.6	21.2
Topic shifters	73	16.9	10.2	59	9.5	5.7
Total	**712**	**164.5**	**100**	**1040**	**167.4**	**100**

The distribution of individual subcategories in the CSSH corpus is in line with Hyland and Jiang's (2020) study of interactive metadiscourse, which shows that sequencing devices are by far the most frequent frame markers in soft fields, followed by goal announcers. Similarly, sequencers were most prominent in Cao and Hu's (2014) study of published research articles in three soft disciplines. In the following sections, each subcategory of frame markers is discussed in detail.

4.2 Sequencing text

Sequencers in Hyland's (2005) model correspond to "linking adverbials" expressing enumeration (Biber et al. 1999) or a category of "listing conjuncts" in traditional Quirk et al.'s (1985) terms. More specifically, they form an open class of "enumerative conjuncts" which give a particular structure to a list; the enumerative function "connotes relative priority and endows the list with an integral structure, having a beginning and an end" (Quirk et al. 1985: 636). Since they explicitly signal connections between units of discourse, they are important cohesive devices.

As mentioned above, only sequencers expressing internal relations in discourse are considered metadiscursive. This is in line with internal and external types of conjunctive relations distinguished by Halliday and Hasan (1976). Therefore, example (3) is regarded as reflexive, since the frame marker organizes writer's arguments, but example (4) is considered propositional, as it expresses the time sequence of events in the world (in this particular case, the process of data collection).

(3)   *Finally*, I will present evidence of change in students' understanding of the CCC and DCI as represented in conceptual models […]. (CSSH17)

(4)   *Finally*, I collected the instructional artifacts Jennifer used while teaching this unit, including presentation slides, handouts, grading rubrics, and assigned readings. (CSSH06)

Table 4 shows the frequency of sequencers in the SciELF-SSH corpus and the reference corpus. The results demonstrate that published RAs contain significantly more sequencers than ELF articles ($G^2 = 24.41$, $p < 0.0001$). The difference is particularly striking with enumerating markers *First/Firstly*, *Second/Secondly* and *finally*. This is in line with Hyland and Jiang (2020), who found that although sequencers have declined in soft knowledge fields (applied linguistics and sociology) over the last 50 years and increased in hard sciences, the forms showing the largest rises overall in their corpus are *first*, *second* and *finally*.

Table 4. Frequency of sequencers in the SciELF-SSH and CSSH corpora

| Marker | SEQUENCERS | | | |
| | SciELF-SSH corpus | | CSSH corpus | |
	Total no.	Frequency per 100,000 words	Total no.	Frequency per 100,000 words
finally	42	9.7	97	15.6
First / Firstly	56	12.9	132	21.2
Last / lastly	13	3.1	11	1.8
Next	6	1.4	9	1.5
Second / Secondly	43	9.9	109	17.5
subsequently	0	0	1	0.2
Then	11	2.6	3	0.5
Third / Thirdly	8	1.8	43	6.9

begin	8	1.8	16	2.6
start	8	1.8	4	0.6
Total	**195**	**45.0**	**425**	**68.4**

It is also interesting to note that CSSH authors overwhelmingly prefer the forms without -ly, i.e. *First* and *Second*, as linking adverbials (95% of cases), while there is more variability among ELF authors. *First* is preferred over *firstly* in 71%, whereas *Secondly* is favoured over the shorter form in 65% of cases in ELF articles. Since –ly is a suffix commonly associated with adverbs, L2 English writers might be more inclined to use this form in the function of a frame marker.

    Sequencing devices fulfil a range of functions in research articles. In the introduction, they can signal the overall organization of the paper, as in example (5). They can also present contributions of the study (6) when "occupying the niche" that exists in previous research (Swales – Feak 2012). In the conclusion, sequencers can help summarize findings, present implications for practice or list limitations of the research (7). In general, sequencing devices structure writer's arguments, explanations, reasons, findings, methodological difficulties, aims etc. Numerical sequencers are especially favoured in the genre of research article, where they contribute to the clearness and legibility of the text, facilitating the reader's understanding of the message (Hyland – Zou 2020).

(5)   The paper unfolds as follows: *First*, I review the literature on age, access, and motivation and then utilize this literature to propose a series of relationships that comprise the model. *Next*, I present the methods and results for a study that provides support for the hypothesized relationships. (CSSH41)

(6)   By more accurately measuring agency costs imposed by criminal prosecutions, the present research contributes to understanding both corporate crime and agency theory in three ways. *First*, confirming predictions made by applying agency theory to a legal context defends agency theory against growing skepticism about its predictive validity and calls to reconceptualize it (e.g., Lan & Heracleous, 2010; Pepper & Gore, 2015). (CSSH26)

(7)   It also appeared that in some respects the results were biased due to the corpus used. *Firstly*, the size of the corpus was not large enough to produce a sufficient number of valid instances. (SSH22)

4.3 Announcing goals

The expression of purpose is a powerful rhetorical tool in research articles, and writers are aware of its value in signalling the direction of their arguments (Hyland – Zou 2020). The data show the importance of goal announcers in both corpora as they account for the largest subcategory in ELF articles and the second largest in the CSSH corpus (see Table 5). While ELF writers use more goal announcers than CSSH writers, and the difference is statistically significant ($G^2 = 5.23$, $p<0.05$), the effect size measured by Phi coefficient is very low ($\Phi = 0.0022$), which suggests a very small correlation.

Table 5. Frequency of goal announcers in the SciELF-SSH and CSSH corpora

Marker	GOAL ANNOUNCERS			
	SciELF-SSH corpus		CSSH corpus	
	Total no.	Frequency per 100,000 words	Total no.	Frequency per 100,000 words
aim*	68	15.7	32	5.2
desire* to	0	0	0	0
focus*	44	10.2	78	12.6
goal*	14	3.2	5	0.8
intend* to	6	1.4	4	0.6
intention	4	0.9	3	0.5
objective*	7	1.6	4	0.6
outline*	7	1.6	16	2.6
propose*	29	6.7	48	7.7
purpose*	32	7.4	28	4.5
seek*	5	1.2	16	2.6
want* to	8	1.8	8	1.3
wish* to	3	0.7	4	0.6
would like to	14	3.2	1	0.2
(in this) section	40	9.2	88	14.2
Total	**281**	**64.9**	**335**	**53.9**

The importance of explicitly stating author's purposes has been corroborated by a diachronic study of Hyland and Jiang (2020) that followed the changes of interactive metadiscourse in research articles over the past 50 years. The

study has shown that the only subcategory of frame markers which increased across all disciplines was announcing goals and purposes, ensuring that objectives can be understood by readers (Hyland – Jiang 2020).

Goal announcers can indicate the overall purpose of the text (8) or a local goal for a specific part of the text (9). Outlining the overall purpose of the text is a key part of journal articles; it is one of the obligatory moves of RA introductions in the "create-a-research-space" (CARS) model (Swales – Feak 2012: 331).

(8)   *The aim of this article is to* analyse family businesses and approaches to them from the perspective of two extremely different cultures. (SSH14)

(9)   *In the next section*, *we* succinctly *present* the theoretical background that supports the development of this study. (SSH48)

Stating purposes can take a wide range of forms and the two corpora differ in this respect. The most common marker among ELF writers is *aim*, accounting for 24% of all goal announcers. As Table 5 indicates, the marker is almost three times more common in the Sci-ELF corpus than in the CSSH corpus. This is in line with Bondi and Borelli (2018), who found that *aim* was one of the positive (overused) keywords in the economics segment of SciELF, especially in its nominal form. Example (8) is thus a typical statement of purpose in ELF articles. In the CSSH corpus, the range of means expressing writer's purposes was more varied. The most frequent marker in the table, *(in this) section*, was followed by a range of verbs, such as:

(10)   *This / The next section discusses / presents / describes / examines* etc. (CSSH)
In this section, we investigate / explore / argue etc. (CSSH)

It also indicates that CSSH writers take greater effort in stating what will be presented in each part of the article, indicating local goals.

Another difference between the corpora can be found in the use of *goal*. This marker was not favoured by either group: the SciELF corpus includes 14 occurrences, out of which 10 are followed by a textual product (11). In published RAs there are only 5 occurrences of *goal* in the function of a frame marker, none of which collocate with a textual product. Rather, the authors talked about goals of *research*, *exploration* or *study* (in the sense of research) (12).

(11)   *The main goal of this paper* is to assess if socio-demographic determinants
        play a different role in European countries characterised by different
        welfare systems […]. (SSH63)

(12)   *One goal of this exploration* is to identify this phenomenon and provide
        a systematic review of how it operates. (CSSH36)

Finally, the corpora differ in the use of the structure *would like to*, which occurs
14 times in the SciELF as a frame marker, but only once in the CSSH. Overall,
ELF writers were more tentative in framing the discourse, using hedges
(e.g. *would like to, try to, endeavour, attempt to)*. This is not to say that they used
more hedges in general, but that they were more tentative in announcing
their goals (13) or labelling discourse stages (14) than Anglophone writers.
This expression of modesty and politeness might help them gain acceptance
of their claims by the discourse community or shield them from potential
criticism.

(13)   *In my paper I would like to show* how Constantine's participation in the
        Council of Nicaea (325) has changed the political thought of Christian
        theologians. (SSH41)

(14)   *We will now try to summarise* the main problems that lexicographers
        have to face when considering to write/publish a dictionary of
        collocations, with regards to the definition of collocation. (SSH42)

4.4 Labelling text stages

Another subcategory of frame markers labels discourse stages. These
expressions "signal the current discourse activity and offer an explicit means
for writers to mark upcoming text stages or rhetorical functions" (Hyland –
Zou 2020: 38). According to Hyland and Jiang (2020), labellers in academic
writing are most commonly used to summarise or draw conclusions from an
argument at certain points in the text. The writers thus explicitly point out
how readers should interpret the preceding message (15). However, labellers
can also pause the discourse and signal what will follow (16).

(15)   *To summarise*, better environmental performance is not primarily
        influenced by SL of any of the three frames. (SSH71)

(16)   However, there are two problems, especially pertinent to my current
        effort, that I would like to mention *at this point*. (SSH11)

The frequency of markers labelling stages is similar in both corpora (see Table 6). It seems that recapitulating key points in the text, or announcing what will come next, are "conventional rhetorical strategies" in research articles (Hyland – Zou 2020).

Table 6. Frequency of labellers in the SciELF-SSH and CSSH corpora

| Marker | LABELLERS | | | |
| | SciELF-SSH corpus | | CSSH corpus | |
	Total no.	Frequency per 100,000 words	Total no.	Frequency per 100,000 words
all in all	2	0.5	0	0
at this point	6	1.4	5	0.8
at this stage	0	0	0	0
briefly	3	0.7	3	0.5
conclu*	67	15.5	52	8.4
in a word	1	0.2	0	0
in brief	1	0.2	0	0
in short	4	0.9	13	2.1
in sum	7	1.6	15	2.4
in summary	0	0	16	2.6
now	29	6.7	11	1.8
on the whole	1	0.2	1	0.2
overall	8	1.8	62	10.0
repeat*	2	0.5	6	0.9
so far	6	1.4	2	0.3
summaris*	21	4.9	29	4.7
thus far	0	0	5	0.8
to sum up	5	1.2	1	0.2
TOTAL	**163**	**37.7**	**221**	**35.6**

Although the overall frequency of labellers across the corpora is very similar, some of the markers are used differently. The greatest difference can be seen in the use of *overall* as a sentence adverb. While this is the most common labeller in the CSSH corpus (10 occurrences per 100,000 words), it is rare in the ELF corpus (1.8 instances per 100,000 words). This expression seems to be a useful tool for writers to summarise their findings (17), summarise results of previous studies or provide a synthesis of arguments.

(17) *Overall*, these results revealed that there are different types of participants who play CRYSTAL ISLAND, and by being able to differentiate and identify these types of participants, we can move toward developing adaptive GBLEs that scaffold participants based on their gameplay behaviors. (CSSH63)

It should be noted that the marker *now* is multifunctional, as is often the case with metadiscourse devices. The respective functions are not connected with the adverb alone but result from the structure in which it occurs, i.e. the semantics of the whole phrase, and the context. If *now* occurred with illocutionary verbs or similar expressions indicating the function of the part of text (such as *mention*, *discuss*, *distinguish* or *introduce*), it was classified as a discourse label (examples 18 and 19). If *now* followed or preceded verbs which clearly indicate the change of the topic (*turn to*, *switch to*, *move on*), it was regarded as a topic shifter (20). Although there were some ambiguous cases, the context usually helped determine which category it belonged to.

(18) As I began to write this article, I invited Anna to read and respond to the analysis of her narrative accounts. […] My sense was that she may still have found the incident uncomfortable to some degree and did not wish to return to it again. I *now* introduce Anna and the setting in which the reported conflict took place. (CSSH20)

(19) However, the three dictionaries mentioned by *now* [fino ad ora] do not include idioms within their entries. (SSH42)

(20) Let me turn *now* to the significance of (rare) exceptions to Conceptual Role Determinism. (CSSH71)

4.5 Shifting topics

The last type of frame markers are topic shifters. They indicate a change in the direction of the discourse, moving from one issue the writer wishes to address to another (Hyland – Zou 2020). Functionally, they correspond to a category of topicalizers introduced by Mur Dueñas (2011) which are used to change the topic, introduce related topics or resume a topic introduced earlier. Topic shifters account for the lowest proportion of frame markers in both corpora, so it is not possible to draw any general conclusions. However, the results in Table 7 indicate that they were used more frequently by ELF writers, and the difference is statistically significant ($G^2 = 10.83$, p<0.01). This

finding is consistent with Mur Dueñas (2011), who reveals that topicalizers (together with endophoric markers) are the only metadiscourse devices which have been found to be significantly more frequent in the Spanish RAs than in English RAs written by American scholars. She concludes that it makes argumentation in Spanish RAs less linear and speculates that it might be one of the traits of texts written in languages favouring "reader-responsible style" (Hinds 1987), such as Spanish.

Table 7. Frequency of topic shifters in the SciELF-SSH and CSSH corpora

Marker	TOPIC SHIFTERS			
	SciELF-SSH corpus		CSSH corpus	
	Total no.	Frequency per 100,000 words	Total no.	Frequency per 100,000 words
back to	5	1.2	4	0.6
digress	1	0.2	0	0
in regard to	1	0.2	0	0
mov* on	2	0.5	1	0.2
now	11	2.5	9	1.4
regarding	31	7.2	8	1.3
resum*	0	0	0	0
return* to	4	0.9	12	1.9
revisit*	2	0.5	1	0.2
shift* to	1	0.2	2	0.3
to look more closely	0	0	1	0.2
turn* to	9	2.1	16	2.6
well	0	0	0	0
with regard to	6	1.4	5	0.8
Total	**73**	**16.9**	**59**	**9.5**

As can be seen in Table 7, the preposition *Regarding* which introduces a new or related topic in the discourse (21) is characteristic of ELF papers, accounting for 42% of all topic shifters. The expression is simple in terms of lexico-grammatical choices and seems to be a useful signal of a change of focus for the reader. In published RAs, the expressions are more varied, including more personal structures such as (22).

(21)   *Regarding* gender differences, men are more prone to volunteer in professional and political activities as well as in education and leisure activities. (SSH63)

(22)   *We now turn to* discussing the theoretical and managerial implications of the research findings. (CSSH44)

This confirms Bondi and Borelli's (2018) observation that ELF articles show an underuse of personal self-reference in comparison with published RAs.

5. Conclusions

The present study has investigated one aspect of discourse reflexivity, frame markers, in English as a lingua franca academic writing. The findings have shown that ELF writers recognise the importance of this kind of discourse organisation since there is no significant difference between the frequency of frame markers in their articles and published RAs written by Anglophone authors. The ELF writers in the corpus are mostly junior or senior academics (not e.g. undergraduate students) who are clearly aware of rhetorical practices in their fields. However, there is considerable variation in the individual types of frame markers across the two corpora, with published RAs containing more sequencers that explicitly organise arguments and fewer topic shifters which indicate digressions and make the discourse possibly less linear.

       The results show that ELF is in many respects similar to ENL as writers must recognize rhetorical expectations of their readers to be successful in their argumentation. Moreover, the research article is a key academic genre which is largely standardized, and writers need to follow disciplinary conventions. Nevertheless, ELF writers tend to use different means of discourse organisation, as the ELF corpus seems to be characterized by a prototypical use of certain devices announcing goals (e.g. *aim*), more tentative presentation of discourse goals and labelling stages, and also simplification as the ELF writers rely on simpler structures (*regarding*) and a limited number of devices. This may support the claim that even though writers from different cultural backgrounds do not have to differ in the overall amount of metadiscourse used, they may diverge in specific realisations of metadiscursive categories (Dontcheva-Navratilova 2021).

The results of this small-scale study should be interpreted with caution since there are several methodological limitations. First, this type of corpus-based metadiscourse research relies on a limited set of metadiscursive items which cannot capture all types of frame markers in the texts. It would also be desirable to investigate other types of interactive metadiscourse, such as endophoric markers or transitions, to get an overall picture of how ELF writers organise their texts. Furthermore, we would need more qualitative studies focusing on specific textual histories (see Lillis – Curry 2010), which would help us understand the role of "literacy brokers" in the process of publishing scientific texts and how metadiscourse is concerned. For example, the case study of Flowerdew and Wang (2016) suggests that revision changes of articles written by Chinese doctoral students concerned among other things textual cohesion and additions of some metadiscursive elements (e.g. endophoric markers). If we knew the final published versions of the articles in the SciELF corpus, we would be able to identify these changes. In spite of these limitations, the current study has identified language variation in metadiscourse patterns on the example of English as a lingua franca academic writing.

REFERENCES

Primary sources

ELFA. The Corpus of English as a Lingua Franca in Academic Settings.
http://www.helsinki.fi/elfa, accessed June 2022.
SciELF. The SciELF Corpus.
https://www2.helsinki.fi/en/researchgroups/english-as-a-lingua-franca-in-academic settings/research/wrelfa-corpus/scielf-corpus, accessed June 2022.

Secondary sources

Anthony, L.
2019   *AntConc* (Version 3.5.8). Tokyo: Waseda University. Available from https://www.laurenceanthony.net/software, accessed June 2022.
Biber, D. et al.
1999   *Longman Grammar of Spoken and Written English.* Harlow: Pearson Education Limited.

Bondi, M. – C. Borelli
    2018   "Publishing in English: ELF writers, textual voices and
              metadiscourse". In: P. Mur-Dueñas – J. Šinkūnienė (eds.) *Intercultural
              Perspectives on Research Writing*. Amsterdam/Philadelphia: John
              Benjamins, 217-235.
              https://doi.org/10.1075/aals.18.10bon.
Cao, F. – G. Hu
    2014   "Interactive metadiscourse in research articles: A comparative study
              of paradigmatic and disciplinary influences", *Journal of Pragmatics* 66,
              15-31. https://doi.org/10.1016/j.pragma.2014.02.007.
Carey, R.
    2013   "On the other side: Formulaic organizing chunks in spoken and written
              academic ELF", *Journal of English as a Lingua Franca* 2 (2), 207-228.
              https://doi.org/10.1515/jelf-2013-0013.
    2015   *SciELF Corpus Manual. ELFA Project*. University of Helsinki.
Čmejrková, S.
    1996   "Academic writing in Czech and English". In: E. Ventola –
              A. Mauranen (eds.) *Academic Writing*. Amsterdam/Philadelphia: John
              Benjamins, 137-152. https://doi.org/10.1075/pbns.41.11cme.
Cogo, A. – M. Dewey
    2012   *Analysing English as a Lingua Franca: A Corpus-driven Investigation*.
              London: Continuum.
Connor, U.
    1996   *Contrastive Rhetoric: Cross-Cultural Aspects of Second Language Writing*.
              Cambridge: Cambridge University Press.
Crismore, A. – R. Markkanen – M. Steffensen
    1993   "Metadiscourse in persuasive writing", *Written Communication* 10 (1),
              39-71. https://doi.org/10.1177/0741088393010001002.
Dahl, T.
    2004   "Textual metadiscourse in research articles: A marker of national culture
              or of academic discipline?", *Journal of Pragmatics* 36 (10), 1807-1825.
              https://doi.org/10.1016/j.pragma.2004.05.004.
Dontcheva-Navratilova, O.
    2021   "Engaging with the reader in research articles in English: Variation
              across disciplines and linguacultural backgrounds", *English for Specific
              Purposes* 63, 18-32. https://doi.org/10.1016/j.esp.2021.02.003.
Flowerdew, J. – S. H. Wang
    2016   "Author's editor revisions to manuscripts published in international
              journals", *Journal of Second Language Writing* 32, 39-52.
              https://doi.org/10.1016/j.jslw.2016.03.004.
Guziurová, T.
    2018   *Metadiscourse in Undergraduate Textbooks and Research Articles in
              Linguistics*. Ostrava: University of Ostrava.

Halliday, M. A. K. – R. Hasan
    1976   *Cohesion in English.* Harlow: Longman.
                https://doi.org/10.4324/9781315836010.

Hinds, J.
    1987   "Reader versus writer responsibility: A new typology". In: U. Connor – R.B. Kaplan (eds.) *Writing across Languages: Analyses of L2 Texts.* Reading: Addison-Wesley, 141-152.

Hu, G. – F. Cao
    2015   "Disciplinary and paradigmatic influences on interactional metadiscourse in research articles", *English for Specific Purposes* 39, 12-25. https://doi.org/10.1016/j.esp.2015.03.002.

Hyland, K.
    2004   "Engagement and disciplinarily: The other side of evaluation". In: G. Del Lungo Camiciotti – E. Tognini Bonelli (eds.) *Academic Discourse – New Insights into Evaluation.* Bern: Peter Lang, 13-30. https://doi.org/10.3726/978-3-0351-0784-5.
    2005   *Metadiscourse: Exploring interaction in writing.* London: Continuum.

Hyland, K. – F. (K.) Jiang
    2016   "'We must conclude that…': A diachronic study of academic engagement', *Journal of English for Academic Purposes* 24, 29-42. https://doi.org/10.1016/j.jeap.2016.09.003.
    2020   "Text-organizing metadiscourse: Tracking changes in rhetorical persuasion", *Journal of Historical Pragmatics* 21 (1), 137-164. https://doi.org/10.1075/jhp.00039.hyl.

Hyland, K. – H. (J.) Zou
    2020   "In the frame: Signalling structure in academic articles and blogs", *Journal of Pragmatics* 165, 31-44. https://doi.org/10.1016/j.pragma.2020.05.002.

Lillis, T. – M. J. Curry
    2010   *Academic Writing in a Global Context. The Politics and Practices of Publishing in English.* London/New York: Routledge.

Lorés Sanz, R.
    2016   "ELF in the making? Simplification and hybridity in abstract writing", *Journal of English as a Lingua Franca* 5 (1), 53-81. https://doi.org/10.1515/jelf-2016-0003.

Mackenzie, I.
    2014   *English as a Lingua Franca. Theorizing and Teaching English.* London/New York: Routledge. https://doi.org/10.4324/9781315890081.

Mauranen, A.
    1993   *Cultural Differences in Academic Rhetoric: A Textlinguistic Study.* Frankfurt: Peter Lang.
    2012   *Exploring ELF. Academic English Shaped by Non-native Speakers.* Cambridge: Cambridge University Press.

2013   "Hybridism, edutainment, and doubt: Science blogging finding its feet", *Nordic Journal of English Studies* 12 (1), 7-36. https://doi.org/10.35360/njes.274.

2017   "A glimpse of ELF". In: M. Filppula et al. (eds.) *Changing English: Global and Local Perspectives*. Berlin/Boston: De Gruyter, 223-254. https://doi.org/10.1515/9783110429657-013.

Mauranen, A. – N. Hynninen – E. Ranta

2016   "English as the academic lingua franca". In: K. Hyland – P. Shaw (eds.) *The Routledge Handbook of English for Academic Purposes.* Abingdon: Routledge, 44-55. https://doi.org/10.4324/9781315657455.

McGrath, L. – M. Kuteeva

2012   "Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices", *English for Specific Purposes* 31, 161-173. https://doi.org/10.1016/j.esp.2011.11.002.

Mu, C. et al.

2015   "The use of metadiscourse for knowledge construction in Chinese and English research articles", *Journal of English for Academic Purposes* 20, 135-148. https://doi.org/10.1016/j.jeap.2015.09.003.

Mur-Dueñas, P.

2011   "An intercultural analysis of metadiscourse features in research articles written in English and in Spanish", *Journal of Pragmatics* 43 (13), 3068-3079. https://doi.org/10.1016/j.pragma.2011.05.002.

2016   "Modal hedging verbs in English as a Lingua Franca (ELF) Business Management research articles", *Kalbotyra* 69, 153-178. https://doi.org/10.15388/Klbt.2016.10371.

2018   "Exploring ELF manuscripts: An analysis of the anticipatory *it* pattern with an interpersonal function". In: P. Mur-Dueñas – J. Šinkūnienė (eds.) *Intercultural Perspectives on Research Writing*. Amsterdam/Philadelphia: John Benjamins, 277-297. https://doi.org/10.1075/aals.18.13mur.

Murillo, S.

2018   "Not the same, but how different? Comparing the use of reformulation markers in ELF and in ENL research articles". In: P. Mur-Dueñas – J. Šinkūnienė (eds.) *Intercultural Perspectives on Research Writing*. Amsterdam/Philadelphia: John Benjamins, 237-253. https://doi.org/10.1075/aals.18.11mur.

Peterlin, A. P.

2005   "Text-organising metatext in research articles: An English-Slovene contrastive analysis", *English for Specific Purposes* 24 (3), 307-319. https://doi.org/10.1016/j.esp.2004.11.001.

Quirk, R. et al.

1985   *A Comprehensive Grammar of the English Language.* London: Longman.

Rayson, P.
    2008   "Log-likelihood and effect size calculator."
          http://ucrel.lancs.ac.uk/llwizard.html, accessed June 2022.

Rowley-Jolivet, E.
    2017   "English as a Lingua Franca in research articles: The SciELF corpus",
          ASp 71, 145-158. https://doi.org/10.4000/asp.4987.

Shchemeleva, I.
    2022   "When the unconventional becomes convention: Epistemic stance
          in English as a lingua franca research articles", *Journal of English as
          a Lingua Franca* 11 (1), 265-279. https://doi.org/10.1515/jelf-2022-2075.

Swales, J. M. – C. B. Feak
    2012   *Academic Writing for Graduate Students: Essential Tasks and Skills*
          (3rd edn.). Ann Arbor, Michigan: University Press.
          https://doi.org/10.3998/mpub.2173936.

Vetchinnikova, S.
    2014   *Second Language Lexis and the Idiom Principle*. Dissertation. Helsinki:
          University of Helsinki.

Wu, X. – A. Mauranen – L. Lei
    2020   "Syntactic complexity in English as a lingua franca academic writing",
          Journal of English for Academic Purposes 43, 100798.
          https://doi.org/10.1016/j.jeap.2019.100798.

Address: Tereza Guziurová, Department of English Language Teaching, Faculty of Education, University of Ostrava, F. Šrámka 3, 709 00 Ostrava, Czech Republic. ORCID code: http://orcid.org/0000-0002-7739-8980.