Encuentro
Revista del Departamento de Filología Moderna

EXPLORING THE LEVEL OF SPECIALISATION OF A CORPUS OF SCIENTIFIC TEXTS AND STUDENTS' PERCEPTIONS AND NEEDS IN AN ESP CONTEXT

EXPLORANDO EL NIVEL DE ESPECIALIZACIÓN DE UN CORPUS DE TEXTOS CIENTÍFICOS Y LAS PERCEPCIONES Y NECESIDADES DE LOS ALUMNOS EN UN CONTEXTO DE INGLÉS PARA FINES ESPECÍFICOS

Antonio Martínez Sáez
Universidad Politécnica de Madrid

Abstract
The incorporation of some original texts combining the academic component with a particular scientific domain can be seen as an opportunity, in higher education and, in particular, in an English for Specific and Academic Purposes (ESP and EAP, respectively) context, to enable students to access authentic examples of the use of the language. This article investigates the degree of specialisation of some scientific texts that could be read and analysed in class in order to determine the pedagogical effect these might have when teaching the reading skill. In so doing, a corpus of twelve recent texts belonging to the domain of Biotechnology was analysed. Moreover, two different empirical studies were conducted in order to know more about students' needs and perceptions throughout two consecutive academic years; 2017-2018 (qualitative analysis of their initial perceptions) and 2018-2019 (qualitative and quantitative analysis of the data provided by a questionnaire). Preliminary findings suggested that the selection of texts showed different degrees of specialisation.
Key words: ESP, EAP, reading, specialisation levels, lexical density

Resumen
La incorporación de textos originales capaces de combinar el componente académico con el de un campo científico concreto puede verse, en el ámbito de la educación superior y, en concreto, en el del inglés para fines específicos y académicos, como una oportunidad para que los estudiantes accedan a ejemplos auténticos del uso de la lengua. En el presente artículo se estudia el grado de especialización de varios textos científicos con el fin de utilizarlos en el aula y determinar su efecto a nivel pedagógico al enseñar la destreza de la comprensión escrita. Para ello, se analizó un corpus de doce textos relacionados con la Biotecnología publicados recientemente y se llevaron a cabo dos estudios empíricos para indagar en las necesidades y percepciones de los estudiantes durante dos cursos académicos; 2017-2018 (análisis cualitativo de las primeras percepciones) y 2018-2019 (análisis cualitativo y cuantitativo de los datos aportados a través de un cuestionario). Los primeros resultados indicaron que los textos seleccionados mostraban diferentes grados de especialización.
Palabras clave: Inglés para fines específicos, inglés para fines académicos, comprensión escrita, niveles de especialización, densidad léxica

1. INTRODUCTION

In order to contextualise the present research, we might refer to the development of new areas of study and the evolution towards a much greater specialisation of knowledge as two of the main factors that lead English for Specific Purposes (henceforth ESP) instructors to face "the challenge of teaching a foreign language to the students of highly specialised disciplines" (Krajka 2009: 209). According to Paltridge and Starfield (2013), the field of English for Specific Purposes has expanded so as to include other areas such as English for Academic Purposes (EAP), English for Occupational Purposes (EOP), English for Vocational Purposes (EVP), etc. The abbreviation EAP, i.e. English for Academic Purposes, might thus be seen as a hyponym of ESP. In this study, we use the label ESP to refer to the teaching and learning of English as a foreign or second language in an attempt to allow students to use it in their particular professional, i.e. scientific and technical, domain. On the other hand, EAP might be defined as "the English that is taught to second-/third-language students preparing to enter an undergraduate and postgraduate course at universities and other institutions of higher education" (McCarter & Jakes 2009: 9). Additionally, in EAP, a language instructor "may be teaching students who are preparing to study or who are already studying subjects like business, sociology, law, finance, science or the arts" (McCarter & Jakes 2009: 9). In this context, language teachers "will have to learn the types of writing and reading assignments that are common to the particular subject areas of the students", which would encourage these professionals to "familiarise themselves with the particular ways their students have to write and the particular texts they have to read" (McCarter & Jakes 2009: 9).

Teaching the two written skills, i.e. reading and writing in higher education and, in particular, in the context of an English for Professional and Academic Communication course is always a challenging process. The combination of the labels "professional" and "academic" in the English classroom can make the teaching process even more encouraging and enriching by taking advantage of the opportunities provided by this dual approach. This article addresses the research conducted in an attempt to identify, implement, and use a corpus of specialised texts representing different genres related to the field of Biotechnology with undergraduate students of English. The final selection of recently published articles has already become the central component of one of the units of the course English for Professional and Academic Communication (henceforth EPAC) at the Technical University of Madrid (Universidad Politécnica de Madrid, henceforth UPM). In order to enrol in this EPAC course, students are required to certify an upper-intermediate level of English, B2 according to the Common European Framework of Reference for Languages. It is a compulsory subject (6 ECTS) at UPM and is mainly focussed on the development of a series of common competences and skills, i.e. note-taking; writing CVs and formal letters; analysing and producing written reports; or familiarising oneself with the situations, tools and guidelines for academic and professional oral communication, for instance, oral presentations, job interviews as well as meetings. Even though this EPAC course aims at providing students with the tools and methods for practicing, developing, and fostering a series of oral and written skills and competences in an academic and professional context, they usually express their interest in using materials and participating in activities closely related and adapted to their corresponding field of specialisation. This is a motivating factor which is especially evident in the case

of the Bachelor's degree in Biotechnology, since it has been perceived that students' level of involvement is very high and most of them expect and demand an English course as adapted to their scientific domain or subject area as possible. Furthermore, in the context of a technical university, in which students have a specialised background, it is relevant and even mandatory to explore and find new ways, methods, and techniques "capable of fostering an integral communicative competence and making the entire learning process more dynamic and attractive" (Gimeno-Sanz & Martínez-Sáez 2016: 31).

2. THEORETICAL FRAMEWORK

The writing and reading that a language instructor teaches "need to fit the academic purposes of their students" (McCarter & Jakes 2009: 9). Consequently, the "focus on preparing students for their specific academic requirements of English for higher education" would be "the main difference between EAP and general English Language Teaching (ELT)" (McCarter & Jakes 2009: 9). More specifically, focussing on the ways of fostering students' competence concerning the reading skill, the incorporation of original texts combining the academic component with the dissemination of knowledge regarding a particular scientific domain could also be seen as a promising opportunity that might allow learners to have access to authentic examples of the language in use. However, choosing the appropriate materials can be one of the toughest steps due to the importance of detecting the suitable level of specialisation and complexity of the written resources.

With regard to the use of appropriate teaching resources in these learning contexts, it might be mentioned that there is a "lack of properly developed commercial materials", which confronts many of the lecturers with the urgent need of adapting those resources (Krajka 2009: 209). As stated by Krajka (2009), there is a significant amount of target language input available on the Internet which would provide language instructors working in technical institutions with different paths. This would enable them to retrieve, evaluate, analyse, and start to implement a compilation of some of those online written materials in a varied range of English learning settings.

Basing students' language practice on a scientific field and implementing learning resources that cater for the students' specific needs have proven to foster their interest, since "the fact of providing learners with specific content is considered as a key factor in stimulating motivation; the more domain-related the content, the more relevant they are considered to be" (Gimeno-Sanz & Martínez-Sáez 2016: 35). Furthermore, learners usually rate the relevance of materials in connection with how much they think these satisfy their future needs and professional expectations (Gimeno-Sanz & Martínez-Sáez 2016). As defined by Räisänen and Fortanet-Gómez (2008: 12), ESP is aimed at providing the learner with the tools to "communicate within a specific academic, occupational or professional domain". These facts led us to start to design the structure of a unit that would pursue 1) the familiarisation with; 2) the understanding; and 3) the eventual production of scientific reports by students of Biotechnology at UPM. It is important to bear in mind that including technical language would require specific knowledge to be understood, and that some second-year students might not be

sufficiently prepared to face a very high level of specialisation. However, most of them are already familiar with complex vocabulary and these types of texts, since their lecturers in other technical courses often share, refer to or ask them to read original English versions. Moreover, the participants were presented with some of the key differences between general English and English for Specific Purposes as well as with relevant features that would enable us to define the language of scientific discourse.

These are some of the main reasons why it was deemed especially relevant to analyse the degree of specialisation of the scientific texts that would be eventually read and analysed in the ESP and EAP classroom context. There are two main ways of addressing the degree of specialisation (Ciapuscio & Kugel 2002; Edo-Marzà 2011: 298). Firstly, some authors refer to a "sharp line" or "clear-cut boundary" which allows for "a fast and easy distinction between general and specialised language" (Edo-Marzà 2011: 298). This distinction would be mainly based on "the specificity of the topics dealt with, addressers and addressees and communicative situations" (Edo-Marzà 2011: 298). In contrast, other authors mention that there is a continuum and that "the difference between general and specific is a matter of degree" (Cabré 1993, 1999; Edo-Marzà 2011: 298). The approach which has been adopted in this specific study coincides with the second perspective, and in particular with the three levels mentioned by Cabré, i.e. highly specialised, specialised and fairly specialised. Therefore, measuring the degree of specialisation was one of the initial priorities in order to present students with a broader variety of resources and different levels of complexity. In order to measure specialisation, lexical richness and variation (Edo-Marzà 2011) as well as function, genre and tenor relationship (Edo-Marzà 2011; Vargas-Sierra 2005) were taken into consideration.

3. METHODS

This section, as well as section four, has been divided into two main parts. On the one hand, we are presenting the method that was followed in order to measure the degree of specialisation of the corpus of twelve texts. On the other hand, we include a description of the way in which the two empirical studies were conducted throughout two consecutive academic years.

3.1. Measuring the degree of specialisation

The initial step consisted in selecting a corpus of twelve relevant texts from a series of online specialised journals related to the field of Biotechnology. Secondly, these texts were classified into three categories based on the genre to which they belong by taking into account the definitions provided for each of these text types, i.e. opinion[1], review[2] (and mini-review), and original[3] research

[1] "Opinion articles present the author's viewpoint on the interpretation, analysis, or methods used in a particular study. It allows the author to comment on the strength and weakness of a theory or hypothesis. Opinion articles are usually based on constructive criticism and should be backed by evidence. Such articles promote discussion on current issues concerning science. These are relatively short articles." (https://www.editage.com/insights/6-article-types-that-journals-publish-a-guide-for-early-career-researchers)

articles. The main topics included in the corpus covered the three biotechnological branches offered at UPM; plant, medical, and computational. The topics were crop modification, genetics, new medical techniques developed for targeting diseases in a more personalised way, bioinformatics or modelling. Thirdly, the need emerged to delimit their corresponding degree of specialisation and more particularly, their level of complexity, define their role and measure how appropriate and effective these could be for their eventual implementation in the classroom.

When measuring specialisation, both lexical richness and variation are seen as defining and key factors (Edo-Marzà 2011). However, lexical variation− "the ratio measured as a percentage between the different words in the text and the total number of running words (number of words x 100/ number of tokens"−can be affected by differences in length and be unstable in the case of short texts (Gregori-Signes & Clavel-Arroitia 2015: 548). Some other criteria are function, genre, and tenor relationship (Vargas-Sierra 2005; Edo-Marzà 2011). These criteria were also taken into consideration when selecting the texts from several prestigious publications and resources dealing with Biotechnology (see Appendix, part 1), through which experts address their research and findings to other scientists and students sharing their knowledge and interest in this particular field.

"Lexical density, lexical diversity or lexical richness are terms which refer to statistical measures that gauge the lexical richness of texts" (Daller, van Hout & Treffers-Daller 2003; Gregori-Signes & Clavel-Arroitia 2015: 547). In the present study, in which we have compared the level of complexity and students' needs and perceptions, we have narrowed the scope and focussed more specifically on lexical density (complexity factor). The factor of complexity "provides a measure of the proportion of lexical items (i.e. nouns, verbs, adjectives, and some adverbs) in the text (Johansson 2008; Gregori-Signes & Clavel-Arroitia 2015: 547). As claimed by these authors (2015), texts with lower density tend to be more easily understood. However, there is a paradox, since we might find a text with high lexical diversity (i.e. "contain many different word types") and with low lexical density (i.e. "contain many pronouns and auxiliaries rather than nouns and lexical verbs") or vice versa (Johansson 2008; Gregori-Signes & Clavel-Arroitia 2015: 547).

We measured lexical density using *Textalyser* (http://textalyser.net), which is presented as a free online text analysis tool able to offer detailed statistics, find out the keyword density, and analyse word groups as well as the prominence of words or expressions. In this study, we focussed our

[2] "Review Articles provide a comprehensive summary of research on a certain topic, and a perspective on the state of the field and where it is heading." "Reviews commonly cite approximately 100 primary research articles."(https://www.springer.com/gp/authors-editors/authorandreviewertutorials/writing-a-journal-manuscript/types-of-journal-articles/10285504).
They "give an overview of existing literature in a field, often identifying specific problems or issues and analysing information from available published work on the topic with a balanced perspective". (https://www.editage.com/insights/6-article-types-that-journals-publish-a-guide-for-early-career-researchers)

[3] "The most common type of journal manuscript used to publish full reports of data from research. It may be called an Original Article, Research Article, Research, or just Article, depending on the journal. The Original Research format is suitable for many different fields and different types of studies." (https://www.springer.com/gp/authors-editors/authorandreviewertutorials/writing-a-journal-manuscript/types-of-journal-articles/10285504)

analysis on the beginnings of the articles compiled and classified for the study, i.e. abstracts as well as other introductory sections. The main reason why we decided not to work on the complete versions was that the original texts were too long to be read and explored in the classroom in the context of a subject designed to cover a wider range of oral and written communicative tools and resources in just one semester. Additionally, this delimitation allowed us to measure lexical density in extracts of a more similar length−between 550 and 650 words per article (total word count provided automatically by *Textalyser*)−and be able to compare more homogenous sections regarding content and structure. This delimitation also helped us to draw attention to the initial parts, as we thought that focussing on the whole article would be much more time and effort consuming for students and might distort their global perceptions.

3.2. Empirical studies conducted in the classroom

In order to offer a more representative view of the comparison of the results of the analysis of the corpus with the results of the first phase of implementation of some of those texts in the classroom context, two different empirical studies were conducted. These studies enabled us to learn more about students' specific needs and perceptions throughout two consecutive academic years; 2017-2018 and 2018-2019. In the first case, we carried out a brief qualitative analysis of the texts with a group of 58 students following the above-mentioned English for Professional and Academic Communication course at the Technical University of Madrid. In so doing, a debate was held right after reading, analysing and exploring the main linguistic patterns of scientific discourse found in the abstract and introductory parts of at least one text per genre. The students were asked to express and share their opinion with the whole group by taking into account the level of difficulty, relevance and adequacy of the texts that had been selected.

A year later, a different group of 52 students who had also enrolled in the same course at UPM participated in the second study. The approach was both qualitative and quantitative in this second stage. We decided to incorporate a more tangible way of measuring students' impressions in an attempt to compare and consolidate the results which had been obtained the previous year, so the learners were asked to fill out a questionnaire. The survey (See Appendix, part 2) was aimed at showing what their main needs and perceptions were with regard to the usefulness and relevance of some of the original texts that had been used in class in relation to their own scientific domain. The online survey was administered through Google Forms. The first part of the questionnaire was designed in an attempt to know if the students of scientific disciplines were aware of the existence of different levels of specialisation and if they were able to detect them easily. We also wanted to determine if students perceived the distinction between general and specific as a continuum and thus based on a matter of degree or, on the contrary, if they were just able to notice a clear-cut boundary between general and specialised language. The participants were also asked about the kind of texts which should play a more dominant role in our EPAC course or about the texts they would be more interested in regarding the level of specialisation. The different options listed in this specific item were the following: highly specialised texts dealing with Biotechnology; highly specialised texts dealing

with Biotechnology and other scientific disciplines; specialised texts dealing with Biotechnology; specialised texts dealing with Biotechnology and other scientific disciplines; general texts including references to Biotechnology, general texts including references to Biotechnology and other scientific fields; and texts dealing with more general topics. We also tried to explore if the level of motivation is higher when reading texts related to one's own Bachelor's degree and the types of texts that are considered more useful and convenient in terms of developing students' reading and communicative skills in a specific academic, occupational or professional context.

As regards the section of the questionnaire dealing with the genres which had been previously analysed, our primary aim was to monitor if the students considered that there were remarkable differences in terms of the level of usefulness, suitability and adequacy of the texts to improve their linguistic or scientific competence. Moreover, they had to indicate if they were capable of detecting a different level of complexity when comparing original, review and opinion articles or if they were familiar with a more complex genre apart from the three types included in the study. They also had to specifically state which one of the texts read in class had been the most complex or hardest to understand. Additionally, the students were told to determine which particular linguistic feature was an indicator of a higher level of complexity. These were the options: complex terminology, complex noun phrases, abbreviations, impersonal structures, the use of the passive voice, the constant use of figures (numbers) or formulae. Furthermore, an open-ended option was provided so that they could emphasise other significant features apart from the ones listed in the survey. Moreover, the participants had to decide on the least complex text seen in class or if they would replace any of the texts provided. They also had to explain the reasons behind that particular decision.

4. RESULTS AND DISCUSSION

4.1. Measuring the degree of specialisation

As it has been previously mentioned, one of the main goals of the present study was to take the first steps and start to delimit the reading materials which were going to be shared with the students in the current and coming years. The initial hypothesis was that these resources would allow them to know more about text structure, degree of specialisation as well as other important features concerning scientific discourse, so that they might have the tools for understanding and writing technical reports in the near future.

Regarding lexical richness, and in particular the lexical density and complexity of the texts analysed, we were able to observe how the complexity factor (lexical density) in opinion articles−ranging from 62.2% to 67.7%−was higher than it was in the selection of review articles, offering more varied results, which ranged from 62.5% to 64.5% in the case of mini-reviews, and from 56.7% to 58.6% in both review articles. The most significant contrast was found in the results provided by the analysis of the corpus of original articles, which showed a more noticeable decrease in the level of

complexity or lexical density –ranging, in this case, from 53.5% to 56.9%–when compared with the two previous genres, which hypothetically means that students would find them easier to understand.

In general, as illustrated in Table 1, these results reflected a trend which can also be seen in other factors such as the average sentence length or readability when the Gunning-Fog and the alternative beta indexes are analysed. When we studied the results provided by the Gunning-Fox Index in the first readability test, we were able to see that the texts with a higher level of complexity or lexical density also showed a more prominent level of difficulty, which is normally equal to or higher than 13.6. The most remarkable data are provided by original articles 3 and 4, which, in spite of offering a lower level of lexical density and thus of complexity, would be more difficult to read than reviews 2 and 3. Similar results were revealed by the alternative beta index, as it illustrated that opinion articles 1 and 2 as well as original articles 3 and 4 would be the hardest to understand. On the other hand, the most accessible texts would be original articles 1 and 2.

As for the sentence length parameter, we observed that the most complex or lexically dense texts tended to be the ones including the longest sentences. However, there were two significant exceptions, since the numbers provided by both mini-reviews regarding the average length of their sentences were much lower than those of other texts with a similar level of lexical density. Moreover, review 2 revealed a more moderate complexity factor in spite of being one of the texts with more words per sentence.

Article	Complexity factor (Lexical density)	Readability (Gunning-Fog Index): (6-easy 20-hard)	Average sentence length (words)	Readability (Alternative) beta (100-easy, 20-hard, optimal 60-70)
Opinion 1	63.9%	14.8	22.08	6.8
Opinion 2	67.7%	15.6	26.35	6.3
Opinion 3	62.2%	14.3	21.68	12.4
Opinion 4	65.9%	14.3	23.64	13.1
(Mini-)review 1	64.5%	11.8	16.8	19.2
Review 2	56.7%	13.8	25	17.6
Review 3	58.6%	13.6	19.78	18.8
(Mini-)review 4	62.5%	10.3	13.37	17.8
Original 1	56.9%	10.2	14.37	30.5
Original 2	54.4%	11.7	19.05	24.7
Original 3	53.5%	14.2	17.57	5.6
Original 4	53.8%	14.6	18.28	8.6

Table 1. Results provided automatically by *Textalyser*

4.2. Empirical studies conducted in the classroom

Firstly, we conducted a brief qualitative analysis of the initial implementation of some of the texts representing each genre included in the corpus with a group of 58 students following the EPAC course at UPM in the academic year 2017-2018. In general terms, most of the students were able to detect different levels of complexity even within the same article and in comparison with other texts. In terms of usefulness and suitability, they shared their own conclusions with the whole group after reading and analysing the introductory sections. Besides, they were asked to think about the role of these extracts and if reading and analysing them could be considered as a useful tool and method to start to attain the above-mentioned linguistic goals. The first and general impressions revealed a favourable outcome.

However, with regard to the apparently high level of specialisation of the first texts read in class, some of the initial hypotheses were also confirmed, since the participants expressed the view that some parts were quite complex and difficult to understand in spite of their good command of English and familiarisation with some of the main ideas and specific topics. Initially, we also thought that those students who are more used to scientific and technical disciplines might be slightly reluctant to do an activity essentially based on linguistic analysis. Nevertheless, they seemed to appreciate the relevance of the contents and analysis carried out in class.

After the first steps of the study, we were aware of the limited scope and obvious restrictions presented by the debate held in class and its subsequent qualitative analysis, along with the impossibility of accepting the results as conclusive evidence, since our validation questionnaire had not been developed yet. A year later, the method that we decided to follow was quite similar in terms of the work done in class and the way the texts were presented and analysed. However, in this second stage of the study, a total of 52 students who had enrolled in the above-mentioned EPAC course submitted the questionnaire.

We were able to observe how 75% of the students considered that there was a clear-cut boundary between general and specialised language, 17.3% did not agree with that statement and 7.7% did not know how to answer the question. If we compare these results with the ones obtained in the following item, we can perceive how the same percentage of students, i.e. 75%, thought that the difference between general and specific was a matter of degree, which can be seen as paradoxical evidence. The results showed that the same number of students were defending the hypothesis based on a clear-cut boundary between the labels “general” and “specific” while stating that there are also different degrees between these two categories. 15.4% of the participants did not think that that distinction consisted in a matter of degree and 9.6 did not know which option was the most appropriate one. The results can be explained by saying that students tend to think that there is a very clear distinction between texts which deal with more general topics and those focussed on a specific domain. At the same time, they can notice that there is a scale within that specific field that would include texts showing a lower or a higher degree of specialisation.

Regarding the kind of texts which should play a more dominant role in the EPAC course, 42.3% of the participants reported that the most suitable option was the use of "specialised texts dealing with the field of Biotechnology and other scientific disciplines". 32.7% agreed that the best option was specialised texts related only to the field of Biotechnology; 11.5% of the students opted for a higher level of specialisation and selected the option "highly specialised texts dealing with the field of Biotechnology and other scientific disciplines" and a less significant number of students (5.8%) thought that using highly specialised texts focussed on the field of Biotechnology would be the most convenient method. Therefore, most of the students reported that a combination between biotechnological and other branches of science would be more enriching in their language learning process. The students were also aware of the level of complexity and difficulty of highly specialised texts and how the use of these original resources could be counter-productive in this learning context.

If we compare the previous results with the kind of texts the students were more interested in, we can observe a similar order of preferences as well as significant differences in the number of students supporting one option or another (See Figure 1). 36.5% of the participants would prefer to read and make use of "specialised texts related to Biotechnology and other scientific disciplines"; 25% opted for specialised texts dealing with Biotechnology, which means that most of the students perceived that specialised texts should be implemented and used in class instead of highly specialised resources focussed on Biotechnology (7.7%) or highly specialised texts combining this field with other disciplines (11.5%). Some students (5.8%) would also prefer to read and analyse more general texts including some references to Biotechnology and a higher percentage (11.5%) would be more interested in the implementation of more general texts combining the field of Biotechnology with other scientific topics.

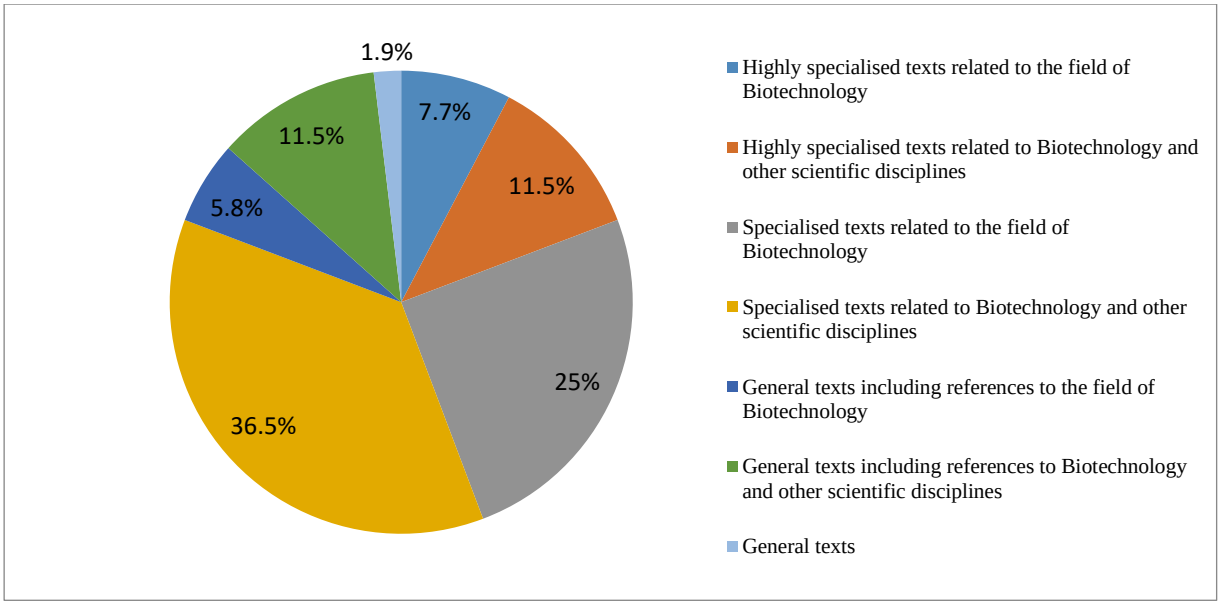

Figure 1. Question 4. What kind of texts are you more interested in?

Regarding students' academic and professional needs, some of the most useful topics listed were genetic engineering, immunology, drug design, and nanotechnology. Figure 2 illustrates the

opinions shared by the participants when asked if their level of motivation is usually higher when they read texts in English related to their specific Bachelor's degree. The percentage of students with no opinion was higher than expected, which had a slight negative impact on the number of students supporting the affirmative option. Nevertheless, the number of students who considered that the fact of using materials related to their own degree would raise their level of motivation was significantly higher.

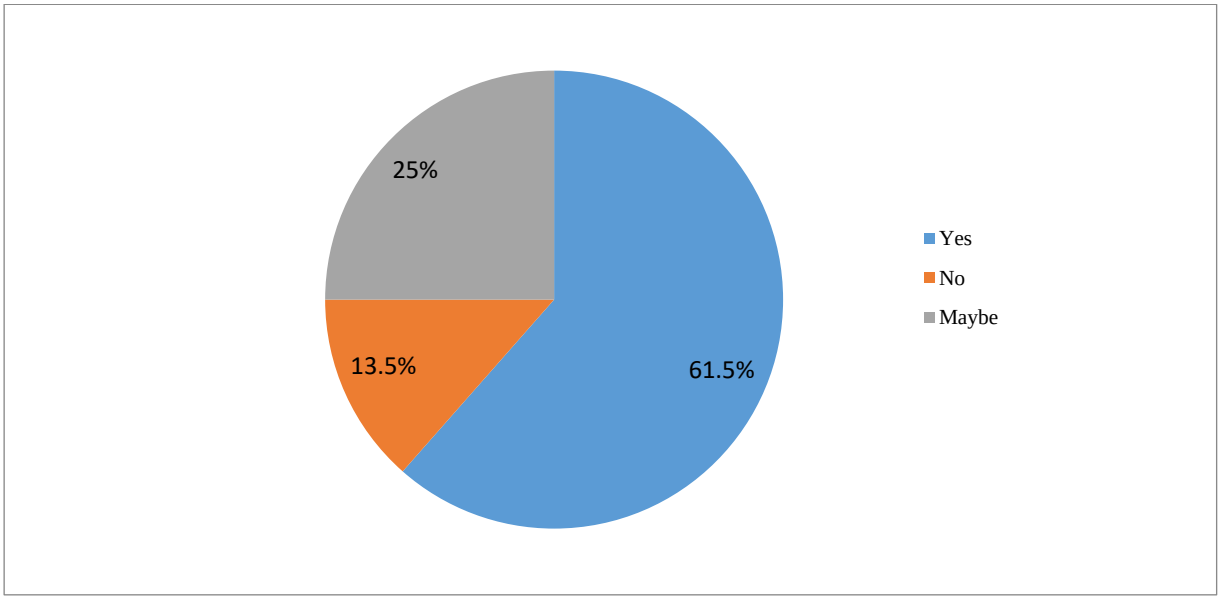

Figure 2. Question 6. Do you think that your level of motivation is higher when you read texts in English related to the Bachelor's degree in Biotechnology?

The following table shows the results of three consecutive questions, which were aimed at measuring students' perceptions regarding different levels of specialisation and how these could foster their communicative skills in a specific academic, occupational or professional context. As we can see, once again students tended to opt for specialised texts when these were compared to highly specialised or more general texts. The percentage of students who preferred a more general level in this particular context was slightly lower if we compare it to the highest degree of specialisation.

Question	Yes	No	Maybe
7. Highly specialised texts	65.4%	5.8%	28.8%
8. Specialised texts	80.8%	1.9%	17.3%
9. More general texts	61.5%	13.5%	25%

Table 2. Questions 7, 8, and 9. Do you think reading (7) highly specialised, (8) specialised, (9) more general texts would allow you to communicate in a specific academic, occupational or professional context?

Regarding the most useful genre to improve students' level of linguistic competence, 36.5% considered that opinion articles would be the most convenient choice; 34.6% supported scientific or original articles as the most effective genre; and 23.1% opted for reviews or mini-reviews. The third position was more surprising, since we did not expect to see such a high number of learners

considering reviews or mini-reviews as the most suitable genre to achieve a higher level of English in this specialised domain. Three students decided to offer their own alternative. Two of them suggested a combination of scientific articles with opinion articles or scientific articles with reviews, and one student emphasised the potential of oral debates held in the classroom context. However, when asked about the genre that would best help them improve their level of scientific competence, an overwhelming percentage (88.5%) deemed the scientific article as the most effective one, followed by reviews/mini-reviews (5.8%) and opinion articles (1.9%). These results allowed us to see a very clear distinction between the tools they would use to improve their language level in this ESP and EAP context and the type of texts they find most reliable when it comes to expanding their scientific knowledge. This was probably due to the fact that scientific articles are usually thought to provide more innovative, up-to-date and captivating data as well as significant findings.

Concerning students' perceptions of the level of complexity and in order to compare them with the results presented in section 4.1, they were asked to point to the most complex genre after reading their introductory sections. An overwhelming percentage of learners (82.7%) stated that the scientific article was the genre showing the highest level of complexity. Only 7.7% of the participants opted for reviews/mini-reviews and a very low number (3.8%) considered that the opinion article was the most complex type of text. There were only three students, out of the 52 participants, who had no opinion or thought that the three alternatives were equally complex. Going back to the results of our corpus analysis, we can observe significant differences between the statistical data provided by *Textalyser* in order to determine the genre showing the highest level of complexity and the results provided by the analysis of students' perceptions. These results coincided with the data retrieved from the questions in which the participants were asked to say which articles had been the most complex and simple ones. We were able to see how students tended to rate original articles as the most complex texts they had had to read, whereas most of them found that the mini-review had been more accessible.

One of the questions centred around students' opinions regarding the level of the articles read in class. The results coincided with the perceptions shared orally a year earlier, since 53.8% of the participants found that the level was high and 40.4% that it was appropriate. 5.8% considered that the level of the texts was low. None of the learners thought that the level was too high or too low.

Furthermore, the questionnaire included an item aimed at showing which linguistic component made students think that a text was more complex than another. In this particular case, a very high number of students (57.7%) considered that there was a clear correlation between the level of complexity and the use of very specialised terminology, whereas 34.6% of the participants thought that the level of complexity was clearly determined by the common use of long and complex noun phrases. Three students pointed to formulae, abbreviations and the passive structure as the cause of a higher level of complexity. One of the participants suggested that it was a matter of content and made reference to the fact they were still in their second year and had to get more familiar with some of the concepts included in the articles.

Lastly, when asked whether they would replace any of the texts read in class or not, a vast majority said that they would not substitute any of them, whereas some other students emphasised the idea that less complex texts and even more general examples might be more encouraging, interesting and useful in order to improve their level of reading competence.

5. CONCLUDING REMARKS

Once the analysis of the texts was completed, preliminary findings suggested that the corpus was able to show different degrees of specialisation. However, we should also mention that this is still a study in progress. Although we are aware of the need to expand and achieve a more representative corpus, we might say that some of the preliminary findings contribute to the initial theory of the existence of a continuum or a progressive degree of specialisation, even within texts addressed by and to experts in the same scientific area. As we have seen, the corpus comprising twelve texts is able to present the reader with different degrees of high complexity depending on the type of text. We have also started to explore the specific pedagogical effects that these resources have when used in an ESP and EAP learning context, and have seen that lecturers' intuition and previous hypothesis about a likely increase in students' motivation and progress is confirmed by the results provided by the questionnaire responded by a total of 52 participants in the second stage of the study.

These initial results have enabled us to see a noticeable effect in terms of the familiarisation of students with the structure of scientific publications and the main features of scientific discourse, since the texts read in the classroom showed some of the main linguistic patterns defining the above-mentioned genres. As it has been previously shown, apart from lexical density, we have also considered factors such as readability or the average sentence length, which has allowed us to detect significant correlations and some other discordant features. For instance, we can highlight the fact that a higher level of complexity does not necessarily correlate with a lower readability index or with higher numbers in relation to the average sentence length. We have also been able to take a very significant step and explore the contrasts and similarities between the results that were automatically obtained and the perceptions shared by the students, who showed their interest and preference for specialised texts instead of for highly specialised resources. Furthermore, there are some other factors concerning the degree of specialisation of scientific resources which should be taken into consideration in the near future in order to offer a broader and complete view of all the components taking part in the present approach to ESP and EAP learning contexts.

Furthermore, taking into account students' perceptions in the present study, one of the concluding remarks would be that the use of authentic and original materials was also seen as an adequate tool to improve students' written comprehension skill in this particular case and context. Additionally, we were able to observe how this practice could allow them to access texts and resources that might have a very positive effect and enhance their careers as scientists or engineers in the coming years. They would learn more about the latest developments and breakthroughs in their corresponding scientific domains while getting familiar with some of the most significant patterns of scientific

discourse. Therefore, it would always be necessary to check the level of adequacy, complexity and suitability of the resources implemented in the ESP and EAP classroom. Bearing students' perceptions and level of motivation in mind has also proven to be a very useful method to finally determine which learning tools could play a more effective role in our EPAC course. In so doing, language instructors might have a more varied and representative view of their students' behavioural characteristics, learning preferences as well as expectations, which would have a positive impact on the whole learning process.

REFERENCES

Cabré Castellví, M. T. 1993. *La terminología: Teoría, metodología, aplicaciones*. Empúries: Editorial Antártida.

Cabré Castellví, M. T. 1999. *La terminología, representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Institut Universitari de Lingüística Aplicada (IULA) – Universitat Pompeu Fabra.

Ciapuscio, G. and Kugel, I. 2002. Hacia una tipología del discurso especializado: aspectos teóricos y aplicados. In J. García and M. T. Fuentes (Eds.), *Texto, terminología y traducción* (pp. 37-73). Salamanca: Ediciones Almar.

Daller, H., van Hout, R., and Treffers-Daller, J. 2003. Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24(2), 197-222.

Edo Marzà, N. 2011. How to Interpret the Type-Token Ratio: A Helpful Measure for Balanced and Representative Corpus Compilation. In S. Maruenda & B. Clavel (Eds.), *Multiple Voices in Academic and Professional Discourse: Current Issues in Specialised Language Research, Teaching and New Technologies*, (pp. 298-308). Newcastle upon Tyne: Cambridge Scholars Publishing.

Gimeno-Sanz, A. and Martínez-Sáez, A. 2016. The design and integration of ESP content in an upper-intermediate online course. *Revista de Lenguas para Fines Específicos*, 22(1), 31-53.

Gregori-Signes, C. and Clavel Arroitia, B. 2015. Analysing lexical density and lexical diversity in university students' written discourse. *Procedia - Social and Behavioral Sciences*, 198, 546 – 556.

Johansson, V. 2008. Lexical diversity and lexical density in speech and writing: a developmental perspective. Lund University, Dept. of Linguistics and Phonetics Working Papers, 53, 61-79.

Krajka, J. 2009. Promoting discovery learning in English for Academic and Professional purposes: Implementing concordancing procedures in online courses. In I. González-Pueyo, C. Foz-Gil, M. Jaime Siso & M. J. Luzón Marco (Eds.), *Teaching academic and professional English online* (pp. 209-228). Bern: Peter Lang.

McCarter, S. and Jakes, P. 2009. Uncovering EAP: How to Teach Academic Writing and Reading. Oxford: Macmillan.

Paltridge, B. and Starfield, S. 2013. *The Handbook of English for Specific Purposes*. Chichester, UK: Wiley-Blackwell.

Räisänen, C., and Fortanet-Gómez, I. 2008. The state of ESP teaching and learning in Western European higher education after Bologna. In I. Fortanet-Gómez & C. A. Räisänen (Eds.), *ESP in European higher education: Integrating language and content* (pp. 11-51). Amsterdam: John Benjamins Publishing Co.

Textalyser.net Tools for Analysis, http://textalyser.net/

Vargas Sierra, C. 2005. *Aproximación terminográfica al lenguaje de la piedra natural. Propuesta de sistematización para la elaboración de un diccionario traductológico* (PhD Dissertation). Alicante, Spain.

APPENDIX

1. LIST OF ARTICLES ANALYSED IN THE PRESENT STUDY (see Table 1):

A. Opinion Articles

A.1. Vervoort, Y., Gutiérrez, A., Roncoroni, M., Liu, C., Steensels, J. & Verstrepen, K. 2017. High-throughput system-wide engineering and screening for microbial biotechnology. *Current Opinion in Biotechnology, 46*, 120-125.
A.2. Tsigkinopoulou, A., Baker, S. M. & Breitling, R. 2017. Respectful Modeling: Addressing Uncertainty in Dynamic System Models for Molecular Biology. *Trends in Biotechnology, 35*(6), 518-529.
A.3. Keidar, M., Yan, D., Beilis, I. I., Trink, B. & Sherman, J. H. 2018. Plasmas for Treating Cancer: Opportunities for Adaptive and Self-Adaptive Approaches. *Trends in Biotechnology, 36*(6), 586-593.
A.4. Malyska, A., Bolla, R. & Twardowski, T. 2016. The Role of Public Opinion in Shaping Trajectories of Agricultural Biotechnology. *Trends in Biotechnology, 34*(7), 530-534.

B. Review Articles

B.1 Regnat, K., Mach, R. L. & Mach-Aigner, A. R. 2018. Erythritol as sweetener—wherefrom and whereto? (Mini-review). *Applied Microbiology and Biotechnology, 102*(2), 587–595.
B.2. Sprague, M., Betancor, M. B. & Tocher, D. R. 2017. Microbial and genetically engineered oils as replacements for fish oil in aquaculture feeds. *Biotechnology Letters, 39*(11), 1599–1609.
B.3. Backofen, R., Engelhardt, J., Erxleben, A., Fallmann, J., Grüning, B., Ohler, U., Rajewsky, N. & Stadler, P. F. 2017. RNA-bioinformatics: Tools, services and databases for the analysis of RNA based regulation. *Journal of Biotechnology, 261*, 76–84.
B.4. Löfblom, J., Rosenstein, R., Nguyen, M., Ståhl, S. & Götz, F. 2017. Staphylococcus carnosus: from starter culture to protein engineering platform (Mini-review). *Applied Microbiology and Biotechnology, 101*(23-24), 8293-8307.

C. Original Research Articles

C.1. Raj, R. S., Thakur, S. V., Hussen, V. S., Joshi, M. N., TyagiSnehal, S. N. & Bagatharia, B. 2017. Development of PR genes panel for screening aphid-tolerant cultivars in Brassica juncea. *3 Biotech, 7*(129), https://doi.org/10.1007/s13205-017-0785-7
C.2. Naresh, P., Reddy, M. K., Reddy A. C., Lavanya, B., Lakshmana Reddy, D. C., & Madhavi Reddy, K. 2017. Isolation, characterization and genetic diversity of NBS-LRR class disease-resistant gene analogs in multiple virus resistant line of chilli (Capsicum annuum L.). *3 Biotech, 7*(114), https://doi.org/10.1007/s13205-017-0720-y
C.3. Nielsen, M., Holst-Fischer, C., Malmgren-Hansen, B., Bjerg-Nielsen, M., Kragelund, C., Møller, H. B., Mørck Ottosen, L. D. 2017. Small temperature differences can improve the performance of mesophilic sludge-based digesters. *Biotechnology Letters, 39*(11), 1689-1698.
C.4. Kroll, P., Stelzer, I. V., Herwig, C. 2017. Soft sensor for monitoring biomass subpopulations in mammalian cell culture processes. *Biotechnology Letters, 39*(11), 1667-1673.

2. SURVEY: Degree of specialisation of a corpus of biotechnological texts (UPM, 2018)

1. Do you consider that there is a clear-cut boundary between general and specialised language?

Yes / No / I do not know.

2. Do you think the difference between general and specific is a matter of degree (specialisation level)?

Yes / No / I do not know.

3. What kind of texts should play a more dominant role in our EPAC course?

- Highly specialised texts related to the field of Biotechnology

- Highly specialised texts related to Biotechnology and other scientific disciplines

- Specialised texts related to the field of Biotechnology

- Specialised texts related to Biotechnology and other scientific disciplines

- General texts related to the field of Biotechnology

- General texts related to Biotechnology and other scientific disciplines

- General texts

- Other:

4. What kind of texts are you more interested in?

- Highly specialised texts related to the field of Biotechnology

- Highly specialised texts related to Biotechnology and other scientific disciplines

- Specialised texts related to the field of Biotechnology

- Specialised texts related to Biotechnology and other scientific disciplines

- General texts related to the field of Biotechnology

- General texts related to Biotechnology and other scientific disciplines

- General texts

- Other:

5. Think about your academic and professional needs. Please, list three useful topics.

6. Do you think that your level of motivation is higher when you read texts in English related to the Bachelor's degree in Biotechnology?

Yes / No / Maybe.

7. Do you think reading highly specialised texts would allow you to communicate in a specific academic, occupational or professional context?

Yes / No / Maybe.

8. Do you think reading specialised texts would allow you to communicate in a specific academic, occupational or professional context?

Yes / No / Maybe.

9. Do you think reading more general texts would allow you to communicate in a specific academic, occupational or professional context?

Yes / No / Maybe.

10. What would you say is the most useful genre to improve your level of linguistic competence?

- Scientific articles

- Opinion articles

- Reviews or mini-reviews

- Other:

11. What would you say is the most useful genre to improve your level of scientific competence?

- Scientific articles

- Opinion articles

- Reviews or mini-reviews

- Other:

12. What do you think is the genre with the highest level of complexity?

- Scientific articles

- Opinion articles

- Reviews or mini-reviews

- I do not know.

- Other:

13. What do you think about the level of the articles in Unit 5 (Moodle)?

- The level was too high.

- The level was high.

- The level was appropriate.

- The level was low.

- The level was too low.

14. What was the most complex text in Unit 5?

15. What makes you think that a text is more complex than others?

- Complex terminology

- Complex noun phrases

- Abbreviations

- Impersonal structures

- Passive voice

- Figures (numbers)

- Formulae

- Other:

16. What was the least complex text in Unit 5?

17. Would you replace any of the texts? Please, say why.