

The rhetorical structure of biomedical sub-genres

Esther Vázquez y del Árbol*

Abstract: The main purpose of this article is to look for differences in the structural organization of two biomedical sub-genres. Accordingly, it pursues two secondary goals: (1) to illustrate the effectiveness of a corpus-based research study in which a genre-based analytical approach was used to represent the rhetorical structure of texts; and (2) to prove the usefulness of this method for teachers, students, and even translators. In order to achieve these goals, we applied genre-based analysis to a corpus of 100 medical texts in English and Spanish (25 Letters to the Editor, 25 *Cartas al director*, 25 Editorials, and 25 *Editoriales*). Using Paltridge's approach, we developed a linguistic method to analyze the rhetorical structure of texts belonging to the field of Medicine. After studying the informational structure of our texts and analyzing the results, we found striking differences in how discourse is organized in texts belonging to the different sub-genres. In fact, such differences abound, and there are very few elements in common. This study shows that each genre and sub-genre has unique linguistic patterns that differ from those found in other genres, and that it is possible to demonstrate these patterns by using a linguistic approach to analyze genre-specific corpora.

La estructura retórica de los subgéneros biomédicos

Resumen: Este artículo presenta como principal objetivo la búsqueda de diferencias en la organización estructural de dos subgéneros biomédicos. Se plantea, para ello, dos metas secundarias: 1) demostrar la eficacia de una investigación basada en el corpus desde el punto de vista del análisis del género, con objeto de simbolizar la estructura retórica textual; y 2) ratificar la utilidad de este método para profesores, estudiantes e incluso traductores relacionados con el ámbito de las lenguas para fines específicos. Con objeto de alcanzar tales objetivos, hemos puesto en práctica un análisis, basado en el género, de un corpus formado por 100 textos médicos en inglés y en español (25 *Letters to the Editor*, 25 *Cartas al director*, 25 *Editorials* y 25 *Editoriales*). Siguiendo la propuesta de Paltridge, hemos desarrollado un análisis, similar al del lingüista, que se ha aplicado a la estructura retórica de textos pertenecientes al área de medicina. Al estudiar la estructura de la información de nuestros textos y analizar los resultados, observamos contrastes significativos tras comparar la organización discursiva en textos de diferentes subgéneros; de hecho, presentan abundantes diferencias y pocos aspectos en común. Ello demuestra que cada género y subgénero presenta sus propias características lingüísticas, las cuales no coinciden con las del resto de los géneros; cuestión que se puede ratificar al aplicar una propuesta de enfoque lingüístico al análisis de corpus pertenecientes a un género concreto.

Key words: genre, contrastive text analysis, scientific genres, letters, editorials. **Palabras clave:** género, análisis textual contrastivo, géneros científicos, cartas, editoriales.

Panace@ 2005; 6 (19): 28-31

1. Introduction

Since genre has become a popular framework for investigating the form and function of both oral and the written discourse, it also serves as a tool for developing educational practices in the field of Language for Specific Purposes (LSP).

Nevertheless, there has been little research on how the various parts of a genre or sub-genre differ from each other (e.g. Ghadessy 1983; Henry & Roseberry 1995, 1996; Swales 1990), on the contrasts between different LSP texts, and on the usefulness of genre analysis for language teaching. Since most research carried out so far has concentrated on the genre of the Research Article and its sub-genres (Research Notes, Survey Papers, Conference Papers, etc.), we have focused on the rhetorical structure of different biomedical sub-genres (Letters to the Editor and Editorials), and studied a random sample from a bilingual corpus of 100 texts.

2. Methodology

2.1. The corpus

To achieve our linguistic aims, we designed a corpus made up of real texts belonging to a specific linguistic code (Biomedicine). All texts were extracted from national or international biomedical journals. We focused on two languages—English and Spanish—and randomly selected 100 opinion texts: 50 English texts (25 Editorials and 25 Letters to the Editor) and 50 Spanish texts (25 *Editoriales* and 25 *Cartas al director*). This corpus was created to help meet the need for situation-specific corpora and study the structure of different sub-genres.

2.2. The analysis and its framework

As noted, our framework is influenced by Paltridge's (1997) proposal, which is based in turn on Swales's (1981) genre-analysis model. In his research Paltridge (1997: 111)

* Autonomous University of Madrid (Spain). Address for correspondence: esther-vazquez@telefonica.net.

analyzed twelve texts from Environmental Studies area using several keys, each one representing a different part of the discourse:

- BI** - Background information
- IG** - Indicating a gap
- RS** - Rationale for the study
- PR** - Previous research
- M** - Materials
- C** - Conclusions
- JS** - Justification for the study
- PS** - Purpose of study
- QR** - Question-raising
- CS** - Context of the study
- R** - Results

In accordance with Hasan’s concept of *generic structure potential* (‘the total range of text structures available within a genre’) (Hasan as cited in Paltridge 1997: 66), Paltridge (1997: 66) demonstrated ‘what elements *must* occur; what elements *can* occur; where elements *must* occur; where elements *can* occur; and *how often* elements can occur’.

This method was chosen for our texts, mainly for three reasons: (1) it is a simple method of rhetorical structure representation; (2) it facilitates counting of the results obtained in an analysis of a large corpus; and (3) after the results are analyzed, more specific conclusions can be easily obtained.

Since our aim was to present a simple schema representative of the rhetorical structure of each sub-genre, we worked with two different units of analysis: (1) triads and (2) keys. The first are larger structural units that consist of three elements, according to position and function: *Introduction*, *Development* and *Final Conclusion*.^a We will refer to each of them by means of a single upper case letter (*Introduction*, section **A**, *Development*, section **B**, *Final Conclusion*, section **C**). The keys are the minimum unit of analysis and the main tool in our analysis of information structure. We therefore represented structural content by assigning at least one key to each sentence. In cases when more than one key element was needed to symbolize the content, we added a + symbol to show that all keys represent the same sentence.

After reading and analyzing the biomedical corpus, we identified a large number of keys:

- A** - Advice
- AP** - Agreement with previous studies
- BI** - Background information
- CO** - Conclusions
- CL^b** - Closure
- CS** - Context of the study
- DP** - Disagreement with previous studies
- H** - Hypothesis
- IC** - Information about the current moment
- IF** - Information about the future
- IG** - Indicating a gap
- MA** - Materials
- ME** - Methods
- NFW** - Need for further work in the future
- PR** - Previous research
- PRS** - Presentation of the study
- PS** - Purpose of the study
- QR** - Question-raising
- R** - Results
- RAP** - Results that agree with previous studies
- RDP** - Results that disagree with previous studies
- R/JS** - Rationale/Justification for the study
- SA** - Salutation
- (Hid)** – Hidden

As noted, owing to the fact that we worked with texts in biomedicine—an area different from that studied by Paltridge (1997), some elements in our list differ from the elements used by this author:

1. We found, in our texts, all the elements described by Paltridge, although we have changed some of their names: C is designated here CO, M is designated MA, and RS and JS have been combined in a single element, R/JS.
2. The element named R/JS is the equivalent for two other elements in Paltridge’s proposal: *Rationale for the study* (RS) and *Justification for the study* (JS). We decided to make this change because we felt that the idea symbolized by these two elements is the same, i.e., the reason for choosing the topic of the text.
3. The corpus textual structure contains some new elements that did not appear in Paltridge’s list: A, AP, CL, DP, H, IC, IF, ME, NFW, PRS, RAP, RDP, and SA.
4. The element modifier **Hid** (*Hidden*) was included to indicate that the idea expressed (and represented by a key) is more implicit than the remaining information.
5. The most frequent elements have been highlighted in bold type.

According to our analysis, there are ten elements which were found to occur in most of the texts (BI, CO, IG, MA, ME, PR, PRS, R, PS, R/JS). These elements were classified as ‘basic elements.’ Obviously, there are also elements which are less frequent and relevant for the text structure, these have been classified as ‘non basic elements’ (A, AP, CL, CS, DP, H, IC, IF, NFW, QR, RAP, RDP, SA). These elements were not included in our list, since they are not fundamental to the rhetorical structure of our texts.

3. Results

Due to space limitations, and for the sake of simplicity, the data from all 100 items in our corpus are not reported here. However, they represent a good source for subsequent, in-depth contrastive analyses.

Here we use a simple schema to represent the structure of the texts in our corpus, where the elements are organized according to their order of appearance in the texts. To this end we decided, not to include frequent combinations of elements (i.e., H+IF+IG) found in a single sentence, because this would enlarge the schema. We also decided to use a list of symbols in order to explain the frequency of occurrence of an element, its repetition, or its location. Before presenting our four schemata (one for each sub-genre), we explain below the symbols used to detail information content within each schema:

1. ++ means that an element may appear more than once in the section indicated. Therefore, an entry of MA++ in section B (*Development*) means that this element appears several times in the *Development* section.

2. * indicates the possibility that the element appears in sections other than the section indicated for a given entry. For example, an entry of PR* in section B means that this element is often found in the Development section, although it has also been used in the other two sections (Introduction, A, and Final Conclusion, C).
3. *++ means that the element modified by ++ may appear in the section indicated and also in the other two sections, and that it tends to appear more than once. For example, an entry of R/JS*++ in section A may also appear, more than once, in section B or C.
4. / means that the keys appear in the same place and section, although the key that appears to the left of the slash is the more common element.
5. According to this classification system, the information structure of the four sub-genres studied here can be summarized as follows:

<i>Cartas al Director</i>	Letters to the Editor
Introduction (A)	Introduction (A)
SA	SA
BI *++	<u>PR</u> *++
R/JS *++	R/JS ++
PR*	<u>BI</u> *++
PS*	
Development (B)	Development (B)
BI	BI
PRS *	PRS
MA ++	MA/PS
ME ++	ME++
R ++	R*++
Final Conclusion (C)	Final Conclusion (C)
CO *++	CO *++
IG=	IG=

<i>Editoriales</i>	Editorials
Introduction (A)	Introduction (A)
BI *++	BI *++
R/JS *	R/JS
Development (B)	Development (B)
BI	BI
R/JS	PRS
PRS*/(PS)*	PR*
PR	MA
R	ME
	R
Final Conclusion (C)	Final Conclusion (C)
CO *++	CO *++
(PS)	PS
IG=	IG=

4. Discussion

The results of our analysis of genre-specific texts show that the pattern of information structure identified in the corpus is the same in all the texts, since all structures are composed of the three elements *Introduction*, *Development* and *Final Conclusion*.

Our analysis of textual structure indicates that these biomedical texts have the *same field*, but they do not share the same communicative purpose. For example, the correspondence sub-genres (*Cartas al director* and Letters to the Editor) deal with topics that are more specific than what is found in Editorials. In contrast, the topics dealt with in *Editoriales* and Editorials are more general, since they deal with a variety of different problems.

The structures of all the sub-genres analyzed are heterogeneous and flexible, especially in *Editoriales* and Editorials. In fact, after comparing the two correspondence sub-genres (*Cartas* and Letters) we found fewer differences than between *Editoriales* and Editorials.

Our comparison of texts in different languages but belonging to the same sub-genre (*Cartas* versus Letters and *Editoriales* versus Editorials) revealed fewer differences in their rhetorical organization than we found when we compared texts belonging to different sub-genres. Nevertheless, on comparing texts in the same language but belonging to different sub-genres (*Cartas* versus *Editoriales* and Letters versus Editorials) greater differences were found.

The content of Editorials and *Editoriales* was less specific, since they do not deal with information about the materials (MA) and methods (ME) used in research. Many differences in textual organization were found between Editorials and *Editoriales*, since their structure is less rigid than those of *Cartas* and Letters.

To conclude, the four schemata we obtained on the basis of an analysis of the contents of each sub-genre show that each sub-genre has its own rhetorical structure.

5. Implications for teaching

Our research suggests that the genre-based approach focused on rhetorical organization is an interesting tool that advanced L-2 students may use to master the functions and linguistic conventions of either oral or written texts that are required in academic settings. This method may also help LSP learners to organize their ideas into set groups which will result in a highly textured writing.

The linguistic characteristics of a biomedical corpus can be described accurately only after an exhaustive analysis. Such an analysis is what makes this method useful for translators. As an aid to translation, two types of corpora can be fruitful: those which contain and study translations, and those which offer different resources for carrying out translations.

In conclusion, as illustrated by our analysis of the corpus used in this research, specialized, genre-specific corpora will become increasingly necessary not only in academic but also in professional settings. This type of analysis can provide not only the LSP learner but also the LSP teacher with knowledge

about the structure of information necessary for recognizing, writing, translating or even teaching any type of genre or sub-genre.

Notes

^a “Final Conclusion” is designated thus in order to avoid confusion with the term *Conclusions*, one of the keys within our schema.

^b In our analysis we also included all Salutations and Closures, since they are formal elements of texts in these sub-genres.

References

Ghadessy M (1983): Information structure in Letters to the Editor. *International Review in Applied Linguistics* 21: 46-56.

Henry A, Roseberry RL (1998): *An Evaluation of a Genre-Based Approach to the Teaching of EAP/ESP Writing*. TESOL Quarterly (Spring Issue), Short Reports and Summaries.

Henry A, Roseberry RL (1996): A corpus-based investigation of the language and linguistic patterns of one genre and the implications for language teaching. *Research in the English* 30: 473-489.

Paltridge B (1997): *Genres, frames and writing in research settings*. Amsterdam: John Benjamins.

Swales J (1990): *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Vázquez y del Árbol E (2002): *Propuesta de un análisis comparado de cien textos biomédicos (español e inglés) desde la perspectiva del género*. Granada: Editorial Universidad de Granada (Dissertation).



¿Quién lo usó por vez primera?

Dardarina

Fernando A. Navarro

Cabrerizos (Salamanca, España)

Hace apenas unos meses, el 22 de octubre del 2004, la versión electrónica en línea de la prestigiosa revista *Neuron* traía una noticia sensacional: la clonación del gen responsable de una variedad de la enfermedad de Parkinson. El artículo, elaborado conjuntamente por tres grupos de científicos dirigidos por Jordi Pérez Tur (Instituto de Biomedicina de Valencia), Nick W. Wood (Instituto de Neurología de Londres) y Andrew B. Singleton (Instituto Nacional de Envejecimiento de los Estados Unidos, en Bethesda), contaba, entre sus 21 firmantes, con un nutrido plantel de investigadores españoles: Coro Paisán Ruiz, Javier Simón, Adolfo López de Munain, Silvia Aparicio, Ángel Martínez Gil, Javier Ruiz Martínez, Itxaso Martí Carrera, Amets Sáenz Peña y José Félix Martí Massó.

Mediante análisis genético de cuatro familias vascas y otra inglesa con antecedentes de una forma familiar de enfermedad de Parkinson, estos investigadores consiguieron identificar un gen mutante, el gen *PARK8*, que codifica una proteína de tipo cinásico y función aún desconocida.

Y si la noticia era en sí sensacional, más llamativo era aún, para quienes nos sentimos atraídos por el lenguaje médico, el nombre que los autores escogieron para la nueva proteína identificada. Llamativo, digo, no porque se hubieran inspirado en el síntoma más llamativo de la enfermedad de Parkinson, el temblor. De hecho, ya en 1817 el descubridor de esta enfermedad, James Parkinson, la bautizó como *shaking palsy* en inglés y *paralysis agitans* en latín (v. *Panace@*, n.º 2, pág. 56, <<http://tremedica.org/panacea.html>>), en virtud del característico temblor de los parkinsonianos. Lo llamativo e interesante del caso es, a mi modo de ver, que los autores del artículo de *Neuron* han recurrido ahora al nombre del temblor, no en inglés ni en latín, sino en vasco, para bautizar la proteína codificada por el gen mutante. De esta forma, la palabra vasca dardara, ‘temblor’, entra por la puerta grande en el lenguaje médico internacional a través del neologismo dardarin, ‘dardarina’.

We describe here the cloning of a novel gene that contains missense mutations segregating with *PARK8*-linked PD in five families from England and Spain. Because of the tremor observed in PD and because a number of the families are of Basque descent, we have named this protein dardarin, derived from the Basque word *dardara*, meaning tremor [Paisán Ruiz C y cols. Cloning of the gene containing mutations that cause *PARK8*-linked Parkinson’s disease. *Neuron* 2004; 44: 595-600].

Se trata de un excelente ejemplo demostrativo de que, en pleno siglo XXI, no sólo el inglés permite formar neologismos en esta era nuestra tan genotecnológica.