Variation across university genres in seven disciplines

A corpus-based study on academic written Spanish

Giovanni Parodi
Pontificia Universidad Católica de Valparaíso

This paper focuses on the identification of academic written genres from two corpora of reading material in Spanish that undergraduate and graduate university students are assigned in order to access specialized information across seven disciplines, namely Biotechnology, Chemistry, Economics, History, Linguistics, Literature and Physics. A group of 31 genres were identified across the discourses of the disciplines under study. Among the most frequent genres, “Textbook”, “Research Article”, and “Disciplinary Text”, were identified. Contrasting reading material, the importance of the “Research Article” is highly relevant, as it was the only genre that emerged in all seven disciplines. What I am proposing is that, regarding generic variation, there are clear-cut differences in the academic discourse of disciplines pertaining to Social Sciences and Humanities (SS&H) and Basic Sciences (BS), as it has been initially proposed (Parodi 2010a, 2014). However, the findings show that Economics academic discourse is closer to BS in terms of genre constitution.

Keywords: discourse genre, academic discourse, written language, discipline, university studies

1. Introduction

A prominent research area in discourse studies and corpus linguistics over the past twenty years has been the developments based on genre theory, applied not only to the description of language in use, but also to the pedagogical approaches of language teaching and learning in academic and professional contexts. Genre studies have moved from broad theoretical conceptualizations to empirical research. This has provided evidence that has reinforced some of the central concepts, while, at the same time, have also helped to reconceptualize others, including, for example,
disciplinarity: “By studying texts within their contexts, we study as well the dynamics of context building. In particular, by understanding texts within the professions, we understand how professions constitute themselves and carry out their work through texts” (Bazerman & Paradis 1991: 3).

One of the key issues that concern researchers in this field is the identification of genre variations across disciplines and languages, which has not received as much attention as required (e.g. Biber 1988, Martin & Veel 1998, Johns 2002, Paltridge 2002, Hyland & Bondi, 2006, Bazerman et al. 2009, Parodi 2010a). The research reported in this paper attempts to fill this gap. It is exploratory and descriptive in nature and examines academic discourse in Spanish. My interest lies in accounting for the written discourse genres that Chilean university students (undergraduate and graduate) employ as reading materials in order to access specialized information. Consequently, the objectives of the study are to identify, describe and quantify the academic discourse genres found in a corpus of 3,272 texts, collected from seven disciplines in the Social Sciences and Humanities (SS&H) and the Basic Sciences (BS): Biotechnology, Chemistry, Economics, History, Linguistics, Literature and Physics. In particular, the study seeks to examine the texts that university students encounter in their routinary academic activities in order to construct disciplinary knowledge and to describe the possible genre variations that emerge across a number of disciplines. It also compares the normalized frequency of occurrence of the identified discourse genres between SS&H and BS.

Three research questions arise at this point:

i. What are the similarities and variations across disciplines in the reading material that university students in Chilean universities deal with as routinary tasks, both at undergraduate and graduate levels?

ii. Are all discourse genres employed in the seven disciplines under study of a similar nature?

iii. What are the most frequent discourse genres that university students are assigned to read in SS&H and in BS?

The outline of the present paper is as follows: Section 2 starts with the theoretical framework of the study; Section 3 describes how the corpora were assembled and the methodology adopted. In Section 4, the most important findings are presented, focusing on the description of the genres identified in the seven disciplinary domains. In Section 5, conclusions, limitations and projections are offered.
2. Discourse genres as multidimensional constructions

‘Genre’ is a highly contested notion, sometimes contrasted and sometimes used synonymously with other related terms such as ‘text type’, ‘superstructure’ and ‘register’ (e.g. van Dijk 1977, Biber 1988, Martin 1992, Paltridge 2002, Egbert 2015). Considering all possible theoretical alternatives, the concept may be elusive, but the wide and rich discussions in the field open many possibilities on how to study discourse genres (e.g. Freedman & Medway 1994, Hyon 1996, Parodi 2010b, 2010c; Bazerman 1994, 2012; Martin & Rose 2008; Swales 1990, 2004). The different approaches and schools of thought differ in terms of emphasis given to the “[…] predictability and dynamism of the genre’s formal features, or the importance given to the wider social context” (Handford 2010: 258). Notwithstanding, Swales (2009, 2012) detects a tendency towards unification, and he proposes that some theoretical underpinnings could be closer to schools of thoughts that could have been perceived as opposed in the past. In this line, Bhatia (2004: 23) advances a definition of genre that highlights related concepts and prioritizes conventions, disciplinarity, and linguistic and discursive forms: “Genre essentially refers to language use in a conventionalised communicative setting in order to give expression to a specific set of communicative goals of a disciplinary or social institution, which give rise to stable structural forms by imposing constrains on the use of lexicogrammatical as well as discoursal resources”.

Complementing Swales’ (2012) and Bhatia’s (2012a, 2012b) approaches, and for the purposes of this study, I follow my own multidimensional conception of discourse genre (Parodi 2010c). In this approach, the emphasis is on the interactions between the linguistic, social and cognitive dimensions of discourse. According to Parodi (2010c), in the three-dimensional conception, genres are complex entities that capture the communicative purposes of speakers and writers in contextual and cognitively situated interactions by means of concrete texts that materialize meanings in construction. The linguistic dimension plays, then, a fundamental role by linking the contextual and the cognitive dimensions. Parodi (2010c) also adds that discourse entities called genres operate in specific social situations; achieve certain stability, due to their eminently cognitive nature; and are stored in the memory of speakers/listeners and readers/writers. Parodi (2010c) proposes that, in the case of academic discourse, texts collected in a corpus represent specific instances of social interaction and their study allows for the identification of their typical lexical and grammatical features. Moreover, the classification of texts according to genres facilitates the observation of patterns in the linguistic cartography of a specific genre (Hasan 1996, Matthiessen 1995, Halliday & Matthiessen 2006). In Parodi (2010c), I claim that genres have a linguistic dimension, which is recognizable like a cartographic map or an organizational pattern, which becomes the concrete
manifestation of a more abstract entity. This cartography, which includes not only patterns of lexicogrammatical features (Parodi 2005) but also the functional and rhetorical organization of texts in terms of, for instance, macro-moves and moves (Parodi 2010d, Ibáñez 2010, Burdiles 2012, Martínez 2012), becomes a key component for the understanding of disciplinary literacies in academic settings.

In university environments, genres have shown diversity across disciplines (e.g. Johns 2002, Hyland & Bondi 2006, Christie & Maton 2011, Gardner & Nesi 2012, Egbert 2015). Firstly, some of them are created specifically for university contexts while others are imported from more prototypically professional or scientific realms (Parodi 2010b, 2014, Bolívar & Parodi 2015). Secondly, some genres display didactic or informative recursive patterns (e.g. “Didactic Guide” and “Textbook”), while others contain specialized and dense prose (e.g. “Disciplinary Text”). Thirdly, some genres are read as a point of access to disciplinary information and are rarely written by university students (e.g. “Textbook”). Likewise, recent research has shown that prototypical genres that are written in graduate and undergraduate university settings are also relevant to professional scientific work, such as the “Critical Review”, the “Research Article” and the “Research Project” (Jarpa 2012).

The principles underlying Parodi’s (2010c) conceptual proposal may be summarized as a few core ideas:

i. Genres are articulated in a dynamic and complex manner from cognitive, linguistic and social dimensions;

ii. The relation among these three dimensions is not symmetrical, but rather an interaction in a dialectal cycle in which the possible differences appear only in some degree, since these dialectical interaction processes are inseparable;

iii. Most essentially, genres are cognitive constructs;

iv. Context is “out there”, but context is mainly considered as a cognitive artefact for this perspective of genre;

v. Genres are acquired socioconstructively, which means that the process of construction is through situated and distributed cognitions in a varying range of social and cultural settings; and,

vi. Genres may be acquired through formal and informal environments, but academic literacy (reading and writing) in academic and professional domains is an educational tool that may greatly help to facilitate the processes of genre construction.
3. Collecting the two corpora and describing the steps of the analysis

In Section 3.1, I first summarize the methodological steps followed to collect the two corpora. In Section 3.2, I then briefly examine the criteria and variables employed to classify texts into genres.

3.1 The PUCV-2010 Corpus and the PUCV-UCSC-2013 Corpus

To achieve the objectives and to answer the research questions, the present study uses two corpora. Although one of the critical restrictions of corpus construction is the consensus on that no corpus — no matter how carefully and ecologically designed or how large it is — can accurately represent and capture the language as a whole (Sinclair 2005, Biber 2005, Miller & Biber 2015), efforts should be made to ensure the most suitable data for the study. Based on this premise, the PUCV-2010 Corpus is a collection of written texts that are obligatory and complementary readings for students of six PhD programs in SS&H and BS (Parodi 2012a, 2012b). The acronym PUCV-2010 stands for the year in which the corpus was compiled (2010) and for the university that sponsored this research project: Pontificia Universidad Católica de Valparaíso (PUCV). The corpus was collected from texts used in 12 PhD programs (two per discipline) in six Chilean universities. The texts were selected by analyzing the syllabi of core and elective courses of each PhD program, except for the materials included in the PhD theses. The disciplines were selected in order to obtain a range of possible variations in written documents from different doctoral programs. Table 1 summarizes the features of texts compiled in the PUCV-2010 Corpus.

Table 1. Number of texts contained in the PUCV-2010 Corpus

<table>
<thead>
<tr>
<th>Areas and disciplines</th>
<th>Number of texts per discipline</th>
<th>Total number of texts per area</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS&amp;H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>History</td>
<td>592</td>
<td>2,603 (82.7%)</td>
</tr>
<tr>
<td>Linguistics</td>
<td>1,065</td>
<td></td>
</tr>
<tr>
<td>Literature</td>
<td>946</td>
<td></td>
</tr>
<tr>
<td>BS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biotechnology</td>
<td>227</td>
<td>544 (17.3%)</td>
</tr>
<tr>
<td>Chemistry</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>Physics</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3,147</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

As shown in Table 1, the PUCV-2010 Corpus is not only one of the most extended academic corpora in Spanish available today, but it is also one of the most diverse in terms the disciplines involved. There are no other corpora similar to this one in Spanish, or in any other language, compiled following these criteria and principles; that is, texts that are given as reading material to doctoral students and
which constitute the academic discourse of six disciplines. The procedure used to compile the corpus is summarized in the following steps:

i. Organization of a database containing full detailed information of the programs curricula (including their obligatory and optional course descriptions);
ii. Organization of a database containing all the obligatory and complementary bibliographic references included in the course descriptions;
iii. Elaboration of a survey submitted to professors of all programs, including a request for complementary course reading materials not included in the course descriptions that was given to be read by students;
iv. Collection of complementary course reading materials, including handouts, didactic guides, digital files and photocopied documents;
v. Internet search to find material already available in digital format, in order to reduce digitization time;
vi. Collection of print texts from libraries and professors’ offices;
vii. Process of photocopying each text in order to create a print-format data base for future reference;
viii. Compilation of all texts in order to have a complete digital data base of the total texts included in the corpus.

Following the same compilation protocol, the PUCV-UCSC-2013 Corpus on Economics in two undergraduate programs was collected (Parodi 2012a, 2012b). In this case, the acronym stands also for the corpus compilation year (2013), but in this case the study was sponsored by two universities: Pontificia Universidad Católica de Valparaíso (PUCV) and Universidad Católica de la Santísima Concepción (UCSC). The research team compiled all the texts that students were given as reading assignments, as part of the academic curriculum in Commercial Engineering and Accounting. The eight steps described were followed, and 124 texts were collected. The total number of texts in both corpora under study is 3,271.

3.2 Methodology: Identifying genres in the corpora

With the aim of identifying and classifying the texts of the two corpora, Parodi et al.’s (2010) methodological orientations to study genres was followed, which contends with what Cooper & Bikowski (2007: 208) have stated for “the difficulty in classifying writing assignments into neat, mutually exclusive categories”. Parodi et al.’s (2010) proposal, which aims at grouping together similar texts to form genre categories, is a useful methodological tool that makes the description of a large corpora more manageable and helps to categorize texts into mutually exclusive types of genres. However, they keep in mind the precaution that one key difficulty is the lack of consensus on criteria and category names as Gardner & Nesi (2013)
Variation across university genres in seven disciplines

point out. For their analysis, they follow Zhu (2004) and Graves et al. (2010), and classify a corpus of 2,858 texts from more than thirty academic disciplines into genre families, written by undergraduate and graduate university students in England. Thus, texts in different contexts may be classified differently. It is not easy or rather impossible, in Ganobcsik-Williams’s (2004) words, to know what differences, if any, exist between research term paper and scholarly article, for example, or between book review and book report.

In order to achieve an accurate distinction of genres in a corpus of texts, Parodi et al. (2010) propose a matrix with criteria and variables that combine deductive and inductive approaches. This means that in their work they operationalize conceptual categories that come from a combination of sources: (i) the theoretical framework, (ii) the researchers’ prior knowledge, and (iii) the empirical data that stem from the texts, in a rather ‘corpus-driven approach’ (Tognini-Bonelli 2001, Biber et al. 2007). The proposal is based on five criteria and twenty-three variables. The criteria are: (i) communicative macro-purpose, (ii) discourse organization mode, (iii) relationship between the participants, (iv) context of circulation, and (v) modality.

The five criteria are operationalized in a set of variables. These variables are analyzed in the texts under the principle of ‘predominance’ of one variable upon the others. This means that, from a multidimensional perspective, genres are neither homogeneous, nor stable units. But within any given genre several communicative macro-purposes coexist, and discourse organization modes vary. In keeping with the aforementioned ideas, it may be observed in Figure 1 that the matrix of criteria is left open to new possibilities in the blank spaces available in the corners. Figure 1 presents the variables established for each criterion.

In order to identifying the genres from the corpora, all texts were analyzed at least by three members of the research team (85% degree of agreement was reached), following the five criteria and the respective twenty-three variables. All texts were available in digital and paper format, so they could be easily studied as complete units. In general, the analysis revealed a tendency towards identifying the least number of genres possible; therefore, texts were grouped into family genres (Bhatia 2002), avoiding unnecessary proliferation and sub-specification. Complementarily, in order to validate the identifications that emerged, a set of examples of all genres were shown to two experts in each of the seven disciplines, once all the texts were assigned to one genre category. A high inter-rater reliability was reached among specialists (90% degree of agreement) using a Likert scale.

As a way to illustrate how, in part, the process of classifying texts into genres was conducted, three genre definitions are presented, based on the matrix of analysis. These definitions show how the criteria and variables interact and are specified for each of these three genres:
i. Textbook: Discourse genre whose communicative macro-purpose is to instruct students about concepts and/or procedures in a specialized topic. Ideally, its circulation context is the educational setting and the relationship between the participants is between expert writer and non-specialist readers. Preferably, discourse is organized in a descriptive mode and it integrates multimodal resources;

ii. Research Article: Discourse genre whose communicative macro-purpose is to persuade its audience about a given point of view, either by means of a review of the theoretical framework or the reporting of results of an empirical study. Ideally, its circulation context is the scientific setting and the relationship between the participants is between expert writer and expert reader. Preferably, discourse is organized in a preeminently argumentative mode and it incorporates multimodal resources;

iii. Disciplinary Text: Discourse genre whose communicative macro-purpose is to persuade about the treatment given to one or several topics within a particular discipline. Its circulation context is the scientific setting and the relationship between the participants is between expert writer and expert reader. Preferably, discourse is organized in an argumentative mode. It also uses multimodal resources.
Considering the difficulties in assigning one particular text to a genre category, genres were named following simple tagging rules, attempting, most of the time, to express their most common usage. When possible, the genre names used in previous studies were maintained (Swales 1990, 2004; Martin & Rose 2008; Parodi & Gramajo 2007; Gardner & Nasi 2013, among others), especially in those on which there was strong agreement, such as “Research Article”, “Textbook”, “Thesis”, and “Biography”. As for relatively new genres, they were named according to the traditions in each discourse community or by the names identified in the texts themselves (as for example, “Graphical Abstract Index”, “Technical Note”, and “Dictionary”).

4. Results and discussion

In Section 4, a step by step description of the genres identified as a result of the analysis is summarized. First, in Section 4.1, the description of the PUCV-2010 Corpus is presented, followed by the description of the PUCV-UCSC-2013 Corpus (Section 4.2). Second, in Section 4.3, a detailed analysis of the genres in each of the seven disciplines is presented and discussed.

4.1 Genre identification in the PUCV-2010 Corpus

In this section, the 30 genres that emerged from the analysis of the first corpus are presented. Table 2 provides the list of all genres identified.

<table>
<thead>
<tr>
<th>Table 2. Genres identified in the PUCV-2010 Corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PUCV-2010 Genres</strong></td>
</tr>
<tr>
<td>1 Anthology</td>
</tr>
<tr>
<td>2 Atlas</td>
</tr>
<tr>
<td>3 Bibliographic Reference</td>
</tr>
<tr>
<td>4 Biography</td>
</tr>
<tr>
<td>5 Museum Guide</td>
</tr>
<tr>
<td>6 Complete Works</td>
</tr>
<tr>
<td>7 Dictionary</td>
</tr>
<tr>
<td>8 Didactic Guide</td>
</tr>
<tr>
<td>9 Disciplinary Text</td>
</tr>
<tr>
<td>10 Dramatic Works</td>
</tr>
<tr>
<td>11 Editorial</td>
</tr>
<tr>
<td>12 Epistolary</td>
</tr>
<tr>
<td>13 Errata</td>
</tr>
<tr>
<td>14 Gospel</td>
</tr>
<tr>
<td>15 Graphical Abstract Index</td>
</tr>
</tbody>
</table>
The empirical findings of the first study show a wide range of genres. Some of the names may not sound familiar to the non-specialist, but there was consensus among the members of the respective discourse communities. Some of the genres identified clearly reveal in their names their academic and pedagogical nature, such as “Textbook”, “Didactic Guide” and “Thesis”. Others stand out due to their more professional approach, and others show their ideal contexts of production and circulation are scientific or workplace settings. Figure 2 shows the normalized occurrence of texts classified in each genre.

Figure 2. General quantification of the genres identified
Figure 2 provides relevant information regarding the academic corpus and the identified genres. Despite the ample heterogeneity of discourse genres found in this academic corpus of texts from six disciplines, it is evident that the discursive mechanisms most commonly used to communicate specialized knowledge at the PhD level are not so varied. Moreover, only two genres stand out: “Disciplinary Text”, with the highest occurrence of 57.23%, and “Research Article”, with 26.06%, followed by “Textbook” with 7.12%.

“Novel” and “Anthology” are the only two that present relatively important percentages in the fourth and fifth place, with 2.76% and 1.91%, respectively. It is worth noting the important presence of these two genres of clearly literary nature as part of academic discourse at the doctoral level. Their possible function within the corpus will be discussed later in a more detailed analysis of the genres of each disciplinary discourse. The other 25 genres identified in the PUCV-2010 Corpus do not reach an occurrence higher than 1%, with absolute numbers very close to zero. This indicates that some or many of these genres appear in just a few instances in the corpus of 3,147 texts.

The “Disciplinary Text” emerges as an important genre in the academic discourse of the disciplines under study, i.e. more than 700 texts in Linguistics, about 600 in Literature and more than 400 in History. These figures show that the “Disciplinary Text” is a fundamental genre in the transmission of specialized information and becomes an indispensable medium that enables the construction of disciplinary knowledge in the three disciplines of SS&H at the doctoral level. Likewise, the singular role of “Disciplinary Text” in SS&H disciplines corresponds with the findings reported by Parodi (2010a) in a corpus of undergraduate university texts in the disciplines of Psychology and Social Work. This genre is characterized by the elaboration, presentation and discussion of theories; and, to a lesser degree, “Disciplinary Text” also gives space to the communication of research findings. Frequently, the texts belonging to this genre are texts written by only one writer, but they may also be co-edited or co-authored editions. The most prototypical focus of “Disciplinary Text” is a theoretical argumentation, not necessarily supported by empirical data, in which one theory is described or competing and alternative theories are explained and discussed (Parodi 2010b, 2010d, 2014, Ibáñez 2010).

In summary, based on the figures reported in this study, there is no doubt that academic discourse in the disciplines under study is constructed by a limited range of genres. Most of the texts in the corpus belong to three of them, apparently very idiosyncratic and easily recognizable: “Disciplinary Text”, “Research Article” and “Textbook”. In other words, the findings reported here reveal that specialized knowledge across these six disciplines is built and transmitted mainly through clearly defined genres that focus on three types of discursive interactions:
i. Elaboration of theories, conceptual reflections, theoretical and methodological discussions, reporting of empirical findings, as in the “Disciplinary Text” (Ibáñez 2010);

ii. Communication of research results and empirical findings, as in the “Research Article” (Parodi 2010b); and,


4.2 Genre identification in the PUCV-UCSC-2013 Corpus

Following the same method of analysis, the PUCV-UCSC-2013 Corpus on Economics at the undergraduate level was studied from a genre perspective. The genre identification, based on the 125 texts, is presented in Table 3.

<table>
<thead>
<tr>
<th>PUCV-UCSC-2013 Genres</th>
<th>DG</th>
<th>RA</th>
<th>SG</th>
<th>TR</th>
<th>TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Didactic Guide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Research Article</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Study Guide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Technical Report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Textbook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This comparatively small corpus of undergraduate academic university discourse on Economics (Parodi & Burdiles 2013, Parodi et al. 2014) is relevant to our present research for two main reasons: on the one hand, it offers a complementary line of research on a discipline only recently studied by our research team on which there are no many published studies conducted on Spanish (e.g. Alcaraz et al. 2007, Martínez 2009, Stagnaro 2011, 2012), although it complements previous corpora collected by our research team on other undergraduate academic discourses (Parodi 2007, 2010a, 2014). On the other, Economics has been a discipline under debate on whether its discursive nature is closer to SS&H language or to BS language (McCloskey 1983, 1985, 1994, Henderson et al. 1993, Bondi 1996a, 1996b, Mosini 2011). This interesting controversy is studied here in terms of written discourse genres, and the analysis may contribute with information from Spanish to the kind of disciplinary discourse involved.

The data presented in Table 3 shows five genres identified. Three of them are of a pedagogical nature (“Didactic Guide”, “Study Guide”, and “Textbook”) and clearly reveal the academic focus of interest of the writer as a teacher responsible for choosing the accessing discursive mechanisms to the construction of specialized
disciplinary knowledge. In these three genres, the role played by the writer and the reader show the prominence given to the interaction between the participants, particularly to the way the writer/teacher guides the reader as a student by providing a cartography of the disciplinary knowledge carefully organized step by step. In these genres, the information is organized in separated units and presented gradually, employing a variety of didactic resources (Bondi 1996a, 1996b, Parodi 2010d, Parodi et al. 2014). Complementarily, the other two genres (“Research Article” and “Technical Report”) are closer to workplace interactions and show the necessity to establish links with the future professional discourse environment the students may encounter in the future. In “Research Article” and “Technical Report”, disciplinary knowledge is not organized with pedagogical purposes of teaching new concepts, but they communicate contemporary scientific information.

4.3 Discourse genres in each of the seven disciplines under study

In Sections 4.3.1, 4.3.2 and 4.3.3 below, I summarize and discuss the results of the analysis of the texts in each of the seven disciplines, which was conducted to examine the genres identified and their frequency of occurrence. Additionally, in Section 4.4, comparisons and implications between SS&H and BS discourses are advanced.

4.3.1 Genres in BS

In this section, I focus now on the results for the academic discourse of the three disciplines in BS at the doctoral level: Chemistry, Physics and Biotechnology. As Figure 3 shows, five genres emerged from the 136 texts on Chemistry. In order of occurrence these are: “Research Article”, “Textbook”, “Disciplinary Text”, “Graphical Abstract Index” and “Review”. It is evident that “Research Article”, “Textbook” and “Disciplinary Text” are the most prototypical and fundamental
genres in the discipline while the other two genres are far less frequent (0.74%) and play a rather marginal role.

The results of genre identification show that highly specialized informative genres, such as “Research Article” and “Disciplinary Text”, are involved in the transmission of knowledge together with other means of a more educational and instructional function (“Textbook”) that apply informative and reformulation strategies. Regarding the latter genre, Parodi (2010d, 2011, 2014) describes its main function in undergraduate disciplines such as Industrial Chemistry and Construction Engineering in the teaching of concepts and procedures by means of exemplification and problem solving in highly illustrative formats. The data presented here for the doctoral level point towards the combination of genres and a tendency to a high specialization, but with an important presence of textbooks, that act as initial written tools that pave the way to new information. Now Figure 4 gives account of the genres found as part of the discourse in Physics.

![Figure 4. Genres in Physics](image)

Similar to Chemistry, five genres were identified in 181 texts of the corpus (Figure 4). The three most frequent genres in Physics were “Textbook” (70.72%), “Disciplinary Text” (24.86%) and “Research Article” (3.31%). The other two genres, “Proceedings” and “Thesis”, reached barely 0.55%. The relationship between the participants can be defined as between experts, just as in the “Research Article” and “Disciplinary Text”. However, it is the “Textbook” genre the one that stands out as the widespread discursive tool in the access to disciplinary contents, through which didactic strategies are deployed, e.g. definition of specialized concepts, exemplification, classification and presentation of problems with their corresponding solution and modeling. In this genre, the type of relationship highlighted is between expert and non-specialist. This means that 70.72% of the texts from the discipline of Physics at the doctoral level belong to a genre that provides readers with a framework that progressively scaffolds access to information and
enables the construction of disciplinary knowledge in a very didactic manner and with the aim of instructing audiences under training. Following with the PUCV-2010 corpus analysis, 14 genres were identified in the disciplinary discourse of Biotechnology (Figure 5).

Figure 5. Genres in Biotechnology

As can be observed, the number of genres in the academic discourse of Biotechnology is considerably higher than in the other two disciplines in BS. Despite the greater diversity of genres identified in the corpus of 227 texts in this disciplinary discourse, as compared to Chemistry and Physics, there are three genres that occur most frequently in the three disciplines: “Research Article” (63.39%), “Disciplinary Text” (9.82%) and “Textbook” (9.38%). The other 11 genres show very low presence.

Unlike the disciplinary discourse of Physics, the hierarchical order of these three genres in Biotechnology is different in terms of pedagogical discursive devices. In this case, “Research Article” is the most relevant genre, the one that presumably gives access to specialized information and scaffolds the construction of knowledge within the discipline. As it is well-known, this genre circulates information among experts, communicates state-of-the-art knowledge and disseminates scientific findings of theoretical and empirical nature. The communicative macro-purpose is to persuade by means of diverse argumentative strategies. Together with “Research Article”, “Disciplinary Text” is the second specialized genre that also establishes an expert relationship between writer and reader. As for “Textbook”, it also holds a relevant place in the discourse of Biotechnology, although it ranks third in terms of occurrence (9.38%). This relevance shows that
different types of genres of different didactic degree are combined, in attention to, in this case, semi-specialized audiences.

The relative heterogeneity of genres in the discourse of this discipline opens a scenario that displays genres such as “Technical Note”, “Technical Report”, “Review”, “Didactic Guide” and “Dictionary”. All of these have an occurrence of more than 1%. This wide diversity of genres in Biotechnology (14 in total) makes this discipline the second most heterogenic in the PUCV-2010 Corpus.

4.3.2 Genres in SS&H

In this subsection, an account of the genres identified in the three disciplines in the SS&H field at the doctoral level is presented. First of all, I will review the genres found in the academic discourse of Linguistics. The normalized frequencies show a clear concentration on only few academic genres (Figure 6).

The analysis of the 1,065 texts of Linguistics discourse resulted in the identification of seven genres. From these seven genres, only two have relevant occurrence: “Disciplinary Text” (71.46%) and “Research Article” (23%). Both genres are prototypical of the transmission of highly specialized knowledge. The rare occurrence of the “Textbook” genre (2.63%), unlike the disciplinary discourses in BS, shows that this is not a prototypical discourse genre in this field of knowledge. Previous studies conducted on the PUCV-2006 Corpus had already highlighted this difference between academic discourse in undergraduate SS&H and in BS and Engineering programs (Parodi 2010d, 2011, 2014; Ibáñez 2010). Now the empirical data hereby presented shows that this difference is also found in academic discourse at the doctoral level. “Disciplinary Text” is as relevant to SS&H as “Textbook” is to BS. Likewise, although not similarly high in occurrence, “Research Article” becomes
the second most important genre in doctoral training discourse in Linguistics. Clearly, in this discipline, knowledge is transmitted through books, followed in percentage of occurrence by research articles published in scientific journals, both in paper and electronic format.

Figure 7 illustrates the wide variety of genres found in the academic discourse of History. 13 genres were identified, where “Disciplinary Text” accounts for 72.64% of the occurrences.

From the 13 genres found in the 502 texts collected in this discipline, only two show relevant occurrences: “Disciplinary Text” (72.64%) and “Research Article” (21.62%). From the remaining 11 genres, only two are above 1% of occurrence: “Atlas” and “Dictionary”. Although an important generic variation is found in the academic discourse of doctoral studies in History, its composition in terms of occurrence is very similar to that observed in Linguistics, except for a considerable lesser heterogeneity. Likewise, academic discourse in Linguistics and History reveals that both disciplines share similar discourse genres for accessing disciplinary information. Notwithstanding, it is necessary to conduct a more detailed study of the texts that integrate these common disciplinary genres. The categorization within a genre does not necessarily mean that such texts are constituted identically or that the functional and rhetoric organization plays the same role. In this regard, Parodi (2010d, 2011, 2014) and Ibáñez (2010) show that, in an analysis of rhetorical moves, “Textbook” and “Disciplinary Text” genres vary in their constitution in the discourse across disciplines, especially between SS&H and BS.
In the discipline of History, the genres identified as “Epistolary”, “Novel” and “Poem” are also worth mentioning. The presence of these genres tells us of two possible types of discourse within the corpus: one genre type that serves the purpose of presenting the analysis of text contents, and other genres that are the object of such analysis. This double articulation reveals the discursive nature of the discipline. This means that, on the one side, some genres serve as the object of study and research. On the other side, genres give an account of such analysis. In other words, there are genres that are the source of information or study matter and others that, in part, are based on the contributions of such source genres. Finally, I present the analysis of the 946 texts that integrated the corpus of Literature, in which 15 genres were identified.

![Figure 8. Genres in Literature](image)

The study of this disciplinary corpus revealed that the academic discourse in Literature is the one with the highest generic diversity. In terms of occurrence, and similar to findings in Linguistics and History, two genres emerged as the most frequent: “Disciplinary Text” (54.55%) and “Research Article” (25.16%). In summary, the results of this study show that the discourse of the three disciplines in SS&H shares a common feature regarding the high specialization of texts and the use of genres where the relationship between writer and reader is from expert to expert. As seen in the data, the academic discourse at the PhD level in Literature shows zero occurrences of genres of pedagogical nature or that display didactic strategies to teach and guide the reader/student. No instances of “Textbook” or “Didactic Guide” were found in the discourse of this discipline. Additionally, the
identification of “Textbook” or “Didactic Guide” in the academic discourse of History and Linguistics is inexistent or of a very low occurrence.

Among the genres identified in the corpus of Literature, there are literary genres such as “Anthology”, “Novel”, “Story”, which are different in nature to the “Disciplinary Text” and the “Research Article”. In order to further explore these distinctions, the texts from this discipline were reclassified in those that belong to a literary or non-literary genre. Figure 9 shows the number of occurrence in percentages.

![Literary / Non-Literary Genres](image)

Figure 9. Global percentage of texts in literary and non-literary genres

As Figure 9 shows, 81% of genres identified in this disciplinary corpus are considered as non-literary. These genres present descriptions, discussions or reflections about theories, models or concepts and analyze texts that belong to genres denominated as literary. The remaining 19% of genres identified for this discipline are constituted by texts considered as literary. This means that these genres that integrate the academic discourse in Literature constitute the matter or focus of study and research in these doctoral programs. This distinction turned out to be quite relevant as it illustrates a particular feature of specialized academic discourse in Literature, in which two types of genres are clearly set apart. The generic constitution of the corpus in Literature reveals a special feature that was also present, although more slightly, in the corpus of History. Thus, there are academic genres used to communicate the result of reflective and analytical activities (“Disciplinary Text”, “Research Article”, “Thesis”, “Review”, among others) upon other genres (“Novel”, “Poem”, “Story”, “Complete Works”, etc.), which are the raw material from which the other genres emerge. In order to further explore this distinction, texts were reorganized in two macro-genres. The data is graphically displayed in Figure 10.
“Novel” (47.98%) and “Anthology” (34.68%) are the most frequently found literary genres in the corpus, followed by “Poem” and “Story”. This means that, at the doctoral level, narrative genres are preferred as objects of study and analysis, followed by poetry. Thus, the analytical work and research in Literature is carried out mainly through “Disciplinary Text” (66.75%) and “Research Article” (30.79%). Both genres show a high degree of specialization and scarce techniques of dissemination to less specialized audiences. Other non-literary genres identified in this corpus show a very scarce occurrence, in which only one of them (“Proceedings”) reaches 1% of frequency and also shares features of high specialization with the “Disciplinary Text” and “Textbook”.

**Figure 10.** Comparison of literary and non-literary genres
4.3.3 Genres in Economics

As part of the second corpora analysis, the 125 texts (PUCV-UCSC-2013 Corpus), collected in Economics courses in two undergraduate programs, were analyzed in quantitative terms. The normalized frequencies show concentration on two genres.

Figure 11. Genres in Economics

As can be observed in Figure 11, the two most frequent genres in this second corpus are “Study Guide” (45.6%) and “Textbook” (36%). The remaining three genres are all below 10% of occurrence: “Research Article” (9.60%), “Technical Report” (6.40%) and “Didactic Guide” (2.40%). This relative generic variation found in the academic discourse of Economics is highly informative, and it offers new data not similar to the findings reported on previous studies conducted on undergraduate genre identification in academic discourse of SS&H (neither for English: Klamer 1990; Henderson et al. 1993; Bondi 1996a, 1996b; Goldschmidt & Szmrecsanyi 2007; nor for Spanish: Alcaraz et al. 2007; Martínez 2009; Parodi 2010d, 2011, 2014; Ibáñez 2010). The fundamental difference is that this generic composition, in which the pedagogical genres (“Study Guide” and “Textbook”) are the most relevant in frequency of occurrence, is closer to the discourse identified in disciplines of BS and not of the group of SS&H. Additionally, the academic discourses of Linguistics, Literature and History (as shown in this paper for graduate studies), and Psychology and Social Work (for undergraduate studies, Parodi 2010d, 2014) have also displayed a considerable variety of genres, while the academic discourse of Economics is now presenting –comparatively- a lesser heterogeneity, similar to what happens in most BS. At the same time, the comparisons between graduate and undergraduate written genres show similarities in the discourse resources identified; that is, no important differences in terms of new genre types were identified in this corpus. Most of the frequent genres show a focus on didactic resources (“Didactic Guide”, “Study Guide”, “Textbook”) or a high specialized informative prose (“Research Article” and “Technical Report”). Therefore, in spite
of the fact that the two corpora under comparison are from different university levels, they are complementary in the cartography they all help to construct across the disciplines.

The similarity in genre composition of the discourse of Economics with those of BS (focus on “Textbook” instead of “Disciplinary Text”) adds new information to genre studies and research on disciplinarity. Also, it reveals that not all academic discourses show similar language resources in supporting knowledge construction and that genres are not as homogeneous as they may be expected, even when being part of one disciplinary domain as part of the SS&H. Additionally, it opens opportunities for detailed analysis of the academic discourse of Economics such as the kind that Parodi & Burdiles (2013), Boudon & Parodi (2014), Parodi et al. (2014) and Parodi & Burdiles (2015) are currently conducting on textbooks and other genres.

4.4 Genre comparisons across disciplines

Figure 12 depicts how these genres from the two corpora (PUCV-2010 Corpus and PUCV-UCSC-2013 Corpus) are found in each of the specialized domains. The Venn diagram illustrates the distribution of the 31 discourse genres identified in the academic discourse of seven disciplines and highlights those that are shared by two or more disciplines and those that belong exclusively to one or another specialized domain.

Based on these data, RA stands out as the only genre that occurs in every one of the seven disciplines of the corpus, both at undergraduate and graduate levels. This finding suggests that research articles published in disciplinary journals is the genre that stands as the most relevant means of information distribution and serves the purpose of communicating disciplinary updated knowledge. The importance of the RA also shows that in these seven disciplines, regardless of generic heterogeneity and variation, specialized knowledge is partially communicated through a relatively similar discursive resource. Additionally, this outcome reveals that both writers and readers from undergraduate and doctoral programs in SS&H and BS construct and employ genres that share similar discourse organization when they want to communicate theories and empirical findings. Nevertheless, considering the number of “Research Article” identified in both corpora, the doctoral level shows the most important concentration of “Research Article”, particularly in Chemistry and Biotechnology (both from BS); therefore, as well-known, genres are not only discipline-bound but, as the findings of this study show, they would also be level-bound (undergraduate or graduate).

Focusing on the academic discourse of Physics, it is worth noting that it does not contain any prototypical genre. From the five genres identified in the discourse
of this discipline, none is exclusive but all are shared with three, four, five or seven of the other disciplines. In the discourse of Linguistics, Economics, and Chemistry, each one has an exclusive genre. The diagram also shows that the discourses of Biotechnology (8), Literature (7), and History (4) have an important number of idiosyncratic genres, although many of them do not have high frequency of oc-
currences. Based on the data collected, Figure 13 shows the distribution of the disciplinary academic discourse genres from the least diverse to the most diverse.

**Figure 13. Diversity continuum of genres across disciplines**

As this figure illustrates, the academic discourses in Economics, Physics and Chemistry have the least diversity in genres (5), while Literature discourse is the most diverse discipline (15), followed by the academic discourse in Biotechnology (14). Along the same lines, History follows with 13 different genres. Linguistics locates at the centre of this continuum with a total of seven genres. This analysis reveals that, regarding generic variation, there are clear-cut differences in the academic discourse of disciplines in SS&H and BS. While the academic discourse in Linguistics is closer to the one in Economics, Physics and Chemistry in terms of generic variation, the discourse in Biotechnology is as varied as the specialized discourse of History and Literature. Considering this distribution, Economics undergraduate discourse reveals interesting features in terms of genre diversity, closer to BS discourse; in other words, the important presence of textbooks and study guides reveals the way in which this discipline constructs knowledge. This process of knowledge construction through genres is performed in similar ways as Chemistry and Physics do and not in the way in which in other SS&H discourses preferably occurs, including Psychology and Social Work discourses (Bondi 1996a, 1996b; Parodi 2010b, 2011, 2014; Maci & Sala 2012).

What the present study and previous research have shown is that in SS&H discourses the most common genres employed in university knowledge construction is the “Disciplinary Text”, while in BS discourses the “Textbook” is the most relevant genre. This is why the finding concerning Economics discourse is so revealing and it is compatible with what McCloskey (1983, 1985, 1994), Klamer (1990) and Balak (2006), among others, have noted about the way economists would prefer to present Economics discourse, that is closer to Physics, Chemistry and Biology discourses. In a detailed analysis of the rhetorical and multimodal resources employed in textbooks in Economics, Parodi et al. (2014) and Boudon
& Parodi (2014) have proposed that this genre in Economics discourse could show features of a ‘discipline in transit’, thus highlighting the tension between a disciplinary discourse that attempts to show mathematical and graphical resources as Physics but still have the persuasive, narrative and open-ended alternative solutions, prototypical of the SS&H discourse.

Examining in detail the data provided by Figure 12, it is also worth restating that the disciplinary discourse of Literature is the only one that presents two macro-genres: literary and non-literary genres. Although this distinction is also present in the academic discourse at the doctoral level in History, it is not similar under the same premises. In this regard, it may be possible to consider the genres in the literary macro-genre as non specialized and general. In many cases, as Parodi (2007) has pointed out, texts belonging to this literary macro-genre show prototypical features (among others, first and third person singular and plural pronouns, second person singular pronouns; periphrastic forms of future tense, imperfect tense; and time and negation adverbs). These lexicogrammatical features have been associated, by opposition, to a lesser degree of complexity and informational density, typical of literary oral and written texts and more distant to specialized technical texts (Parodi 2007, Egbert 2015).

5. Conclusions

It is hoped that this corpus-based study, combining two academic corpora from undergraduate and graduate university settings, has contributed valuable insights into the variation of written academic discourse genres across disciplines. The findings presented here have helped to corroborate the powerful explanatory value of the discipline in order to predict genre variation. The results of this study strongly support the distinction between SS&H genres and BS genres, although Economics discourse seems to stand in between.

As part of the findings, 31 discourse genres in a corpus of 3,272 texts of academic written discourse in seven disciplines were identified. The findings show a high generic variation within the corpora. However, the quantification of texts pertaining to each genre showed a high frequency of occurrence of a few prototypical genres of academic and disciplinary discourse at undergraduate and graduate levels. The most frequent and fundamental genres from this quantitative perspective are: “Disciplinary Text”, “Research Article”, “Textbook”, and “Study Guide”. Conversely, from a cross-disciplinary perspective, “Research Article” was identified as the only genre that is present in the discourse of these seven disciplines (Biotechnology, Chemistry, Economics, History, Linguistics, Literature and Physics).
Another important contribution of the present study is that it offers data that do not follow the more traditional and significant strong generic distinction, previously identified between the academic discourse of SS&H and the BS at undergraduate level (Parodi 2011, 2014; Ibáñez 2010), in which “Disciplinary Text” was the most prototypical of the SS&H and “Textbook” was the most relevant for BS. In this vein, the more classical diversity continuum and distinction of genres across disciplines have now been altered: the discourse of Economics reveals features that will require deeper analysis because it is supposedly classified as a SS&H discipline, although the most frequent genre is the “Textbook”, a genre that is most prototypical of BS. As well-known, textbooks play a crucial role in the process of knowledge construction by providing a forum in which new concepts are taught and exercises are solved in a step-by-step manner, typical of Chemistry and Physics discourses.

Additionally, the double articulation of genres observed in Literature and History was another relevant finding. The written discourse in Literature presents literary and non-literary macro-genres. Both disciplines have in common that one macro-genre addresses the study and analysis of the other macro-genre. What this finding shows is the primarily discursive nature of research in Literature and History, as compared to a more experimental approach in the other disciplines included in the corpora under analysis.

Despite all of the above, it is also reasonable to expect that these written genres be not homogenous across disciplines and showed particular idiosyncratic and prototypical features according to their discipline of origin. This means that, while in the corpora four specific genres (“Textbook”, “Research Article”, “Disciplinary Text” and “Study Guide”) were identified as important discourse mechanisms in the construction of disciplinary knowledge, they could show, for example, rhetorical move organization differences or variation on the predominance of a more theoretical or a more empirical approach or differences in their lexicogrammatical constitution. Parodi (2010d, 2011, 2014), Ibáñez (2010) and Martínez (2012) have shown that textbooks, disciplinary texts and doctoral theses show idiosyncratic features across disciplines. On the other hand, Burdiles (2012) studies a genre that showed sub-discipline variation (the “Clinic Case” in Medicine). This hypothesis, for the two corpora under study in this paper, should be tested in future research.

It can be inferred then that the data obtained in this study offer a wide variety of projections and open the door for more complementary studies. These findings pave the way for planning and creating ‘reading across disciplines’ (RAD) programs from a perspective in which psycholinguistics be informed by corpus linguistics. At the same time, replication of the design used in this study with other disciplines and other languages would reveal whether these findings may have cross linguistic validity as they are extended to new horizons.
Finally, it is essential to recognize theoretical and empirical limitations to a study such as this one. The matrix employed to identify genres should include more criteria and variables so as to enrich the analysis of each text and capture the broad scope of the written material under study. Additionally, the task of grouping and classifying texts and including them into genre categories should face new complex challenges. In some areas, these processes could be relatively easy, particularly in which the communicative purpose and the relationship between participants were accurately revealed through section headings or were explicitly stated in the texts. In other areas, it was more difficult to identify predominant elements, such as the discourse organization mode. Nevertheless, the corpus is indicative of actual reading practices across these disciplines and the two levels of study.

Acknowledgements

This research has been funded by FONDECYT 1130033 and 1090030

References


DOI: 10.1075/scl.28


DOI: 10.1016/j.jeap.2007.09.008


DOI: 10.1177/0741088310371635


**Author’s address**

Giovanni Parodi
Instituto de Literatura y Ciencias del Lenguaje
Pontificia Universidad Católica de Valparaíso
Av. El Bosque 1290
Viña del Mar, 2530388
Chile

gparodi@ucv.cl