A lexical bundle analysis of art-related crowdfunding projects

Alberto A. Vela-Rodrigo
University of Zaragoza (Spain)
vela@unizar.es

Abstract

In the context of Open Science, crowdfunding projects are gaining increasing attention. They are becoming an alternative way of funding research and an opportunity for researchers across the disciplines to share and disseminate their work widely while engaging their potential backers. The aim of this paper is to analyze crowdfunding projects online by examining language-in-use at the level of phraseology, understanding the latter from a lexical bundle approach. Using corpus analytical approaches, the study findings show that there is a recurrence of lexical bundles conveying deontic meanings used to persuade the potential backers that the project and the research methods proposed for carrying it out are reliable and therefore trustable. Lexical bundles expressing gratitude and politeness are also recurrent, not unexpected considering that crowdfunding proposals aim to prompt the audience's participation through donation. The findings further reveal how distinct discourse style and language features especially frequent in the conversational register realise the main communicative purpose of the genre, namely, to build credibility and trust in research with a view to persuasively enticing the backers' audiences to donate money.

Keywords: crowdfunding, digital genres, discourse functions, lexical bundles, Open Science

Resumen

Un análisis de paquetes léxicos de proyectos de micromecenazgo en el ámbito artístico

En el contexto de la Ciencia Abierta, los proyectos de crowdfunding (micromecenazgo) están acaparando cada vez mayor atención. Se están convirtiendo en una forma alternativa de financiación para la investigación y en una oportunidad para que los investigadores de todas las disciplinas compartan

y difundan ampliamente su trabajo al tiempo que involucran a sus posibles patrocinadores. El objetivo de este artículo es analizar proyectos de crowdfunding online examinando el lenguaje en uso a nivel fraseológico desde el punto de vista de los paquetes léxicos. Usando enfoques analíticos de corpus, los hallazgos del estudio muestran que hay una recurrencia de paquetes léxicos que transmiten significados deónticos utilizados para persuadir a los patrocinadores potenciales de que el proyecto y los métodos de investigación propuestos para llevarlo a cabo son fiables y, por lo tanto, financiables. Los paquetes léxicos que expresan gratitud y cortesía también son recurrentes, lo que no es inesperado si se tiene en cuenta que las propuestas de crowdfunding tienen como objetivo promover la participación de la audiencia a través de la donación. Los hallazgos revelan, además, cómo el estilo discursivo distintivo y las características del lenguaje, especialmente frecuentes en el registro conversacional, cumplen el principal propósito comunicativo del género, esto es, generar credibilidad y confianza en la investigación con el fin de atraer persuasivamente a las audiencias de los patrocinadores para que donen dinero.

Palabras clave: micromecenazgo, géneros digitales, funciones discursivas, paquetes léxicos, Ciencia Abierta

1. Introduction

Digital technologies have been developed that give researchers the opportunity to use new channels for sharing and disseminating their work and for interacting with broad publics in various ways. In the context of Open Science and participatory science, Web 2.0 functionalities allow researchers and broad publics to construct new knowledge, share it and disseminate it widely (Luzón & Pérez-Llantada, 2022). Among all the emerging digital genres, the crowdfunding proposal stands out for allowing especially direct and easy communication between researchers and society, providing a space in which parties can both collaborate and support research in different ways (Mehlenbacher, 2017, 2019; Handke & Dalla Chiesa, 2022). This genre responds to new social exigences such as the need to disseminate scientific knowledge beyond expert audiences and support and enhance the broad publics' understanding of science (Herring, 2013; Miller & Kelly, 2017). Moreover, it allows researchers to communicate their outcomes to society through different platforms and make them credible and accessible, by this means supporting Open Science (Bonney et al., 2009; Follett & Strezov, 2015). From a genre theoretical standpoint, the crowdfunding project proposal has been defined as a genre type that falls under the

category of hybrid genre (Herring, 2013; Mehlenbacher, 2019). As Mehlenbacher argues, it adopts the rhetorical and linguistic conventions of the research grant proposal (i.e., complex grammar, elaborated phrasal structures and explicitness in meaning relations) and also exhibits distinct features of other online genres of science popularization, such as blogs and citizen science projects (for example, colloquial features associated with conversation and the use of stance markers to express judgments) (Caliendo, 2012, p. 101).

Different platforms have become popular to launch crowdfunding projects, such as *Kickstarter*, *Precipita*, *Artistshare*, *GoFundMe*, *Experiment*, *Consano* or *Davincicrowd*, among others. In the discipline of Arts, the number of art museums and galleries and people who claim to be interested in art have dramatically increased over the last decades, especially with the advent of mass exhibitions (McCarthy et al., 2005; Kabassi, 2017). Students in Museum Studies programmes and Art students are developing new strategies to fund their projects out of the traditional national grant programmes, being one of these strategies the use of crowdfunding platforms. As Handke and Dalla Chiesa (2022, p. 278) put it, crowdfunding is rising as a "flexible tool that changes restrictions for participants in the cultural sector, which entails opportunities to develop new empirical insights". It is important, therefore, to characterise the deployment of language resources that understand how language in these texts serves to inform about research and persuade their backers to give funds for scientific research.

The existing literature has mainly examined the rhetorical information organization conventions and main discourse features of crowdfunding proposals but, to the best of my knowledge, has not yet complemented these analyses with the analysis of language at a phrasal level. Another research gap is that these analyses have not been done for disciplinary fields other than STEM and health sciences. Investigating the grammar and syntax at the level of the phrase in this genre is vital to better understand the functionality of linguistic features that contribute to grammatical complexity in discourse vs grammatical elaboration and construct either a grammatically compressed or grammatically elaborated style for rhetorical effects. This can help more clearly discern how these proposals are written to 'sell' their scientific or humanistic research results. Previous studies of crowdfunding proposals for engineering, business, medicine or biomedical research, among the study of other science popularization genres, have come to light in the past years (e.g., Paulus & Roberts, 2018; Pérez-Llantada, 2021). These studies have provided

empirical evidence to claim that linguistic features associated with formal academic prose coexist with features especially common in conversation in this genre. On the other hand, the scarce existing research on the discipline of arts (Bondi, 2009; Lazzeretti, 2016) has focused fundamentally on the discourse of the manuals of art history, exhibition catalogues or even in digital environments, e.g., art blogs and museum websites. The few studies on crowdfunding proposals in the discipline of art have shown the existence of promotional and informational discourse types within this genre (Donelli et. al., 2022), or have been approached from an economic perspective (Gürler & Çağlar, 2021; Handke & Dalla Chiesa, 2022), but they have not explored language at the level of the phrase.

The current study focuses on lexical bundles in art-related crowdfunding projects. Defined as "the most frequently recurring sequences of words" (Biber & Barbieri, 2007, p. 264) that "show statistical tendency to co-occur" (for example, *a better understanding of the, it should be noted that, thank you very much for*) (Biber et al., 1999, pp. 989–990), lexical bundles make up one of the main categories of recurrent word combinations. Lexical bundles serve important discourse functions in both spoken and written language, such as guiding readers to particular interpretations of the information, organizing details along a text (for example, framing expressions or linking previous and coming details) or providing identification of new data (Cortes, 2004; Biber & Barbieri, 2007). Lexical bundles have been investigated in different written or spoken genres of academic discourse (e.g., Biber et al., 2004; Cortes, 2004; Hyland, 2008) and in digital genres such as blogs and crowdfunding projects online (Barbieri, 2018; Pérez-Llantada, 2021). To contribute to this line of research, the purpose of this paper is to explore bundle usage in crowdfunding proposals online in the field of Art, a disciplinary field that has been little investigated in comparison with other disciplines in the STEM fields. The research questions that guided the study were the following:

> RQ1. What are the recurring lexical bundles of crowdfunding projects?
>
> RQ2. What are the main structures and discourse functions of these bundles?
>
> RQ3. More broadly, how do these recurring lexical bundles help fulfil the communicative purpose of these projects?

2. Literature review

All languages are formulaic or repetitive, in the sense that sequences of words or strings of linguistic items that are more or less fixed in form recur (Hunston & Francis, 2000, p. 7). In addition to the general language, there are also different languages typical of discursive communities. Demonstrating membership of a specific discourse community requires mastery of its language rules and conventions that only its members know.

Several terms have been used in the literature to define and describe the formulaicity of language: 'lexical phrases', 'formulas', 'routines', 'fixed expressions', 'pre-fabricated patterns', or 'lexical bundles', among others. These terms all foreground the fact that they are recurring sequences of words that seem to be stored in memory and ready to be retrieved when use requires it (Wray & Perkins, 2000). Among all of them, this study focuses on lexical bundles. A large number of studies on traditional academic genres have dealt with lexical bundles related to discourse style and register variation. Some have compared the characteristics of different registers, such as textbooks and classroom discourse (Biber et al., 2004; Biber & Barbieri, 2007), or the behaviour of native and non-native English speakers (De Cock, 2000; Chen & Baker, 2010; Ädel & Erman, 2012; Pérez-Llantada, 2014) and student vs. expert academic writers (Cortes, 2004; Hyland, 2008; Chen & Baker, 2010).

Taking the case of textbooks and book reviews, Römer (2010) reports that lexical bundles are not always mono-functional but they can have different meanings. Furthermore, word bundles or word units serve to express four different functions in a text: evaluation, structure, content and discourse, so their examination is essential for identifying a text type's phraseological profile. Hyland (2008) and Pérez-Llantada (2014) have also relied on lexical bundle analysis to examine the formulaicity of the research article from the perspective of expert writing for academic and research publication purposes. The lexical bundle approach is important for its pedagogical implications, since formulaicity in academic writing is not part of the native writer's innate language ability and is thus far from being a linguistic universal skill. Hyland concludes that the forms, structures and functions of lexical bundles in four different disciplines (electrical engineering, biology, business studies and applied linguistics) offer an important means of differentiating written texts, noticing that these recurring bundles occur and behave in different ways, hence indicating disciplinary variations. For example, the

recurring lexical bundles used in hard sciences articles indicate that the emphasis is put on precision to ensure the accurate understanding of proceedings and results, while social sciences articles are rich in stance bundles, allowing writers to present information tentatively rather than as complete commitments (Hyland, 2008; Dontcheva-Navratilova, 2012). These studies conclude that a high frequency of occurrence and a high degree of systematicity at the level of the phrase is a key feature of successful academic writing, enabling communication among distinct members of the academic discourse community. Academic genres show a high level of formality in the use of the language, being very explicit, and lexical bundles have an important role in facilitating academic text construction and reception across academic genres.

Particularly salient is the recent work by Cortes (2022), who has adopted a new perspective in lexical bundle analysis, moving the focus from the traditional 4-word bundle analysis to the functioning of 3-word lexical bundles. In her analysis of 4.5 million words in a corpus of research articles in different disciplines, Cortes addressed the problem of overlapping bundles, that is to say, shorter expressions that may be subsumed in longer ones, varying their structure or function. Cortes realized that among the 120 3-word lexical bundles she extracted, only 35% were autonomous 3-word bundles, while 65% of the expressions formed part of 4-word or longer bundles. Independent 3-word bundles included expressions such as *in order to, a set of,* and *the amount of*, to mention just a few.

Turning to digital genres, features of discourse (Herring, 2013; Barbieri, 2018), lexicon (e.g., lexis in TED talks) (Scotto di Carlo, 2014) and pragmatics (Carter-Thomas & Rowley-Jolivet, 2017) have been studied. These studies conclude that some linguistic features (i.e., the use of more nominal than verbal patterns) in online communication are similar to those of written and spoken registers. Corpus linguistics research has also claimed that there is an increasing colloquialization of written texts in digital media (Biber & Gray, 2016) while constructing credibility online through the use of the linguistic features associated with these texts (e.g., deictics, person pronouns 'I-you-we', use of stance markers, among others). Few studies have examined phraseology in digital genres. For example, using a corpus of crowdfunding projects, Pérez-Llantada (2021) concludes that, at the clause/phrase level, nominal phrases and verbal phrases are especially frequent in these projects. This points to a hybrid discursive style, with typical features of academic writing and features of conversational language.

The present work gives continuity to this line of research in digital genres, drawing attention to crowdfunding proposals and to one of the least investigated disciplinary fields in the crowdfunding genre, that of art.

3. Methods

As also done in previous studies, this study takes a "radical corpus-driven approach" (Biber, 2009, p. 281) and makes no prior assumptions on the use of the language. As the approach is entirely inductive, the linguistic constructs emerged directly from the analysis of 40 crowdfunding project proposals that were specifically collected for this work from the platform Experiment.com, one of the largest platforms supporting science crowdfunding. This platform also supports other disciplines but always from a scientific perspective in relation to their methods. It is worth highlighting that crowdfunding projects can be difficult to analyse because they may be subject to intrageneric changes. These are fragmented texts of a hybrid nature, which present their information in this website divided into different sections or tabs (Overview, Methods, Lab Notes and Discussion). Figure 1 shows an example of the homepage of a project in Experiment.

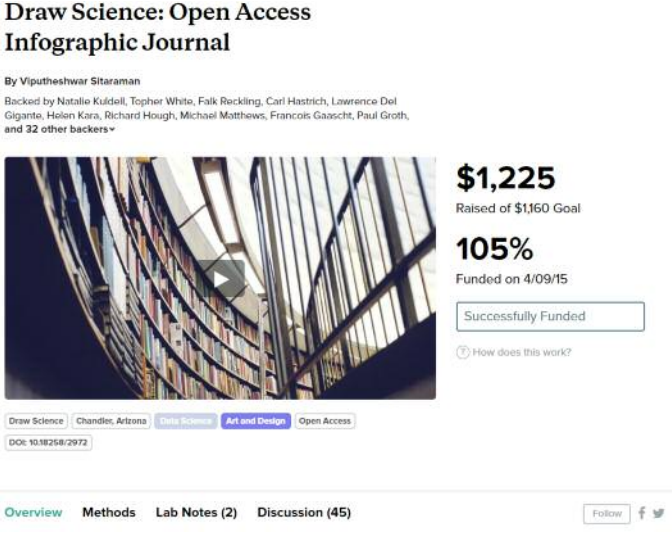

Figure 1. Homepage of the project 'Draw Science: Open Access Infographic Journal', by Viputheswar Sitaraman (doi: 10.18258/2972; https://experiment.com/projects/draw-science-open-access-infographic-journal?s=discover, last accessed on 7 January 2023.

The overview tab allows researchers to present a summary or synopsis of their projects, their main goals, benefits for science and society, data about the Team, the number of dollars pledge and the time required. This is a tab firmly guided by the template offered by the website, with researchers answering a series of established questions (*What is the context of this research? / What is the significance of this project? / What are the goals of the project?*). The Methods and Lab Notes tabs are not mandatory; they offer a space for researchers to give updates on the project to their potential backers, with additional information, photographs and comments. Finally, the Discussion tab is an interactive page that takes the form of a discussion forum with which researchers and their backers interact. On the other hand, this platform offers guidelines by way of recommendation on content, style and register when starting a project (https://experiment.com/start) and also a section for frequently asked questions (FAQ) that informs and can also guide the writing process (https://experiment.com/faq). Even so, how researchers organize and upload information to the web is a task that depends exclusively on them and their writing and communication abilities.

As for the nature of the corpus analyzed, it is representative of two related disciplines, 'art and design' and 'anthropology', available on the Experiment.com website. As it is well known, art is closely linked to anthropology (Zitzewitz & Ciotti, 2022) and together they made up a total of 40 projects. The corpus totalled 105,006 words (see Annex 1 for a complete description of the number of words per project). The content of all blocks containing information was extracted and "cleaned" before being saved as .txt documents for subsequent analysis, eliminating manually all those words and expressions present in titles or not belonging to natural language.

Following the frequency-first principle (Biber et al., 1999; Biber, 2009) as a key criterion for the extraction of lexical bundles, a frequency list of 3-word lexical bundles was created to ensure the independence of observations. By this means, the issue of overlapping bundles (i.e., the fact that some 3-word bundles are subsumed in longer bundles) has been addressed.

For this study, intended to be an exploratory analysis of crowdfunding proposals, I used *Antconc 4.2.2* (Anthony, 2023) to retrieve the especially frequent lexical bundles. A cut-off point frequency (>10) and range (>5) were established to work with a manageable number of bundles. Following Cortes (2015) and Bestgen (2019) I provide raw counts since, as these

authors argue, normalization of frequencies does not work well with data-driven formulaic expressions. I carried out a manual analysis of all resulting bundles, extracting those 3-word bundles which were embedded in longer bundles (n-bundles) and according also to previous studies (Chen & Baker, 2010; Cortes & Lake, 2023), assuring, this manner, that the functions they played in the texts were correct and they were not distorted by their apparent belonging to longer constructions.

To further explore how the resulting highly frequent bundles frame semantic meanings, I drew upon previous functional taxonomies of bundles in academic written genres (Biber et al., 2004; Cortes, 2004; Hyland, 2008). Following these taxonomies, the bundles were classified according to i) types of structure (phrasal/clausal), ii) types of words forming the bundle (content/function words) and discourse functions that they perform (referential, text-organizing and stance bundles). Referential expressions were those that provided identification of new information or make reference to physical or abstract entities, including time (*the end of; in the past*) place (*at the site*), quantifiers (*an important part of*) and descriptive expressions (*this project is; the results of*). Text organizing expressions were those reflecting relationships between prior and coming discourse. This category comprised lexical bundles that establish comparison/contrast (*in addition to; some of the*). Finally, stance and interactional expressions were identified. These aimed at guiding readers to particular interpretations, conveyed by epistemic bundles such as *be able to* or *will keep you* and attitudinal ones (*look forward to; we hope to*).

Although previous studies of lexical bundle usage in academic prose refer to these three categories, for the present study special conversational bundles (Lorenzo, 2011) were also extracted. The main reason for extracting them is that these bundles perform a set of functions for showing politeness (*thanks so much*), inquiring and reporting (*learn more about; if you want*). In the present study it was hypothesised that because the Discussion section of these projects is built upon an interactivity tool with which the researchers and their publics can interact and exchange comments, these bundles were likely to be present.

These different discourse functions were identified manually using close reading of the texts and context-sensitive analysis, following Chen and Baker (2010) and Pérez-Llantada (2021). This last category was included on the assumption that this is a digital genre on the Internet aimed at non-specialist audiences with a very important space for the exchange of gratitude from

both sides, in which the fact of showing support, not only economic, plays a fundamental role.

4. Results

4.1. Frequencies and structures of lexical bundles

A total of 84 3-word recurring lexical bundles were retrieved (Table 1). The most highly frequent 3-word combinations (*as well as, in order to, one of the*) are also especially frequent in a similar analysis carried out by Cortes (2022) in a 4.5-million-word corpus of research articles in Biology, Economics, Business, Engineering and History. This is an important finding, since it suggests that crowdfunding proposals exhibit word combinations similar to those used in academic written genres, which supports the idea of the crowdfunding proposal genre as a hybrid genre (Herring, 2013).

Rank	Freq	Range	3-word bundle	Rank	Freq	Range	3-word bundle
1	57	21	as well as	43	13	9	you for your
2	46	20	one of the	44	12	9	at the same
3	32	19	be able to	45	12	9	forward to seeing
4	31	17	in order to	46	12	9	in addition to
5	28	15	some of the	47	12	9	in the past
6	26	11	i want to	48	12	11	in this project
7	26	14	thank you for	49	12	10	interested in the
8	26	19	this project is	50	12	10	is one of
9	25	18	this project will	51	12	10	learn more about
10	23	12	a lot of	52	12	12	that can be
11	23	9	the th century	53	12	9	the end of
12	22	14	part of the	54	12	10	to do this
13	21	6	of the site	55	12	6	we have been
14	20	15	i am a	56	12	8	we will be
15	20	12	i will be	57	12	6	will help us
16	19	12	looking forward to	58	11	8	allow us to
17	18	13	be used to	59	11	7	as part of
18	18	14	the results of	60	11	7	due to the
19	18	15	this is a	61	11	10	good luck with
20	17	11	a variety of	62	11	9	i have a
21	17	11	for your support	63	11	6	proud of you
22	16	10	goal is to	64	11	6	the analysis of
23	16	8	in the world	65	11	5	the site is
24	16	7	we hope to	66	11	9	the study of
25	16	9	will be used	67	11	5	we will keep
26	15	5	at the site	68	10	5	and we will
27	15	10	for this project	69	10	6	been able to
28	15	9	i have been	70	10	9	it can be
29	15	7	in the field	71	10	7	it has been
30	15	11	look forward to	72	10	7	of the most
31	15	11	the use of	73	10	10	project will be
32	15	9	would like to	74	10	10	so much for
33	14	10	it is a	75	10	8	thank you so
34	14	11	project is to	76	10	8	thank you to
35	14	9	thanks so much	77	10	9	the goal of
36	14	7	to be able	78	10	6	the process of
37	14	11	we need to	79	10	5	the site of
38	14	6	will keep you	80	10	8	to be a
39	13	8	best of luck	81	10	7	to participate in
40	13	7	i plan to	82	10	9	to understand the
41	13	8	the amount of	83	10	8	we want to
42	13	9	to create a	84	10	8	you so much

Table 1. Frequency and range of 3-word lexical bundles.

4.1.1. Bundle overlapping

From the list of 3-word bundles in Table 1 11 bundles (13.09%) also made up longer bundles (n-bundles), as illustrated in Table 2.

3-word bundle (subsumed)	N-word bundle
thank you for so much for for your support	thank you so much for your support
looking forward to forward to seeing	looking forward to seeing
look forward to forward to seeing	look forward to seeing
this project will project will be	this project will be
we will keep will keep you	we will keep you (informed)

Table 2. Instances of longer bundles made up from 3-word bundles.

The result of this simplification, that took on board instances of overlapping, was a revised list of 5 longer bundles. For instance, 3-word bundles *thank you for, so much for* and *for your support*, give rise to *thank you so much for your support*. Example 1[1] allows to see this in context, with researchers of a project about the use of photogrammetry in old colonies in East Florida thanking the support of backers in the Lab Notes section.

(1) **Thanks so much for your support**. There is so much to be done and so little that is new or up-to-date.

This type of interaction and the use of this bundle is common in other sections too, such as the Discussion tab, a space reserved for communication between researchers and backers. Most of these longer bundles found in the corpus are present in this type of sections in which conversational and colloquial features of language seem to be the norm.

Comments of support and comments of greeting can happen in writer-backer direction, as in example 2,

(2) This approach will indeed bring to light the biases that are a structural part of research. **We will keep you informed**!

Or in the backer-writer direction, as in example 3, a project about an archaeological excavation and drawing in Byzantine Athens, in which a backer writes the following comment in the Discussion section,

(3) **Looking forward to seeing** this project fully funded and the results of your work!

4.1.2. Structures of lexical bundles

The analysis of the structure of the bundles showed that clausal bundles (45 bundles, 53.57%) were slightly more frequent than phrasal bundles (39 bundles, 46.43%) (Table 3). Clausal bundles are especially frequent in the conversational register. These clausal bundles were mainly verb phrase fragments (33.34%) (e.g., *project will be; learn more about; this project is*) and verb phrase fragments with finite clause fragments (14.28%) and, to a much lesser extent, dependent clause fragments (5.95%) (e.g., *to do this; to understand the; to participate in*). Phrasal types were mostly formed by noun phrase fragments (22.61%) (e.g., *the end of, the results of*), prepositional phrases (13.09%) (e.g., *at the site; in the past*), adjective phrases (4.76%) (e.g., *proud of you*), adverb phrase fragments (3.57%) (e.g., *so much for; as well as*) and conjunction phrase fragments (e.g., *in order to; and we will*) (2.4%).

As also shown in Table 3, more than three quarters of all clausal bundles (88.82%) corresponds to verb phrase fragments (62.16%) and its subcategory of verb phrase fragments with finite clause fragments (26.66%), and the rest (11.18%) to dependent clause fragments. As for phrasal types, 48.71% were noun phrases, 28.2% prepositional phrases, 10.25% adjective phrase fragments, 7.69% adverb phrase fragments and 5.15% conjunction phrase fragments.

		SUBTYPES	
STRUCTURAL TYPES	CLAUSAL	Dependent clause fragment	to do this; to create a; to be a; to participate in; to understand the
		Verb phrase fragment	this project is; this project will; i am a; i will be; this is a; will be used; i have been; it is a; interested in the; that can be; be able to; thank you for; thanks so much; is one of; project will be; will keep you; learn more about; we have been; we will be; will help us; due to the; i have a; the site is; we will keep; it can be; it has been; thank you so; thank you to;
		Verb phrase fragment with finite clause fragment	i want to; be used to; we hope to; we need to; looking forward to; look forward to; would like to; project is to; i plan to; forward to seeing; allow us to; we want to
	PHRASAL	Adjective phrase fragment	proud of you; been able to; to be able; best of luck
		Prepositional phrase fragment	of the site; in the world; at the site; in the field; you for your; in addition to; in the past; in this project; of the most; for your support; as part of
		Noun phrase fragment	a lot of; the th century; part of the; the results of; a variety of; goal is to; for this project; the use of; the end of; the amount of; the end of; good luck with; the analysis of; the study of; the goal of; the process of; the site of; one of the; some of the
		Conjunction phrase fragment	in order to; and we will;
		Adverb phrase fragment	as well as; so much for; you so much

Table 3. Structures of 3-word lexical bundles in the corpus.

In Biber's study (2009) of formulaicity in academic prose phrasal bundles are more frequent than clausal bundles, with 74% of academic prose lexical bundles consisting of noun phrase expressions (e.g., *the nature of the*) or prepositional phrases (e.g., *as a result of the*) contributing to the sense of coherence in a text (Biber, 2009). In the present study the findings showed that clausal types were slightly more frequent (53.57%) than noun phrases, with verb phrase fragments and verb phrase fragments with embedded non-finite clause fragments amounting together to almost 50% of all analysed lexical bundles, which suggests that the language used in these proposals does not only resemble the language of academic written genres. In other words, the slightly higher presence of clausal fragments strongly suggests that, although the audience is being informed of scientific content (since roughly 46% of all bundles are phrasal), the language deployed in the proposals also resembles the language of conversation. These results reinforce the hybrid nature of this genre and are consistent with previous studies (Pérez-Llantada, 2021).

Hence, the fact that the percentage of clausal bundles is almost similar to that of phrasal bundles suggests that the crowdfunding proposals combine features especially frequent in academic prose and features especially frequent in conversation, being the style of these projects, both grammatically elaborated and grammatically compressed, that is, less "noun-centric" (Swales, 2008, p. v). The syntax of the academic written register tends to "employ embedded phrases rather than fuller dependent clauses" (Biber & Gray, 2010, p. 7). Noun phrase fragments with prepositional phrase post-modifiers (e.g., *the end of; the presence of*) and prepositional phrases post-modified by other prepositional phrases (e.g., *in the context of; at the beginning of*) usually reflects the grammatical compression typical of academic writing (Pérez-Llantada, 2014), a grammatical feature that is not salient in the analysed projects. This is not an unexpected finding if we consider that these proposals do not only target a specialised community, but also the broad public, that is, diversified audiences.

By contrast, it is worth highlighting the relatively frequent use of 'to' infinitive clausal-fragments (26.66%), which seems to indicate that 'to'-infinitive non-finite clauses are an important syntactic feature in these proposals. As shown in example 4, extracted from a project about the Palaeolithic cultural artifacts in Jordan, 'to'-infinitive clauses make clear to the readers the aims of the investigation and expected results.

(4) In order to begin testing this hypothesis, **we need to** excavate more. In particular, **we need to** obtain larger, more representative samples

Making explicit the purpose of the project may facilitate understanding of it and eventually trigger donation. This is a grammatical structure also recurrent in other science popularization genres, such as didactic reportages, news in online journals or citizen science projects (Pérez-Llantada, 2021).

4.2. Discourse functions

4.2.1. Overall findings

The 3-word bundles identified in the corpus performed several discourse functions. Table 4 shows the distribution of functional categories, following Biber et al.'s (1999) taxonomy.

Functional category	Functional subcategory	3-word bundles
Referential bundles	Time	at the same; in the past; the th century; the end of; the process of; forward to seeing
	Place	in this project; of the site; in the world; at the site; in the field; as part of; the site of
	Descriptive	for this project; the use of; to do this; i am a; this project is; this is a; it is a; project is to; the results of; to create a; to be a; to participate in; to understand the; will be used; that can be; i have been; we have been; i have a; the site is; it has been; the analysis of; the study of; the goal of; goal is to
	Quantifiers	a lot of; part of the; a variety of; of the most; some of the; the amount of; is one of; one of the; you so much
Text-organizers	Comparison / Contrast	as well as; due to the; in order to; in addition to
Stance and interactional expressions	Epistemic / Deontic	we need to; be used to; to be able; be able to; it can be; project will be; we will be; this project will; i will be; will help us; would like to; allow us to; been able to; and we will; will keep you; we will keep
	Attitudinal stance	good luck with; interested in the; looking forward to; look forward to; we hope to; i plan to; proud of you; best of luck
Special conversational functions	Politeness	for your support; thank you for; so much for; you for your; thanks so much; thank you to; thank you so
	Simple inquiry	i want to; learn more about; we want to

Table 4. Distribution of functional categories.

Figure 2 summarises the salience of each functional category, which indicates that referential bundles, and to a lesser extent, stance lexical bundles are particularly important in these proposals. The fact that more than half of the bundles in the texts are referential is significant. These bundles are especially associated with the rhetorical section Overview, in which researchers must make it very clear why they are researching and what

their purposes are. In this section, all this information must be very clear so that potential backers understand the project and decide to donate their money. On the other hand, the context sensitive analysis indicates that the presence of special conversational bundles in the corpus builds a type of language that is especially (but not exclusively) associated with the Discussion section. In this tab, researchers and the public interact, establishing a dialogue.

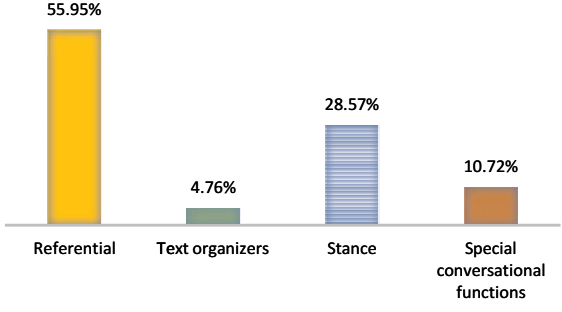

Figure 2. Distribution of functional categories (%).

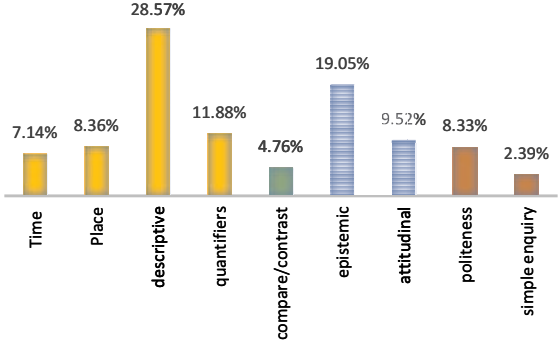

Figure 3. Distribution of functional subcategories (%).

Among referential bundles (55.95%), the most representative subcategory is that of descriptive bundles (e.g., *this project is; the use of; the site is*), amounting to 28.57% of all 3-word bundles (Figure 3). It is also worth noting that the very structure of the Experiment.com website, with a template with numerous sections (short modular texts) and headings (e.g., *What is the context of this research?; What is the significance of this project?* etc.) guides (and also constrains) the space for writing each project proposal, and therefore the

distribution of recurring lexical bundles may differ across the text blocks of the platform. Therefore, the materiality of the medium accounts for the well-established organization of the information into short modular texts, which might explain the relatively rare presence of text organizers (4.76%). This may be due to the lack of necessity of being redundant directing readers around the text or specifying limitations (Hyland, 2008).

Regarding stance bundles, epistemic/deontic (e.g., *been able to; in order to; will help us*) are particularly salient (19.05%), followed by attitudinal expressions (9.52%) (e.g., *we hope to; look forward to*) (Figure 3). Special conversational bundles represented 10.72% of all the analyzed lexical bundles, a presence significantly lower than those of referential, stance and interactional expressions. Other referential subcategories described by Hyland (2008) such as framing expressions (e.g., *with respect to the; in the basis of*) and inferential bundles (e.g., *it was found that*) were very rare in these texts. A possible explanation is that the organization of the hypertext into modular texts (text blocks/pages/subpages) does not require any framing expressions or transitions to move from one text block in a page/subpage to another.

4.2.2. Functions of referential bundles

Referential bundles are the most frequent ones, amounting to 55.95%, and are those that help the most to build the informative discourse, providing identification of new information and determining "the way of looking at things" (Cortes, 2004, p. 401). The very high frequency of referential expressions also suggests that the researchers aim at putting the emphasis on precision, to ensure the accurate understanding of procedures and results. Among them, descriptive bundles (e.g., *for this project; the use of, to do this*) are the most common (28.57%), being very important to accomplish the social/rhetorical action that the genre enacts. They serve to inform about the main data of the projects, that is to say, setting the procedures to be taken and informing about results that must be both economically and scientifically credible (example 5):

(5) When we compare **the results of** this study to our previous work testing people with little art experience, this will tell us […].

Example 6 provides historical data of an archaeological site in Macedonia, situating the potential backer in a historical context and helping to understand the importance of the excavation carried out in this project.

(6) **The site is** famous due to an event that occurred in 354 BC. Philip of Macedon, the father of Alexander the Great, was laying siege to the town.

The abundance of descriptive bundles helps to situate the reader in the context of the project, providing objective data and an idea of scientific rigour and seriousness. Since this genre aims to reach a diversified audience, informing and providing data is essential for the project to be understood both by experts and non-experts, therefore potential backers can measure its importance in order to donate funding.

4.2.3. Functions of stance expressions

Focusing particular attention on stance bundles now, those containing deontic expressions were the most frequent in its category (Figure 3) (e.g., *will be used; will allow us*) together with epistemic expressions (e.g., *would like to*) (19.05%). These bundles concern "the ways writers explicitly intrude into the discourse to convey epistemic and affective judgements, evaluations and degrees of commitment to what they say" (Hyland, 2008, p. 18). At a semantic level, epistemic modals are related to a tentative language, being very frequent in academic prose, while deontic modals are associated with an assertive language, typical of conversation (Biber et al., 1999). The presence of both types of bundles in the texts analysed indicates that these proposals are built upon a discourse that exhibits conversational features of the language to communicate scientific knowledge to diversified readers. In turn, the presence of lexical bundles with assertive functions also implies a clear persuasion strategy on the part of the researchers. For example, the bundles containing the deontic modal *will* can be considered a rhetorical strategy to build credibility and trust, as shown in example 7. In this example, the text explains how the monitoring of the impact of modern construction will be a consequence of discovering city ruins, a fact that will become true if they receive the funds for studying those old excavations.

(7) Not only has this (the location of older excavations) given us a more complete understanding of the city than ever before, it **will allow us** to establish a base line for monitoring the impact of modern construction.

Furthermore, stance bundles create a sense of security, efficiency and control on the part of the researcher, which can be transmitted through the

explanation of methodological issues or any other related aspect of the project development in an impersonal manner. In example 8, from a project dealing with the study of the ingredients and methods of the food of sailors in 17th century, in the Overview text, and more specifically in section 'What are the goals of the project?', stance bundles conveying deontic meanings perform rhetorically. They help to persuade the backer that his/her money is going to be used with appropriate scientific rigour.

(8) […] Biostatistics **will be used** to analyze the findings and find trends in the microbiological and nutritional values. This information will then be translated to what it would likely mean for the sailors on 17th-century ships who actually ate these items.

In example 8, the text draws on the lexical bundle containing a deontic verb phrase fragment (*will be used to*) to express that researchers will create their project in order to get to know the nutritional and microbiological data of the diet for a sailor in that time. The function of the deontic unit here is that of persuading the public to donate funds, otherwise the expected anthropological results will not be able to come to light. Therefore, the scientific community and society in general would get a benefit of this potential funding and the future enriching debate it will generate.

On the other hand, attitudinal expressions (e.g., *look forward to; proud of you; good luck with*) have an evaluative function for showing delight, desire or compromise. This function is located especially in the Discussion tab but also in Lab Notes tab (in which researchers share updates about their projects) and is carried out mainly by backers, as in example 9. This is so because backers sometimes happen to be part of the researcher's network (Experiment, 2014) or very committed with the project for personal interests and they show their feelings about it as a way to support morally. Moreover, attitudinal expressions indicate the presence of conversational elements typical of spoken discourse and informal interaction (first person pronoun 'I', verbs of emotion such as *feel, look forward, love*, contracted forms such as *can't, don't, won't*, or repetitions),

(9) Hi all, **Thank you for** supporting this project! We are currently still in the volunteer recruitment and data collection phase. We've reached out to local schools and community members through neighbourhood pop-up markets. I **look forward to** sharing more updates regarding results and outreach soon!

Stance expressions are present in different tabs and sections of Experiment.com as they are also characteristic of other genres such as the discussion forum or the microblog, which seems to corroborate the hybrid nature of the crowdfunding proposal genre. For example, in the Fanteca Project about rehabilitation through art among users of opiates in certain areas and communities of New York (Curtis et al., 2017), some followers of the project exclaim in the Discussion tab, *Congratulations, team! I* ***look forward to*** *seeing this great project get off the ground!* The same can be generalized to 100% of the projects analysed, in which appreciative comments are always shown by the community of backers in both sections, Lab Notes and Discussion.

4.2.4. Functions of special conversational bundles

These bundles are presumed to recur especially in the Discussion Tab, an area reserved for communication between backers and researchers. Normally, this communication occurs in backer-writer direction (but not only, see example 9), with expressions of politeness (e.g., *thanks so much; thank you for*), which denotes a certain familiarity and desire to support not only economically but also morally. This may be due to the fact that "the researcher's professional and personal networks play a large role in getting the project off the ground" (Experiment, 2014). Therefore, we could imply that many of the backers who post comments on the Discussion forum are relatives, friends or acquaintances of the researchers. Having a discussion tab offers recognition and importance to audiences that directly appeals to their emotions but also a space for the recognition of researchers' work. For example, in the discussion tab of all projects it is representative to find gratitude comments by the backers such as example 10,

(10) **Thank you for** this and everything else you do for this community. May many others follow your legacy and lend a helping hand.

The presence of these lexical bundles (especially in the Discussion section) indicates that these backers are sharing their emotions, establishing a direct contact with the researchers and telling them that the project is important to them, and that they perceive a social contribution that benefits them, either personally or as human beings within a community. From this it follows that persuasion in this section is higher than in other rhetorical sections of these projects. In turn, they express expectations of continuity of the project, providing the researcher with the idea of a second project and showing the

persistence of their support if this happens. In this way, through these recurring lexical bundles, in addition to interacting and being able to respond to their backers, researchers can distinguish the degree of follow-up and acceptance of their project and the emotional impact it generates.

5. Discussion

The purpose of this study has been to explore how crowdfunding projects online are constructed at the level of phraseology. For this purpose, a selection of crowdfunding artistic projects was analysed to understand how phraseology, a key feature enabling effective communication (Hyland, 2008; Pérez-Llantada, 2014), contributes to achieving the rhetorical/communicative goals of the genre, to inform about academic topics and to persuade the audience to support the project with donations (Mehlenbacher, 2019; Pérez-Llantada, 2021). The phraseological analysis focused on their main structures and discourse functions.

In response to RQ1, 'What are the recurring bundles of the selected crowdfunding projects?', the findings of the present study have shown that more than a half of all lexical bundles are clausal bundles, strongly suggesting that the discourse style of these texts is not grammatically compressed or, to put it differently, it is less "noun-centric" than academic writing. The study findings also suggest that the primary discourse function of highly frequent bundles formed by verbal phrase fragments is to convey information or researchers' procedures at the same time that they build convincing arguments. By this means, they helped to convey professionalism, build credibility and trust in research and persuade the backers to fund a particular project, reaching the communicative purposes of the genre.

As shown in previous studies, lexical bundles in academic writing are predominantly phrasal (Biber & Gray, 2010; Ädel & Erman, 2012; Hyland & Jiang, 2015). Findings here suggest a combination of little syntactic elaboration and scarce formality in some tabs (Lab Notes and Discussion) and a more elaborated discourse in others (Overview) built upon recurring VP structures. Therefore, the crowdfunding genre seems not to rely exclusively on phraseological features especially frequent in academic writing to achieve its communicative goals. However, these projects also display some features associated with the formal register that characterises academic

prose built upon NP structures performing referential functions in discourse. For example, when the researchers wish to objectively show data and their collection criteria, or when the budget planning and the research methodology are presented using descriptive lexical bundles. All this shows the importance of the research project and an image of seriousness and execution capacity, in the same way that the traditional grant proposal does in order to obtain funds from funding agencies (Biber & Gray, 2010; Salazar, 2014). Thus, both features (grammar compression and the presence of linguistic features associated with informal prose) seem to support the view of this genre as 'hybrid'. This is also the case of other digital genres that support public communication of science and public engagement in science such as TED Talks, blogs and citizen science projects (Luzón, 2013; Scotto di Carlo, 2014; Luzón & Pérez-Llantada, 2022).

Regarding RQ2: 'What are the main structures and discourse functions of these patterns?', this study showed that these proposals rely heavily on descriptive bundles for helping potential backers to understand a project and how beneficial their support can be, and therefore appealing directly to their value as entities with the might to make it successful. In science popularization genres such as didactic reportages and research group websites, informing about scientific topics to broad publics is vital, so they can interpret and form a personal opinion about the research that is communicated. Descriptive lexical bundles build a discourse full of details that enables the potential backers to get a clear idea of the project's topic, scope and impact. The fact of offering a detailed account of the project's objective data and explaining its interest allows the project launchers to fulfil their communicative intentions. This finding is consistent with previous studies of participatory genres (Mehlenbacher, 2017; Paulus & Roberts, 2018; Pérez-Llantada, 2021, 2023). It is worth recalling that non-finite fragments ('to'-infinitive clausal bundles) were especially frequent and expressed the relevance of the research and are associated with functions such as "conveying credibility and demonstrating that what researchers (in crowdfunding projects) do has potential for addressing social concerns" (Pérez-Llantada, 2021, p. 88).

On the other hand, the fact that stance bundles perform an important role as well, being epistemic expressions the most common functional subcategory suggests that researchers' positioning plays an important role to rephrase or recontextualize specialised information in an assertive (and hence persuasive) tone, as this is also the case for other web mediated genres,

such as blogs (Luzón, 2013), participatory citizen-science projects (Pérez-Llantada, 2023) and also edutainment genres such as TED Talks (Scotto di Carlo, 2014). The social exigence of all these genres is to adapt expert knowledge and make it accessible to broad audiences.

Lastly and more broadly, in response to RQ3: 'How do these recurring lexicogrammar patterns help fulfil the communicative goals of these projects?', this exploratory study has shown that phraseology is key to understanding crowdfunding proposals. In light of the results of the present analysis, one would argue that the presence of bundles especially frequent in formal academic prose co-occurring with bundles that are especially frequent in conversation indicates that this genre merges two different discourse styles to inform about science, to build trust in science and by this means persuade the readers so that they donate and back the projects. The high presence of bundles related to descriptive and deontic meanings recall the traditional genre of the research proposal. But accompanying them, the inclusion of stance bundles conveying attitudinal meanings and pragmatic politeness resembles the colloquial style that characterise science popularization genres such as didactic reportages, news in online journals, science blogs or infomercials (Caliendo, 2012).

The studies on bundle usage in academic prose and, in particular, journal article writing, indicate that formulaic expressions of stance are frequent because they have to do with the primary communicative purpose of specialized writing, namely persuading the scientific community of the relevance of the results of the study (e.g., *will keep you; allow us to, can be used*). This is a strategy that allows writers to present information as an opinion rather than an objective fact, being therefore closer to all types of publics (Hyland, 2008; Dontcheva-Navratilova, 2012). The results of the study indicated that the crowdfunding proposals for the discipline of art seems to also draw on phraseological bundles to convey stance while also performing referential functions. However, although Hyland's (2008) classification of discourse functions varies slightly from Biber's (2009) or the one used in this study, it is interesting to note that stance expressions are part of what Hyland conceptualises as participant-oriented bundles. These bundles are focused on the writer or reader of the text and not exclusively on one of them, such as research-oriented bundles or text-oriented bundles (Hyland, 2005). The fact, therefore, that a predominant presence of formulaic stance expressions conveys an assertive tone is related to the participation of both the project launcher and the backer that characterises the crowdfunding genre, a

participatory genre itself. As such, the exposition of activities and experiences typical of the investigation converge with others of an organizational and persuasive type of the text, as well as with the exchange of ideas with the project participants (the backers and the scientists) in tabs such as Lab Notes and Discussion. On the other hand, the study findings suggest that the affordances of the digital medium (hypertextuality, modularity and interactivity) would minimize the need to use other lexical bundles such as text organizers in high proportions to allow the relationship of a message with the surrounding discourse (Cortes, 2004), since the different sections and headings in the Experiment website would minimise the cognitive effort of the audience. Thus, special conversational bundles may be common in the crowdfunding genre due to its digital affordances too, being Experiment configured also as a space for discussion and digital forum, something not possible in other studied genres such as dissertations, research articles or doctoral thesis. In other words, this digital genre adopts characteristics of conversation, which build a colloquial tone that creates proximity to a wider audience with a more significant use of fuller dependent clauses than embedded ones. Nonetheless, an intrageneric study on the distribution of the taxonomy of bundles in the different sections (tabs) of Experiment.com should help to gain a better understanding of the functioning of this genre.

Despite the limitations of the small corpus size, this exploratory study lends credence that phraseology is key to understanding how the communicative goals of a given genre are realized by language forms at the level of lexicogrammar. However, how phraseology succeeds in meeting the specific communicative purposes important for the field of art remains exploratory. Although the present analysis has been carried out with projects in the field of art, variations in phraseology across the disciplines would possibly occur, being this area open to future research. To gain further insight into language-in-use at the level of phraseology, it would also be desirable to complement the findings of this study with qualitative (ethnomethodologically-informed) research. The insiders' perspectives could allow first-hand knowledge of the motivations, modes and choices that researchers make when composing a crowdfunding project proposal online.

Acknowledgements

The research conducted in this article has been funded by the Spanish Ministry of Economy and the State Research Agency under the project "Digital Genres and Open Science: an analysis of processes of hybridization, innovation and generic interdiscursivity" (project code PID2019-105655RB-I00). I also thank the support of the Government of Aragon to the research group CIRES "Comunicación internacional y retos sociales"/ "International communication and social challenges" (code H16_20R). This article is also a contribution to the research conducted in the Institute of Biocomputation and Physics of Complex Systems (BIFI) of the University of Zaragoza (Spain) and a contribution to the project Digital Language and Communication Resources for EU Scientists (DILAN) (project code 2022-1-ES01-KA220-HED-000086749) funded by the European Commission.

Article history:
Received 28 February 2023
Received in revised form 5 July 2023
Accepted 20 July 2023

References

Ädel, A., & Erman, B. (2012). Recurrent Word Combinations in Academic Writing by Native and Non-Native Speakers of English: A Lexical Bundles Approach, *English for Specific Purposes*, *31* (2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Anthony, L. (2023). *AntConc* (Version 4.2.2) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Barbieri, F. (2018). I don't want to and don't get me wrong: Lexical bundles as a window to subjectivity and intersubjectivity in American blogs. In J. Kopaczyk, & J. Tyrkkö (Eds.), *Applications of pattern-driven methods in corpus linguistics* (pp. 251–276). John Benjamins.

Belleflamme, P. P., Lambert, T., & Schwienbacher, A. (2013). Individual crowdfunding practices. *Venture Capital*, *15*(4), 1–21. https://doi.org/10.1080/13691066.2013.785151

Bestgen, Y. (2019), Comparing Lexical Bundles across Corpora of Different Sizes: The Zipfian Problem, *Journal of Quantitative Linguistics*, 27(3), 272–290. https://doi.org/10.1080/09296174.2019.1566975

Biber, D. (2009). A corpus-driven approach to formulaic language in English. Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, *14*(3), 275–311. https://doi.org/10.1075/ijcl.14.3.08bib

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*, 263–86. https://doi:10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical Bundles in University Teaching and Textbooks, *Applied Linguistics*, *25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*(1), 2–20. https://doi:10.1016/j.jeap.2010.01.001

Biber, D., & Gray, B. (2016). The competing demands of popularisation vs. economy: Written language in the age of mass literacy. In T. Nevalainen & E. C. Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 314–328). Oxford University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Ltd.

Bondi, M. (2009). Perspective and position in museum websites. In S. Radighieri, & P. Tucker (Eds.), *Point of view. Description and evaluation across discourses* (pp. 113–127). Officina Edizioni.

Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy, *BioScience*, *59*(11), 977–984. https://doi.org/10.1525/bio.2009.59.11.9

Caliendo, G. (2012). The popularization of science in web-based genres. In G. Caliendo & G. Bongo (Eds.), *The language of popularization: Theoretical and descriptive models* (pp. 101–132). Peter Lang.

Carter-Thomas, S., & Rowley-Jolivet, E. (2017). Open science notebooks: New insights, new affordances. *Journal of Pragmatics, 116*, 64–76. https://doi.org/10.1016/j.pragma.2016.12.003

Carrier, D. (2003). *Writing about visual art*. Allworth Press.

Chen, Y. (2004). *Educating art venue professionals: The current state of museum studies programs in the United States*. The Florida State University (unpublished doctoral dissertation).

Chen, Y., & Baker, P. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning and Technology*, *14*(2), 30–49. https://doi.org/10125/44213

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology, *English for specific purposes, 23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora, 3*(1), 43–57. https://doi.org/10.3366/E1749503208000063

Cortes, V. (2015). Situating lexical bundles in the formulaic language spectrum. In V. Cortes & E. Csomay (Eds.), *Corpus-based Research in Applied Linguistics: Studies in Honor of Doug Biber* (pp. 197-216). John Benjamins Publishing.

Cortes, V. (2022). Lexical bundles in academic writing. In R. Jablonkai, & E. Csomay (Eds.), *The Routledge handbook of corpora in language teaching and learning* (pp. 220–233). Routledge. https://doi.org/10.4324/9781003002901

Cortes, V., & Lake, W. (2023). LBiaP: A solution to the problem of attaining observation independence in lexical bundle studies. *International Journal of Corpus Linguistics, 28*(2), 263–277. https://doi.org/10.1075/ijcl.21100.cor

Curtis, K., Fry, M., Shaban, R. Z., & Considine, J. (2017). Translating research findings to clinical nursing practice. *Journal of Clinical Nursing, 26*, 5–6. https://doi.org/10.1111/jocn.13586

De Cock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair, & M. Hundt (Eds.), *Corpus Linguistics and Linguistic Theory* (pp. 51–68). Rodopi.

Donelli, C. C., Mozzoni, I., Badia, F., & Fanelli, S. (2022). Financing Sustainability in the Arts Sector: The Case of the Art Bonus Public Crowdfunding Campaign in Italy. *Sustainability*, *14*, 1641. https://doi.org/10.3390/su14031641

Dontcheva-Navratilova, O. (2012). Lexical bundles in academic texts by non-native speakers. *Brno Studies in English, 38*(2), 37–58.

Fletcher, C. (2002). Performance Appraisal and Management the Developing Research Agenda. *Journal of Occupational of Individual*, *74*, 473–487. https://doi.org/10.1348/096317901167488

Follett, R., & Strezov, V. (2015). An Analysis of Citizen Science Based Research: Usage and Publication Patterns, *PLoS ONE, 10*(11). https://doi.org/10.1371/journal.pone.0143687

Gledhill, C. (2011). The 'lexicogrammar' approach to analysing phraseology and collocation in ESP texts. *ASp*, *59*, 5–23. https://doi.org/10.4000/asp.2169

Gürler, C., & Çağlar, M. (2021). Success Prediction of a Crowdfunding Project in Art Categories. In R. Lenart-Gansiniec, & J. Chen (Eds.), *Crowdfunding in the Public Sector. Contributions to Finance and Accounting* (pp. 147–166). Springer. https://doi.org/10.1007/978-3-030-77841-5_10

Handke, C., & Dalla Chiesa, C. (2022). The art of crowdfunding arts and innovation: the cultural economic perspective. *Journal of Cultural Economics*, *46*, 249–284. https://doi.org/10.1007/s10824-022-09444-9

Herring, S. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen, & A. Trester (Eds.), *Discourse 2.0* (pp. 1–25). Georgetown University Press.

Hunston, S., & Francis, G. (2000). *Pattern Grammar: a corpus-driven approach to the lexical grammar of English,* Studies in Corpus Linguistics. John Benjamins.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173–192. http://dx.doi.

org/10.1177/1461445605050365

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation, *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/0.1017/S0261444808005235

Hyland, K., & Jiang, K. (2015). Academic lexical bundles: How are they changing*?, International Journal of Corpus Linguistics*, *23*(4), 383–07. https://doi.org/10.1075/ijcl.17080.hyl

Kabassi, K. (2017). Evaluating websites of museums: State of the art. *Journal of Cultural Heritage, 24*, 184-196. https://doi.org/10.1016/j.culher.2016.10.016

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, *99*(3), 422–431.

Lake, W., & Cortes, V. (2020). Comparing lexical bundles in Spanish and English psychology methods sections. In E. Friginal, & J. Hardy (Eds.), *Routledge Handbook of Corpus Approaches to Discourse Analysis* (pp. 334–353). Routledge.

Lazzeretti, C. (2016). *The language of museum communication: A diachronic perspective*. Palgrave Macmillan.

Lorenzo, D. (2011). *Lexical bundles in scientific English. A corpus-based study of native and non-native writing* (doctoral dissertation). Universidad de Barcelona.

Luzón, M. J. (2013) Public communication of science in blogs: Recontextualizing scientific discourse for a diversified audience. *Written Communication, 30*, 428–457. https://doi.org/10.1177/0741088313493610

Luzón, M. J., & Pérez-Llantada, C. (2022). *Digital Genres in Academic Knowledge Production and Communication: Perspectives and Practices*. Multilingual Matters.

McCarthy, K. F., Ondaatje, E. H., Brooks, A. & Szantó, A. (2005). *A portrait of the visual arts: Meeting the challenges of a new era*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/monographs/2005/RAND_MG290.pdf

Mehlenbacher, A. R. (2017). Crowdfunding Science: Exigencies and Strategies in an Emerging Genre of Science Communication, *Technical Communication Quarterly, 26*(2). https://doi.org/10.1080/10572252.2017.1287361

Mehlenbacher, A. R. (2019). *Science Communication Online: Engaging Experts and Publics on the Internet*. The Ohio State University Press.

Miller, C. R., & Kelly, A. R. (Eds.). (2017). *Emerging Genres in New Media Environments*. Palgrave Macmillan/Springer.

Parodi, G. (2015). Variation across university genres in seven disciplines. A corpus-based study on academic written Spanish. *International Journal of Corpus Linguistics*, *20*(4), 469–499. https://doi 10.1075/ijcl.20.4.03par

Paulus, T. M., & Roberts, K. R. (2018). Crowdfunding a "real-life superhero": The construction of worthy bodies in medical campaign narratives. *Discourse, Context & Media*, *21*, 64–72. https://doi.org/10.1016/j.dcm.2017.09.008

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent use. *Journal of English for Academic Purposes, 14*, 84–94. http://dx.doi.org/10.1016/j.jeap.2014.01.002

Pérez-Llantada, C. (2021). Grammar features and discourse style in digital genres: The case of science-focused crowdfunding projects. *Revista Signos, 54*(105), 73–96. https://revistasignos.cl/index.php/signos/article/view/363

Pérez-Llantada, C. (2022). Online data articles: The language of intersubjective stance in a rhetorical hybrid. *Language Communication, 39*(3), 400–425. https://doi.org/10.1177/07410883221087486

Pérez-Llantada, C. (2023). 'Help us better understand our changing climate': Exploring the discourse of Citizen Science. *Discourse & Communication*, *0*(0) (online first) https://doi.org/10.1177/17504813231158927

Römer, U. (2010). Establishing the phraseological profile of a text type: the construction of meaning in academic book reviews. *English Text Construction*, *3*(1), 95–119. https://doi.org/10.1075/etc.3.1.06rom

Salazar, D. (2014). *Lexical Bundles in Native and Non-native Scientific Writing. Applying a Corpus-based Study to Language Teaching*. John Benjamins.

Scotto di Carlo, G. (2014). The role of proximity in online popularisations: The case of TED talks. *Discourse Studies*, *16*(5), 591–606. https://doi.org/10.1177/1461445614538565

Swales, J. M. (2008). *Genre Analysis. English in Academic and Research Settings*. New York: Cambridge University Press.

Wray, A., & Perkins M. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20*, 1–28. http://dx.doi.org/10.1016/S0271-5309(99)00015-4

Zitzewitz, K., & Ciotti, M. (2022). Art and anthropology: Twenty-five years of The Traffic in Culture. *Journal of Material Culture*, *27*(1), 3–9. https://doi.org/10.1177/13591835221074151

Alberto A. Vela-Rodrigo is Assistant Lecturer at the Department of English and German Philology, University of Zaragoza, where he is developing his PhD dissertation with a 4-year grant of the Government of Aragon, Spain. His main research lines focus on corpus analysis and the role that digital genres play in scientific dissemination in Web 2.0 environments. He has also published academic and informative articles in the field of Oriental Arts.

NOTES

[1] The sources of all the excerpts in this article are listed in Annex 2.

ANNEX 1. Projects of the corpus and number of words (tokens) per project

Title	Researcher	N. of words (tokens)
Single-cell analysis of whole-body pattern formation: Mama had a baby, why did its head pop off?	Morgan Q. Goulding	3,825
Youth on the Move: Re-Storying Urban Communities with Public Art and Digital Media	Stephen Goss et al.	1,598
Does artistic experience affect the mental processes used when viewing visual art?	John Mullennix	2,280
Fresco Painting at the University of South Carolina: Medium of the Past, Art of Today	Taylor Tynes	8,69
VR for Virtually Everybody: Help us teach humanists and artists how to make virtual reality experiences	Marisa Parham et al.	2,396
Reconnecting Indigenous Collections to their Source Communities	Sydney Collins	983
Designing touch surfaces to help the visually impaired operate home appliances	Anne DeWitte	4,703
Draw Science: Open Access Infographic Journal	Viputheshwar Sitaraman	1,909
Animating Spinal Cord Damage: Building an Educational Website for Kids	Jeff Day	2,043
One step closer to the perfect chocolate chip cookie	Cindy Wu	1,603
Reducing anxiety with art therapy: A study of men with prostate cancer	Anna Celander	1,140
The Neuroscience of Digital Product Experiences	Lance & Juel	1,082
Drawing Archaeology in Byzantine Athens	Laura Leddy	1,258
What makes British and American choirs sound so different?	Anna Carro	1,109
Of Monks and Men: How medieval construction brought monasteries and lay communities together	Alexis Jackson	2,232
Can music improve memories in patients with brain damage?	Amy	1,132
Re-membering nuclear stories from a Maohi lens	Ganivet & Villierme	1,830
Can we hear how mycelia communicate?	Rita Hamm	2,074
Improving the qualitative properties of mycelium skins through fungal-bacterial biocomposites	Catherine Euale	1,292
Exploring material properties and fabrication processes of a mycelium based surfboard	Jessica Dias et al.	1,208
Understanding the terminal ballistics of Stone Age hunting projectiles	Devin Pettigrew	2,268
Using community science to evaluate the intersection of social, racial, and economic injustices in North Birmingham, AL	Rose Albert	806
How are rural communities on the edge of the largest tropical peat swamp forest in the Congo Basin using peat resources?	Shona Jenkins	2,450
Kumeyaay landscapes, knowledge transfer, and land stewardship in San Diego	Catherine Eng	1,572
How did Rome create new cities to build an Empire?	Hanna Friedman	2,392
How has modernization affected the population of rural Yuzawa, Niigata, Japan?	Steve Moody	1,339
Excavation of El Toril Cave to delve into the evolution in the Pleistocene of the Iberian Peninsula	Daniel García	1,142
How has Viking knitting evolved up to the 17th century?	Lindsey Howell	1,655
Friends and food...how does an artificial termite mound affect the social behavior of gorilla groups?	Caroline Jones	2,331
Does tattooing benefit the immune system? The inking of immunity	Christopher Dana	13,615
Where Was Vínland? Tracking Viking Explorers in the Americas Using Trace Element Analyses	Kevin Philbrook	6,190
Coquina in my Backyard: Can photogrammetry and 3D scanning be used to answer a 250 year old enigma?	Rebecca Harris	2,068
"Hog Meat en Rabbit en Fish en Such as Dat": Pre- and Post-Emancipation Foodways in the South Carolina Lowcountry	Brandy Joy	2,414
Troy: Archaeology of Archaeology	Gert van Wijngaarden	6,493
Fanteca Project: Student-led study of opiates and overdose in NYC	Ric Curtis	3,993
What Did 17th-Century Sailors Really Eat?	Grace Tsai	4,082
How did Paleolithic Hunter-Gatherers Use and Consume Plant Resources in Eurasia?	Aaron Jonas	4,939
UAV Infrared Mapping of Archaeological Sites in Greece	Hugh Thomas	4,847
How can a specialized diet affect the shape of the primate neck?	Abigail Nishimura	1,050
Who's there? Oh, it's you! Vocal recognition in western gorillas	Roberta Salmi	2,794

ANNEX 2. Crowdfunding proposals in Experiment.com referenced in this article

[1] Coquina in my Backyard: Can photogrammetry and 3D scanning be used to answer a 250 year old enigma? (2018) By Rebecca Harris and Kate Wright (https://experiment.com/projects/coquina-in-my-backyard-can-photogrammetry-and-3d-scanning-be-used-to-answer-a-250-year-old-enigma)

[2] Troy: Archaeology of Archaeology (2018) By Gert Van Wijngaarden, Ailbhe Ni Thoirealaigh, Vita Gerritsen, Nina Magdelijns, and Bart Rendering (https://experiment.com/projects/troy-archaeology-of-archaeology)

[3] Drawing Archaeology in Byzantine Athens (2014) By Laura Leddy (https://experiment.com/projects/drawing-archaeology-in-byzantine-athens)

[4] How did Paleolithic Hunter-Gatherers Use and Consume Plant Resources in Eurasia? (2017) By Aaron Jonas Stutz, Chantel White, and Liv Nilsson Stutz (https://experiment.com/projects/what-were-hunter-gatherers-doing-as-anatomically-modern-humans-spread-into-eurasia)

[5] Does artistic experience affect the mental processes used when viewing visual art? (2016) By John Mullennix (https://experiment.com/projects/does-artistic-experience-affect-the-mental-processes-used-when-viewing-visual-art)

[6] UAV Infrared Mapping of Archaeological Sites in Greece (2017) By Hugh Thomas (https://experiment.com/projects/infrared-mapping-of-archaeological-sites-in-greece)

[7] How did Rome create new cities to build an Empire? (2020) By Hannah Friedman, Katherine V Huntley, and Nicole Inghilterra (https://experiment.com/projects/how-did-rome-create-new-cities-to-build-an-empire)

[8] What Did 17th-Century Sailors Really Eat? (2018) By Grace Tsai and Elizabeth Latham (https://experiment.com/projects/what-did-17th-century-sailors-really-eat)

[9] Using community science to evaluate the intersection of social, racial, and economic injustices in North Birmingham, AL (2022) By Rose Albert (https://experiment.com/projects/community-science-to-evaluate-the-intersection-of-social-racial-and-economic-injustices-in-north-birmingham-al)

[10] Fanteca Project: Student-led study of opiates and overdose in NYC (2017) By Ric Curtis, Leonardo A. Dominguez Gomez, Sheng Li, Camila Gelpi-acosta, Douglas Goldsmith, Anjelica Maria Camacho, Rob Freeman, Popy Begum, Tabrina Youmans, Joana Bakiasi, Jonathan Rupay, Raymond Ruggiero, Sylvia Hamielec, Ashley Park, and David Frank (https://experiment.com/projects/fanteca-project-student-led-study-of-opiates-and-overdose-in-nyc)