Revista de Lenguas para Fines Específicos

Reformulation and its markers in unpublished research articles: Some evidence on the rhetorical patterns of written academic ELF

Silvia Murillo Ornat[1] ·

https://orcid.org/0000-0003-1369-3590

Universidad de Zaragoza

Departamento de Filología Inglesa y Alemana, C/ San Juan Bosco, 7 (3.ª planta) · Facultad de Filosofía y Letras · Spain

ABSTRACT

This paper intends to contribute to the description of written academic English as a Lingua Franca (ELF) from an endonormative perspective (Seidlhofer, 2011). Reformulation markers (*that is*, *that is to say*, *in other words*, *namely* and *i.e.*) fulfil an interactive metadiscourse function (Hyland, 2007) and have been considered indicators of certain rhetorical aspects of different languages, specifically, whether expansions, clarifications, adjustments, etc. are frequent or not (Cuenca, 2003). Here I examine the frequency, functions, and (non-) parenthetical uses of these markers in the components of a corpus of unedited ELF research papers (the SciELF corpus, University of Helsinki). The findings indicate that the frequency of reformulation markers varies in the different L1 groups, with high rates in the Romance languages. The results also point to other ELF-related trends such as simplification/ specialization of the use of one marker (*i.e.*), and discourse explicitation, closely associated to the functions specification and explanation. Other outcomes may be related to a global academic context (with different disciplinary areas), of which ELF forms part (and they would not be ELF-specific). In conclusion, formal written academic ELF seems to constitute an "endorhetorical" use of the language.

RESUMEN

En este artículo se pretende contribuir a la descripción del inglés como lengua franca (ILF) en el ámbito académico desde una perspectiva endonormativa (Seidlhofer, 2011). Los marcadores de reformulación (*that is*, *that is to say, in other words*, *namely* y *i.e.*) tienen una función metadiscursiva interactiva (Hyland, 2007) y han sido considerados indicadores de ciertos aspectos retóricos en diferentes idiomas, concretamente, de si son o no frecuentes las expansiones, aclaraciones, ajustes, etc. (Cuenca, 2003). En este trabajo examino la frecuencia, funciones y usos (no) parentéticos de estos marcadores en los

1   Corresponding author · Email: smurillo@unizar.es

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

27

componentes de un corpus de artículos de investigación escritos en ILF y no editados (el corpus SciELF, Universidad de Helsinki). Los resultados indican que la frecuencia de los marcadores de reformulación varía en los diferentes grupos de L1, con niveles altos en las lenguas romances. Los resultados también apuntan a otras tendencias relacionadas con el ILF, tales como la simplificación/ especialización del uso de un marcador (*i.e*) y la explicitación discursiva, asociada a las funciones de especificación y explicación. Otros resultados pueden relacionarse con un contexto académico global, con diferentes disciplinas científicas, del que el ILF forma parte (y no serían específicos del mismo). Como conclusión, se puede decir que el ILF académico y formal parece constituir un uso "endoretórico" de la lengua.

1. Introduction

In the academic world, it is of primary importance to publish research articles in English, and this is the case in most countries and in most disciplines. Research articles are also assessed in terms of their impact, which is generally associated with the number of citations they receive. In general, an article will potentially have more impact if it is accessible by a high number of readers, and the use of a common language no doubt plays an important role regarding such accessibility. Nevertheless, despite the wide use of L2 English in the academic sphere, L1 English has been considered the norm to be followed at the linguistic and rhetorical levels, however difficult that may have been for L2 English scholars (cf. Ingvarsdóttir & Arnbjörnsdóttir, 2013).

In the recent past, intercultural rhetorical studies in EAP have (progressively) taken several approaches, incorporating different sets of corpora. Comparisons were drawn initially between published research articles written in L1 English and in other L1 languages: English and Polish (Duszak, 1994; Golebiowski, 1998), English, French and Norwegian (Breivega et al., 2002; Dahl, 2003, 2004; Fløttum et al., 2006; Vold, 2006), English and Italian (Molino, 2010), and English and Spanish (Moreno, 1997, 1998, 2004; Mur-Dueñas, 2007). All these studies contrasting English and other languages revealed intercultural differences in the use of rhetorical and lexico-grammatical features between the research articles in English as used by native speakers and in the different native languages.

Furthermore, some studies concentrated on the comparison of discourse features between research articles written in L1 English, in another L1, and in L2 English (i.e., by scholars whose L1 language was the other language under analysis). These studies tried to establish whether there was transference of certain features from the writers' L1 to the L2 English texts, or whether an adaptation process had taken place. Transference processes were reported in English articles written by L1 Bulgarian academics, with regard to boosting and hedging devices (Vassileva, 2001), by L1 Danish speakers, in an analysis of evaluative features (Shaw, 2003), and by L1 Spanish speakers, concerning the use of epistemic modal verbs (Pérez-Llantada, 2010), first person pronouns (Lorés-Sanz, 2011), and engagement markers (Lafuente-Millán, 2014). On the other hand, Mur-Dueñas (2009) found adaptation processes in the use of logical markers by L1 Spanish academics writing in L2 English, and Murillo (2012) found both adaptation and transference

processes in the use of reformulation markers by writers with the same linguistic background, giving way to discursive hybridity (cf. Mauranen et al., 2010).

Research contrasting published papers written in L2 English by speakers of other languages with papers written by authors affiliated to Anglo-American institutions has also revealed interesting outcomes. Mur-Dueñas (2015) analysed evaluative it-clauses in Business Management research articles written in L1 English and in L2 English by L1 Spanish academics, and found differences in their frequency and lexicogrammatical realizations (specifically, in the choice of adjectives and in the use of modal verbs). Povolná (2016) also found differences in the use of discourse markers in L1 English and L2 English (by L1 Czech and L1 Slovak authors) in a corpus of Linguistics research papers. Edited and published L2 English materials seem to generate linguistic changes even when they have been edited and reviewed, so they may be considered instances of academic ELF writing, and such argument could lead us to reinterpret the research on transference processes reported above.

In what can be considered an intermediate stage, studies on the use of ELF in research articles have been made in a somewhat indirect way, focusing on the processes of correction/ edition of the papers, and they have yielded some promising preliminary results. Anderson (2010) analysed working papers on the area of Social Sciences, with respect to 'deviations' from the norm (as judged by expert raters), and found 'variability' in certain areas, specifically in function words like articles and prepositions, and also in the position of adverbs and in tense choice. Mur-Dueñas (2013) studied several text histories of papers of the same discipline, written by L1 Spanish scholars, which had undergone an editing process, and concluded that frequent corrections could be taken as "potential salient features of ELF" (2013, p. 334). These corrections included changes and additions at different levels: lexical (choice of items and phraseology), grammatical (articles, prepositions, pronouns, verb tenses, and word order) and discursive (breaking long sentences, use of discourse markers –mainly correction of contrastive connectors and addition of endophoric markers and sequencers–, and retrospective labels).

While these previous lines of study have provided insightful aspects regarding the use of ELF in research articles, differences with L1 English articles, deviations from the norm as pointed out by linguistic experts, or even so-called 'improvements', cannot stand alone to characterize the dynamic nature of written academic ELF. ELF is a vehicular language between users who do not share a first language, that is, who have different L1 language backgrounds or "similects" (Mauranen, 2012, 2017, 2018). As Mauranen explains, ELF is based on contact between users from different L1 language groups. In other words, "the languages involved are each in contact with English and it is these hybrids […] that are in turn in contact with each other" (2017, p. 739). ELF is thus a "hybrid of similects" (2012, p. 30). Yet, most of the previous studies of research articles have been restricted to the analysis of L2 English with one L1 language background, or two at most.

Academic ELF as such has been mostly explored by means of studies on oral encounters (or interactive written genres) where mutual understanding is more important than correctness (Mauranen, 2010, 2012). Thus, a salient feature in spoken academic ELF is the tendency to use

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

29

strategies that are emphatically explicit, such as metadiscourse and self-rephrasing (2010, pp. 14-18). Mauranen's studies show that metadiscourse or discourse reflexivity is fundamental to academic discourse, helping "increase clarity and explicitness among speakers from different linguistic and cultural backgrounds" (Mauranen et al., 2016: 46); self-rephrasing can likewise enhance explicitness by retaking what has already been said to make it clearer (Mauranen, 2010, p. 17). Other ELF-related trends are lexicogrammatical simplification (e.g. regularisation of morphology and increased use of the most common items) and approximation (e.g. approximate use of articles, prepositions, and phraseology) (Mauranen, 2010, 2012).

Written academic ELF is thus open ground for empirical research (Mauranen et al., 2016). The WrELFA corpus Project (Corpus of Written English as a Lingua Franca in Academic Settings), coordinated by A. Mauranen at the University of Helsinki, has started to yield some results (Carey, 2013), and the compilation of the SciELF corpus has opened up the possibility of carrying out systematic analyses of non-edited research articles (Rowley-Jolivet, 2017). Specifically, Carey's study (2013) suggested that the proportion of phraseological approximations (*in my point of view*, *in my view point*) was similar in speech and writing, and clearly ELF-specific. He also found that canonical expressions were more frequent than approximations, and that frequent standard expressions in native writing increased their frequency in ELF. Rowley-Jolivet (2017) has carried out a brief analysis of enabling verbs (*allow, enable, permit*) in the SciELF corpus, and has found non-canonical patterns in some of the similects included in the corpus, mainly among the Romance L1 writers.

The present paper should be understood as a contribution to the description of written academic ELF, from an endonormative perspective (as emphasized by Seidlhofer et al., 2006; Seidlhofer, 2011, among others). This study specifically seeks to explore the use of reformulation markers (Cuenca, 2003; Murillo, 2004; Hyland, 2007) in a corpus of unedited research papers, the SciELF corpus, arguing that these items may be taken as indicators of some discourse and rhetorical processes and tendencies that take place in ELF.

The outline of this paper is as follows. In section 2, I briefly review some relevant aspects of the literature on reformulation markers and I present my approach for classifying their functions (Murillo, 2012), using examples from the SciELF corpus. Section 3 describes the corpus in detail, highlighting its potential features as a resource for this type of study, and explains the methodological procedures followed. In section 4, the results are presented, including the general frequency of the reformulation markers, their types, functions, and (non-)parenthetical uses, taking into account where possible the L1 components and the broad disciplinary groups of the corpus. Finally, in section 5, I conclude with some final remarks concerning the extent to which the results could respond to ELF-related tendencies.

2. Reformulation markers

Studies on discourse markers in ELF have focused mainly on spoken communication (House, 2009; Baumgarten & House, 2010). The items under study, reformulation markers, include *that is*, *that is to say*, *in other words*, *namely* and *i.e.* (Chalker, 1996; del Saz, 2007), which are common

in academic prose (Biber et al., 1999, p. 884; Hyland, 2007; Murillo, 2012). They fulfill an interactive metadiscourse function: by using them, authors relate a text to its context by taking into account "their readers' needs, understandings, existing knowledge, intertextual experiences, and relative status" (Hyland, 2007, p. 284). They can thus be seen to draw up a map of the writers' assessments or perceptions about their readers.

Reformulation markers have also been characterized as indicators of the rhetorical organization of the different languages, specifically, whether expansions, clarifications, adjustments, etc. are frequent or not. The results of Cuenca's (2003) corpus study on reformulation markers in Linguistics articles in English, Spanish and Catalan indicate that the English writers use fewer reformulation markers than the Spanish in academic texts. This lower frequency in the English language can be associated with a writer-responsible/ formal-oriented culture, which favours linearity over digressions or explanations; Spanish would form part of a reader-responsible/ content-oriented culture, which allows for more digressions and where the inclusion of content is favoured over linearity. With a frequency of reformulation markers higher than English and lower than Spanish, Catalan would occupy an intermediate position between these two languages (cf. Kaplan, 1966; Hinds, 1987; Clyne, 1994, in Cuenca, 2003).

In Murillo (2004, 2012) reformulation markers are characterized as elements that help in the process of utterance interpretation (cf. Sperber & Wilson, 1986/ 1995). I provide a framework (Murillo, 2012) which integrates the typologies elaborated by previous research (mainly Gülich & Kotschi, 1983; Charolles & Coltier, 1986; Murat & Cartier-Bresson, 1987; and Fløttum, 1994), and includes the categories of identification, specification and explanation, in relation to the interpretation of explicit meaning; definition and denomination, in relation to conceptual knowledge; and conclusion and mathematical operation, in relation to implicit meaning. Thus, while all these categories are related to the use of metadiscourse and to rephrasing, that is, to discourse explicitness (cf. Mauranen, 2010), there are some functions that are primarily related to the explicit side of discourse. The categories operate cross-linguistically, as shown in some English-Spanish contrastive studies (Murillo, 2009, 2012, 2016a).

Regarding the different categories, an identification process is used to help the reader in reference assignment. In the following example from the SciELF corpus, the reformulation marker *that is* introduces the referent of the previous pronoun "us":

(1) In an earlier interview, Kirill explains that "us," **that is** "the Orthodox people," includes all his bap-
tised Orthodox compatriots who are not convinced atheists (Metropolitan Kirill 2002). (SSH57; Rus-
sian L1)

Specification is similar to identification, but it includes the presence of a cataphoric element, in the following case, "3 distinct time windows". This process serves as a discourse-organizing device:

(2) This is the first study to examine rates of bone loss across menopause in HIV women evaluating 3
distinct time windows, **namely** transitional period, early and late menopause. (Sci48; Italian L1)

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

31

In example (3), a previous expression is reformulated and clarified by means of an explanation:

(3) In addition to rupturing the logical hypotactic construction, by translating the hypotaxis into an adjunct, <FOREIGN> jovialmente </FOREIGN> forms a misleading picture in terms of the meaning provided by laughing. **That is**, it only expresses a person who is happy and friendly, and not a person who feels an exquisite pleasure for having given such a brilliant and dissenting opinion on moral edicts. (SSH47; Portuguese L1)

In a definition, the writer provides the necessary implicit contextual or encyclopedic information to understand a given term, making it explicit; in the following example, we find a definition of the word "agalmata":

(4) As Alcibiades states it himself, without grasping the meaning of his statement, Socrates' spiritual goods are only <FOREIGN> agalmata </FOREIGN>, **that is** images reflecting divine beauty without being identical to beauty itself. (SSH35; French L1)

Conversely, in (5) a denomination process the writer provides implicit conceptual information (how to name a concept):

(5)   [...]: two heliostats in the outer next row at both sides of the radial axis of the blocked heliostat, named "shoulder" blocking, and one more in the outer second row just in front (on the same radial axis) of the problem heliostat **i.e.**, "nose" blocking. (Sci66; Spanish L1)

In the following conclusion, the clause introduced by *in other words* constitutes an implication arising from the initial sequences[2]:

(6) [...] St. Augustine responded, "It is true that Christians pay religious honour to the memory of the martyrs, both to excite us to imitate them and to obtain a share in their merits, and the assistance of their prayers" (Contra Faustum Manichæum XX, 21). **In other words**, for a Christian believer saints are role models and helpers. (SSH57; Russian L1)

The last process, mathematical operation, is a particular type of conclusion in which a calculation is performed and made explicit:

(7) Thus, a total of six heliostats are checked for shadowing and three ones for blocking **i.e.**, a total of nine projections on the surface of the analysed heliostat. (Sci66; Spanish L1)

Finally, previous studies (Murillo, 2007, 2012) have dealt with reformulation markers in parenthetical sequences in English, which tend to be frequent. These sequences would be a way of inserting explicit content without disrupting the rhetorical linearity of the English language (cf. Cuenca, 2003). They may be signalled by parentheses, dashes, or commas, as in examples (8), (9) and (10) from the SciELF corpus:

---

[2]   In examples (1), (4) and (6) we find a reinterpretation of a previous discourse member which has been formulated by a different voice, by means of the use of direct or indirect reported speech. For an account of reformulation and polyphony, see Murillo (2016a). For the relationship between conclusive reformulation and polyphony, see Murillo (2016b).

(8) The five companies included in the sample are the largest European enterprises (**i.e.** in terms of market capitalisation) in the oil and gas sector. (SSH50; Romanian L1)

(9) Our focus is on how the shape of the status function -- **i.e.,** how social status is computed and eva-luated -- can affect the equilibrium outcome of the model, and in particular the relationship between inequality and wasteful conspicuous consumption. (Sci49; Italian L1)

(10) The main goal of this study was to evaluate 87Sr/86Sr ratio in different matrices, **namely** soils, branches, and grape juices, of an oenological food chain in order to develop a robust analytical strategy able to link the investigated food to its territory of origin. (Sci47; Italian L1)

3. Corpus and methodology

The SciELF corpus (2015) is a component of the WrELFA corpus (University of Helsinki); it consists of 150 unedited research papers by authors with ten different L1 backgrounds, of both hard science disciplines ('Sci') and soft science disciplines ('SSH'), with a total of 759,300 words. The fact that these papers have not been proofread professionally or edited by a native English speaker makes this corpus a valuable resource for the research of written academic ELF.

The corpus includes papers written by academics from a wide range of L1 backgrounds or similects: Chinese, Czech, Finnish, French, Italian, Portuguese, Romanian, Russian, Spanish, and Swedish (table 1)[3]. Similects have been defined as "L1-based group lects that derive from parallel cross-linguistic influence in individual speakers, identifiable as similar features in their second language repertoires" (Mauranen, 2014, p. 229). As Lorés-Sanz (2016) explains, written English used as an international language for academic communication in research articles may also be studied from this ELF perspective, because "language contact may involve face-to face interaction among groups of speakers using different languages, but it may also involve non-personal contact of persons with texts available in the written medium" (Braunmüller & House, 2009, cited in Lorés-Sanz, 2016, p. 57). Therefore, this L1-group component can provide new relevant data for the study of written academic ELF.

---

[3] The requirements for inclusion in the corpus were that "the author(s) should not have English as an L1 and the text should not have undergone professional proofreading services or language checking by an English native speaker" (*SCiELF Corpus Manual*, p. 8). The unedited final versions of the papers were collected by international partners in their respective L1 countries (Rowley-Jolivet, 2017, p. 6). The authors hold several academic roles, the distribution of which is as follows: junior staff (418,366 words or 55%); senior staff (172,075 words or 23%); research student (107,998 words or 14%); masters student (19,745 words or 3%); unknown (41,116 words or 5%) (*SCiELF Corpus Manual*, p. 3). Although the level of expertise is, of course, a relevant issue (cf. Tribble, 2017), I believe that the fact that authors are using L2 English and the influence of the L1 background may also be considered worthy of study, and the SciELF corpus offers a representative sample of researchers from different countries whose objective is to publish in L2 English.

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

33

	Articles	**Number of words**
Chinese	21	84,807
Czech	22	109,173
Finnish	25	123,153
French	16	91,186
Italian	11	58,685
Portuguese	12	56,625
Romanian	4	25,197
Russian	13	71,376
Spanish	13	79,038
Swedish	13	60,060
Total	150	759,300

Table 1. L1 categories in the SciELF corpus.

As specified on the corpus website, the Sci (sciences) category comprises 326,463 words, and the SSH (social sciences and humanities) texts amount to 432,837 words (table 2). 79% of the Sci contents are drawn from the natural sciences and 18% from medicine. The contents in SSH are drawn as follows: 45% from social sciences, 34% from humanities, and 21% from behavioural sciences. The compilers' purpose was to achieve a balanced sample of papers between the sciences and the social sciences and humanities, and, as a result, the design of this corpus can allow us to provide a better characterization of ELF paper writing, taking (broad) disciplinary variation into account (cf. previous research by Hyland, 2001, 2002; Harwood, 2005; Lafuente-Millán, 2012, among others). As the compilers indicate, "[t]his categorisation is by no means unproblematic, but it tends to work best for the big picture, and a more fine-grained division would not be justified for a corpus of this size" (*SciELF Corpus Manual*, pp. 8-9).

	Articles	**Number of words**
Sci	78	326,463
SSH	72	432,837
Total	150	759,300

Table 2. Distribution of the broad binary categories in the SciELF corpus.

In order to extract the occurrences of reformulation markers in the corpus, the items *that is*, *that is to say*, *in other words*, *namely* and *i.e.* (Chalker, 1996; del Saz, 2007) were searched automatically in the articles of the corpus and the texts were also checked manually. Additional searches were made for sequences such as "to put it", "putting it", "differently", "simply", "way", "words", and "terms", in order to find less grammaticalized markers such as *to put it/putting it differently*, *to say the same thing differently*, *putting it simply*, *to put/say the same thing a different way, in plain words, in simpler terms* (cf. Cuenca, 2003; del Saz, 2007), and also possible new phraseological variations. The different fragments including reformulation markers were extracted from the texts, with sufficient context (as in the examples above), and compiled separately.

The analysis considered the ten L1 background subcorpora and the Sci and SSH subcorpora, as well as the variables specific markers (types), functions, and parenthetical uses. The categories for the function variable were identification, specification, explanation, definition, denomination, conclusion, and mathematical operation, as explained above (examples 1-7). For parenthetical

uses, the three possible cases illustrated in examples (8-10) were included. Due to their more limited length, the analysis of the L1 background subcorpora focused on the general frequency of the markers (and their types). The Sci and SSH subcorpora were large enough to yield relevant results, with respect to broad disciplinary differences. A data set was generated with R-Commander (version 3.5.1), which was used in order to cross the variables and display the data in tables. The data were standardized to a common basis (occurrences per 10,000 words), to ensure the comparability of the results, and chi-square calculations were performed to check the statistical significance of the different comparisons[4].

4. Results and discussion

The search for the grammaticalized markers *that is*, *that is to say*, *in other words*, *namely* and *i.e.* yielded 588 tokens, or 7.74 per ten thousand words (table 3). However, not all the similects contributed to the same extent to the rates of use of reformulation markers in the corpus. Indeed, a clear continuum can be seen from Chinese (2.48 reformulation markers per 10,000 words), then Swedish (3.66), Finnish (4.63), Russian (6.02), Portuguese (7.59), French (7.57), Spanish (9.11), Czech (10.08), and Romanian (11.91), to Italian (20.62). The higher frequency in the L1 groups corresponding to the Romance languages is quite revealing, as it is in line with previous research on L1s (cf. Cuenca, 2003; Murillo, 2007, 2016a). The L1s are ranked following the ascending order in the table. There are statistically significant differences between the total results for the different subcorpora, that is, between the total number of cases of reformulation markers in each similect (p-value: <0.001).

	That is	That is to say	In other words	Namely	I.e.	Total/ Per 10,000 words
Chinese	5	1	1	6	8	21 / 2.48
Swedish	1	0	0	1	20	22 / 3.66
Finnish	9	1	7	7	33	57 / 4.63
Russian	3	0	7	3	30	43 / 6.02
Portuguese	15	0	5	3	20	43 / 7.59
French	6	5	21	4	33	69 / 7.57
Spanish	3	11	8	17	33	72 / 9.11
Czech	9	4	14	6	77	110 / 10.08
Romanian	1	0	0	5	24	30 / 11.91
Italian	19	5	3	21	73	121 / 20.62
Total	71 (12.1%)	27 (4.6%)	66 (11.2%)	73 (12.4%)	351 (59.7%)	588 / 7.74
Chi-square (L1 totals): 202.495; degrees of freedom: 9; p-value: <0.001						

Table 3. Frequency of reformulation markers in the SciELF similects.

---

[4] These operations were made using the online statistical calculator provided by Preacher (2001). In order to mitigate the effects of the different number of words in the components of the corpus (L1 and Sci/SSH groups), size figures were taken into account in the contingency tables for the calculations. If the p-value was <0.05, the threshold level usually set in Linguistics, the results were considered statistically significant. Very low p-values are represented as <0.001 in the tables.

These L1 groups reflect different frequencies of marked reformulations, that is, the authors go back explicitly to previous discourse fragments in different degrees, with significant variation. These language groups present in-group features and, at the same time, they are contributing to ELF, which is a hybrid by definition.

With reference to the frequency of the specific markers, *i.e.* was the most recurrent one in all the similects, the rest of the markers presenting a much more limited occurrence. There were some other markers that might reflect the use of the markers of the L1s due to formal correspondences. For instance, *es decir* is very frequent in general Spanish (Murillo, 2007, 2016a) and *that is to say*, which is a formal equivalent, was also frequent in the L1 Spanish subcorpus, unlike in L1 general English (Murillo, 2009), or academic L1 English corpora (Murillo, 2012).

Regarding the two disciplinary subcorpora Sci (containing hard-science research articles) and SSH (Social Sciences and Humanities, with soft-science research articles), they show a similar overall frequency of reformulation markers (table 4): 7.78 tokens per 10,000 words in the Sci subcorpus, and 7.72 in the SSH subcorpus.

	That is	That is to say	In other words	Namely	I.e.	Total/ Per 10,000 words
Sci	15	2	11	28	198	254 / 7.78
SSH	56	25	55	45	153	334 / 7.72
Total	71	27	66	73	351	588 / 7.74
	(12.1%)	(4.6%)	(11.2%)	(12.4%)	(59.7%)	

Chi-square (totals): Yates' chi-square: 0.003; degrees of freedom: 1; Yates' p-value: 0.9563199[5]
Chi-square (types): chi-square: 72.886; degrees of freedom: 5; p-value: <0.001

Table 4. Frequency of reformulation markers in the SciELF corpus (Sci and SSH).

As seen in tables 3 and 4, the item *i.e.* is the most frequent reformulation marker in the whole corpus (59.7% of the tokens). This frequent use of *i.e.* would be an instance of specialization or simplification, an ELF-related trend in which a frequent item in L1 English (Murillo, 2007) becomes even more frequent in ELF (cf. Mauranen, 2012; Carey, 2013).

In table 4, this marker predominates with a very high frequency in the Sci subcorpus, while in the SSH subcorpus a certain frequency of other markers like *that is*, *in other words*, and *namely* was found. In fact, the chi-square test confirmed statistically significant differences regarding the frequency of the different types (p-value: <0.001). These results are consistent with those of Hyland (2007, p. 273), and can be attributed to disciplinary differences, as the SSH papers tend to be more interpretative, allowing for more variation in the markers used.

Only 13 occurrences of less grammaticalized markers were found in the corpus, and most of them correspond to the SSH subcorpus (10 cases). They include mainly the adverb *simply* or the adjective *simple*: *to put it simply, simply put, put simply, simply,* and *in simple terms.* Other instances include the adverb *differently*: *to put it differently, put differently,* and *said differently*. The markers *to put it another way* and *another way to put it* were also documented. No phraseological sub-

---

[5]   As this calculation corresponds to a 2x2 contingency table (degrees of freedom: 1), Yates' correction was applied.

standard variants were documented in the corpus. However, as Carey points out, "the tendency of academic ELF users to deploy approximated chunks should not be overstated" (2013, p. 226). These less grammaticalized markers introduce processes of conclusion (example 11) or explanation (example 12).

(11) It has long been suggested that the vocal apparatus and auditory circuitry are actively involved in language comprehension including silent reading (Baddeley, Eldridge, and V. Lewis 1981). In line with this theory, recent studies have shown that listening to speech activates the recipient's tongue muscles (Watkins, Strafella, and Paus 2003; Fadiga et al. 2002),[10] that verbal auditory imagery activates the auditory cortex (for a review, see Hubbard 2010), and crucially, that silent narrative reading activates the temporal voice areas associated with speech perception (Yao, Belin, and Scheepers 2011). **Simply put**, silent reading entails "voices" in one's brain. (SSH11; Czech L1)

(12) The potentiodynamic polarisation curves were obtained and a comparison between the pitting potential listed in Table 4 indicates that the values didn't show significant differences. **In simple terms**, the oxides films of the steels submitted to the saline environment were just erupted above 1.35 V. (Sci54; Portuguese L1)

If we consider the different functions of the markers, as we can see in table 5, explanation (in 44.9% of the cases) and specification (in 30.9%) are the most frequent cases in the corpus, and no statistically significant differences were found in this comparison between the Sci and the SSH subcorpora (p-value: >0.05)[6]. In other words, the markers can be associated with functions primarily related to explicit content in both subcorpora, following the ELF tendency towards discourse explicitation strategies (Mauranen, 2012).

	Sci	SSH	Total	
Identification	6	8	14	(2.4%)
Specification	92	90	182	(30.9%)
Explanation	115	149	264	(44.9%)
Definition	16	28	44	(7.5%)
Denomination	4	7	11	(1.9%)
Conclusion/ Mat. Operat.	21	52	73	(12.4%)
Chi-square: 11.289; degrees of freedom: 6; p-value: 0.07984479				

Table 5. Functions of reformulation markers in the SciELF corpus (Sci and SSH).

In line with previous results (tables 3, 4 and 5), the most frequent combination in the whole SciELF corpus is *i.e.* performing the functions of either explanation or specification, with 176 and 113 cases, respectively (table 6).

[6] The functions conclusion and mathematical operation have been subsumed in order to calculate the chi-square, as it is necessary to have frequencies higher than 0 in all the cells of the contingency table, and these two functions can be considered to be close enough.

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

37

	That is	That is to say	In other words	Namely	I.e.
Identification	4	0	0	2	8
Specification	8	0	0	61	113
Explanation	42	14	23	9	176
Definition	9	2	3	1	29
Denomination	0	1	0	0	10
Conclusion	8	10	40	0	14
Mat. Operation	0	0	0	0	1

Table 6. Functions associated to reformulation markers in the SciELF corpus.

Example (13) illustrates *i.e.* introducing a (parenthetical) explanation, and example (14) shows this marker in a specification.

(13) Our focus is on how the shape of the status function -- **i.e.**, how social status is computed and evaluated -- can affect the equilibrium outcome of the model [...] (Sci49; Italian L1)

(14) Verbal imagery and conscious conceptual thought share a fundament, **i.e.** the verbal format. (SSH11; Czech L1)

Namely can be associated to specifications (example 2 above), with 61 cases in the corpus (table 6), *that is* tends to appear in explanations, with 42 cases (example 3), and *in other words* in conclusions, with 40 cases (example 6).

Regarding deviant cases, only one was found in the corpus (example 15). It corresponds to the following use of *that is to say*, in which the author is explaining research data from a table. The logical sequence would have been to present the data first, and then to draw the conclusion from these data, but here it is done in reverse order, which renders an odd sequence. In other words, this would be a pragmatically deviant use, in relation to the function of the marker.

(15) From Table 1, we can see clearly that learners in experimental classes are more autonomous in previewing the text. **That is to say** most students (81.1%) read the text at least twice before they visit the website courseware while only 23.7% students preview the text twice. (SSH04; Chinese L1)

Finally, table 7 displays the frequency of parenthetical and non-parenthetical uses of reformulation markers in the Sci and SSH subcorpora. Although in general non-parenthetical uses are more frequent than parenthetical uses, the latter are more frequent in the Sci corpus (3.49 parenthetical uses per 10,000 words) than in the SSH subcorpus (2.24). The distribution in the two disciplinary subcorpora yields statistically significant differences (p-value: <0.001).

	Parenthetical uses	Non parenthetical uses
Sci	114 / 3.49	140 / 4.29
SSH	97 / 2.24	237 / 5.48
Chi-square: 15.762; degrees of freedom: 2; p-value: <0.001		

Table 7. Parenthetical uses of reformulation markers in the SciELF corpus (Sci and SSH).

These findings, like those of table 4 which contrast the frequency of the different types of markers in the Sci and SSH subcorpora, can also be accounted for by disciplinary differences (cf. Hyland, 2007). The Sci papers tend to be concise and usually involve more succinct parenthetical reformulations.

5. Conclusion

This research focused on the characterization of written academic ELF in unedited papers, by analysing the use of reformulation markers as indicators of discourse processes and rhetorical patterns. The results would indicate that the frequency of these markers in the L1 subcorpora seems to reflect the hybrid nature of EFL, and that the use of such items follows certain tendencies accounted for in other (mainly oral) ELF (academic) genres or communicative situations. Finally, other outcomes point to trends that may be related to the global academic context of which ELF forms part (and would thus not be ELF-specific).

The results seem to depict a "melting pot" with regard to reformulation rhetorical patterns in ELF research articles, which would be reflecting the hybrid nature of EFL, that is, a common language conformed by the contribution of a range of similects, which appear to be transferring their own L1 rhetorical patterns. The similects display a different frequency in the use of reformulation markers, and they all contribute to and form part of ELF: writers whose first language is other than English are communicating in English, and these writers are somehow interacting, or having contact, with one another. In Mauranen's words "we can characterize ELF as what could perhaps be called 'second-order language contact': a contact between hybrids" (2014, p. 229).

Furthermore, the use of these markers may be seen to follow certain tendencies identified in other academic ELF communicative contexts, that is, part of the results would point to general processes that take place in ELF. There seems to be a common specialization of the use of one particular marker (*i.e.*), and two particular functions (explanation and specification) which are related to the interpretation of explicit meaning. The trend towards specialization/ simplification may be a cost-effective use of the language on the part of the L2 English academics, who in many cases may have become familiar with these markers and their uses by reading other papers on their subject fields. *I.e.* is indeed a very simple marker, with no phraseological complexities (Mauranen, 2012, pp. 30-31; Mauranen et al., 2015), and it is perhaps treated as an "island of reliability" (Granger, 1998, cited in Carey, 2013, p. 226), which would account for its high frequency. As explained above, the function explanation is used to rephrase a previous discourse member in order to make it more understandable; specification clarifies a previous referent, and it is very often associated to discourse organization by means of cataphoric elements and enumerations. This tendency towards explicitation has been pointed out as a salient feature of academic spoken ELF, as a way of enhancing clarity and understanding (Mauranen, 2012, 2014, p. 243).

On the other hand, the results of the contrast between the disciplinary components of the SciELF corpus, the Sci and SSH subcorpora, may be related to a global academic context, of which ELF forms part. There are statistically significant differences regarding the frequency of the different

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

39

types or specific markers and the incidence of the parenthetical sequences versus non-parenthetical ones, which is relevant in terms of promoting rhetorical linearity (or not). In the science papers, the marker *i.e.* is the most frequent one, and, while this is also the case in the other subcorpus, there is a higher frequency of some markers such as *that is*, *in other words* and *namely* in the social sciences and humanities papers, and less grammaticalized markers are also used. There are more parenthetical cases in the science subcorpus. These results concerning the specific types of markers and their parenthetical uses are consistent with those reported by Hyland, and can be attributed to disciplinary variation, or to "different knowledge making practices" (2007, p. 284). Science papers tend to be of a more concise nature, whereas social sciences and humanities papers tend to be more interpretative, allowing for more variation.

The ultimate aim of this research was to know about L2 English academic writing, that is, to contribute to the description of written academic ELF. There seem to be differences in the L1 groups despite the writers' common exposure to ELF, and also some collective ELF-specific trends. Thus, written academic ELF can be studied in its own right, from an "endonormative perspective" (Seidlhofer, 2011). This approach, however, does not mean we should talk about absence of norms, or deviations from grammar rules, particularly in published academic ELF (Rowley-Jolivet, 2017, p. 10). We have in fact seen in this paper that there were almost no linguistic deviations in the use of reformulation markers[7], so the label "endonormative perspective" should perhaps be delimited and, in that case, it would have to include an endorhetorical component. Formal ELF academic writing seems to constitute an endorhetorical use of the language, with its own rhetorical patterns. These patterns would be motivated by different factors such as the influence of the L1s and the writers' need to communicate effectively in ELF.

Regarding future research, as the present study was limited to reformulation markers, it would be interesting to extend this analysis to other ways of introducing reformulations and also to unmarked reformulations. It would also be of interest to do similar analyses of other discourse markers or other discourse features. Further, contrastive studies with academic L1 English corpora should be carried out. I believe this latter line of research would not undermine the perspective adopted in this paper; on the contrary, it would complement and strengthen it.

In addition, as the cases in the corpus analysed correspond to standard correct English, the extent to which such uses are respected by reviewers and editors (cf. Anderson, 2010) could be assessed, i.e. by carrying out a contrastive study between edited and non-edited ELF papers. Mur-Dueñas (2013) reported corrections of discourse markers whose use was deviant from standard English in revised research papers, mainly addition of connectives and changes in cases involving *on the contrary* and *on the other hand,* which are problematic for L1 Spanish speakers; however, no quantitative data are available, and no qualitative or quantitative data in relation to reformulation markers or reformulations, in any case. There are indeed a lot of aspects that remain to be described and categorized with regard to the use of written academic ELF.

---

[7] However, deviations from the norm have been found even in papers published in prestigious journals in some specific disciplines (Rozycky & Johnson, 2013).

40

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

About the author

Silvia Murillo Ornat is a senior lecturer in the Department of English and German Studies at the Faculty of Arts of the Universidad de Zaragoza (Spain). Her research includes contrastive studies (English-Spanish) on general corpora and on academic corpora (covering L1 and L2 English), focusing on the use discourse markers, particularly reformulation markers.

Acknowledgements

References

Anderson, L. (2010). Standards of acceptability in English as an Academic Lingua Franca: Evidence from a corpus of peer-reviewed working papers by international scholars. In R. Cagliero & J. Jenkins (Eds.), *Discourses, Communities, and Global Englishes* (pp. 115-143). Bern: Peter Lang.

Baumgarten, N. & House, J. (2010). *I think* and *I don't know* in English as lingua franca and native English discourse. *Journal of Pragmatics, 42*(5), 1184-1200.

Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Breivega, K., Dahl, T. & Fløttum, K. (2002). Traces of self and others in research articles. A comparative pilot study of English, French and Norwegian research articles in medicine, economics and linguistics. *International Journal of Applied Linguistics*, *12*(2), 218-239.

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.

Chalker, S. (1996). *Collins Cobuild English guides 9: Linking words*. London: Harper Collins Publishers.

Charolles, M. & Coltier, D. (1986). Le contrôle de la compréhension dans une activité rédactionnelle: Élements pour l'analyse des reformulations paraphrastiques. *Pratiques*, *49*, 51-66.

Cuenca, M.J. (2003). Two ways to reformulate: A contrastive analysis of reformulation markers. *Journal of Pragmatics*, *35*(7), 1069-1093.

Dahl, T. (2003). Metadiscourse in research articles. In K. Fløttum & F. Rastier (Eds.), *Academic Discourse: Multidisciplinary Approaches* (pp. 120-138). Oslo: Novus Press.

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

41

Dahl, T. (2004). Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics*, *36*(10), 1807-1825.

Del Saz, M. (2007). *English discourse markers of reformulation*. Bern: Peter Lang.

Duszak, A. (1994). Academic discourse and intellectual styles. *Journal of Pragmatics*, *21*(3), 291-313.

Fløttum, K. (1994). À propos de *c'est-à-dire* et ses correspondants norvégiens. *Cahiers de Linguistique Française*, *15*, 109-130.

Fløttum, K., Dahl, T. & Kinn, T. (2006). *Academic Voices – Across Languages and Disciplines*. Amsterdam/ Philadelphia: John Benjamins.

Golebiowski, Z. (1998). Rhetorical Approaches to Scientific Writing: An English-Polish Contrastive Study. *Text*, *18*(1), 67-102.

Gülich, E. & Kotschi, T. (1983). Les marqueurs de reformulation paraphrastique. *Cahiers de Linguistique Française*, *5*, 305-351.

Harwood, N. (2005). 'We Do Not Seem to Have a Theory… The Theory I Present Here Attempts to Fill This Gap': Inclusive and Exclusive Pronouns in Academic Writing. *Applied Linguistics, 26*(3), 343-375.

House, J. (2009). Subjectivity in English as Lingua Franca discourse: the case of *you know*. *Intercultural Pragmatics, 6*(2), 171-193.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes, 20*, 207-226.

Hyland, K. (2002). Options of identity in academic writing. *ELT Journal*, *56*, 351-358.

Hyland, K. (2007). Applying a gloss. Exemplifying and reformulating in academic discourse. *Applied Linguistics*, *28*(2), 266-285.

Ingvarsdóttir, H. & Arnbjörnsdóttir, B. (2013). ELF and academic writing: a perspective from the Expanding Circle. *Journal of English as a Lingua Franca*, *2*(1), 123-145.

Lafuente Millán, E. (2012). A contrastive study of generic integrity in the use of attitudinal evaluation in RAs written for different audiences. *Brno Studies in English*, *38*(2), 79-92.

Lafuente-Millán, E. (2014). Reader engagement across cultures, languages and contexts of publication in business research articles. *International Journal of Applied Linguistics*, *24*(2), 201-223.

Lorés-Sanz, R. (2011). The study of authorial voice: Using a Spanish–English corpus to explore linguistic transference. *Corpora*, *6*(1), 1-24.

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.

Mauranen, A. (2010). Features of English as a Lingua Franca in academia. *Helsinki English Studies*, *6*, 6-28.

Mauranen, A. (2012). *Exploring ELF: Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press.

Mauranen, A. (2014). Lingua franca discourse: shaped by complexity. In J. Flowerdew (Ed.), *Discourse in Context: Contemporary Applied Linguistics*. Vol. 3 (pp. 225-245). London; New York: Bloomsbury.

Mauranen, A. (2017). Second-Order Language Contact: English as an Academic Lingua Franca. In M. Filppula, J. Klemola & D. Sharma (Eds.), *The Oxford Handbook of World Englishes* (pp. 735-753). Oxford: Oxford University Press.

42

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

Mauranen, A. (2018). Conceptualising ELF. In J. Jenkins, W. Baker & M. Dewey (Eds.), *The Routledge Handbook of English as a Lingua Franca* (pp. 7-24). London: Routledge.

Mauranen, A., Carey, R. & Ranta, E. (2015). New answers to familiar questions: English as a lingua franca. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 401-417)*.* Cambridge: Cambridge University Press.

Mauranen, A., Hynninen, N. & Ranta, E. (2016). "English as the academic lingua franca". In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 44-55). New York: Routledge.

Mauranen, A., Pérez-Llantada, C. & Swales, J.M. (2010). Academic Englishes: A standardized knowledge? In A. Kirkpatrick (Ed.), *The Routledge Handbook of World Englishes* (pp. 634-652). London: Routledge.

Molino, A. (2010). Personal and impersonal authorial references: A contrastive study of English and Italian linguistics research articles. *Journal of English for Academic Purposes*, *9*, 86-101.

Moreno, A.I. (1997). Genre constraints across languages: Causal metatext in Spanish and English RAs. *English for Specific Purposes*, *16*, 161-179.

Moreno, A.I. (1998). The explicit signalling of premise-conclusion sequences in research articles: A contrastive framework. *Text*, *18*(4), 545-585.

Moreno, A.I. (2004). Retrospective labelling in premise-conclusion metatext: An English-Spanish contrastive study of research articles on business and economics. *Journal of English for Academic Purposes*, *3*, 321-339.

Murat, M. & Cartier-Bresson, B. (1987). *C'est-à-dire* ou la reprise interpretative. *Langue Française*, *73*, *La reformulation du sens dans le discours*, 5-15.

Mur-Dueñas, P. (2007). 'I/we focus on...': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, *6*, 143-162.

Mur-Dueñas, P. (2009). Logical markers in L1 (Spanish and English) and L2 (English) Business research articles. *English Text Construction*, *2*(2), 246-264.

Mur-Dueñas, P. (2013). Spanish scholars' research article publishing process in English medium journals: English used as a lingua franca?. *Journal of English as a Lingua Franca*, *2*(2), 315-340.

Mur-Dueñas, P. (2015). Looking into ELF variants: A study of evaluative it-clauses in research articles. *ESP Today*, *3*(2), 160-179.

Murillo, S. (2004). A relevance reassessment of reformulation markers. *Journal of Pragmatics*, *36*(11), 2059-2068.

Murillo, S. (2007). *A Contribution to the Pragmalinguistic Contrastive Study of Explicatory Reformulative Discourse Markers in Contemporary Journalistic Written English and Spanish*. Unpublished PhD thesis. Universidad de Zaragoza.

Murillo, S. (2009). Los marcadores de reformulación explicativa en español y en inglés: Estudio contrastivo de 'o sea' y sus traducciones 'that is (to say)' e 'in other words'. In M.P. Garcés Gómez (Ed.), *La reformulación del discurso en español en comparación con otras lenguas (catalán, francés italiano, inglés, alemán e islandés)* (pp. 137-161). Madrid: BOE/ Universidad Carlos III de Madrid.

Murillo, S. (2012). The use of reformulation markers in Business Management research articles: An intercultural analysis. *International Journal of Corpus Linguistics*, *17*(1), 69-90.

Murillo, S. (2016a). Reformulation markers and polyphony: a contrastive English-Spanish analysis. *Languages in Contrast*, *16*(1), 1-30.

Murillo, S. (2016b). Sobre la reformulación y sus marcadores. *Cuadernos AISPI*, *8*, 237-258.

Silvia Murillo Ornat · Reformulation and its markers in unpublished research articles:
Some evidence on the rhetorical patterns of written academic ELF

43

Pérez-Llantada, C. (2010). The 'dialectics of change' as a facet of globalisation: Epistemic modality in academic writing. In M.F. Ruiz-Garrido, J.C. Palmer-Silveira & I. Fortanet-Gómez (Eds.), *English for Professional and Academic Purposes* (pp. 25-42)*. Amsterdam: Rodopi.

Povolná, R. (2016). A cross cultural analysis of conjuncts as indicators of the interaction and negotiation of meaning in research articles. *Topics in Linguistics, 17*(1), 45-63.

Preacher, K.J. (2001). Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from <http://quantpsy.org>.

Rowley-Jolivet, E. (2017). English as a Lingua Franca in research articles: the SciELF corpus. *ASp, 71*, 145-158.

Rozycky, W. & Johnson, N.H. (2013). Non-canonical grammar in Best Paper award winners in engineering. *English for Specific Purposes*, *32*, 157-169.

Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford: Oxford University Press.

Seidlhofer, B., Breiteneder, A. & Pitzl, M.-L. (2006). English as a Lingua Franca in Europe: Challenges for Applied Linguistics. *Annual Review of Applied Linguistics*, *26*, 3-34.

Shaw, P. (2003). Evaluation and promotion across languages. *Journal of English for Academic Purposes*, *2*, 343-357.

Sperber, D. & Wilson, D. (1986/ 1995). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

Tribble, C. (2017). ELFA vs. Genre: A new paradigm war in EAP writing instruction? *Journal of English for Academic Purposes*, *25*, 30-44.

Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes*, *20*, 83-102.

Vold, E.T. (2006). Epistemic modality markers in research articles: A cross-linguistic and crossdisciplinary study. *International Journal of Applied Linguistics*, *16*(1), 61-87.

Primary data

SciELF 2015. *The SciELF Corpus*. Director: Anna Mauranen. Compilation manager: Ray Carey. <http://www.helsinki.fi/elfa/scielf.html>.

Carey, R. (2015). *SciELF Corpus Manual. ELFA Project*. University of Helsinki.