UNIVERSIDADE FEDERAL DE SANTA CATARINA

PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA CORRESPONDENTE

TACKLING GENRE CLASSIFICATION:
THE CASE OF HTML RESEARCH ARTICLES

GRACIELA RABUSKE HENDGES

Tese submetida à Universidade Federal de Santa Catarina em cumprimento
parcial dos requisitos para obtenção do grau de

DOUTOR EM LETRAS

FLORIANÓPOLIS

Dezembro 2007

Livros Grátis

Esta tese de Graciela Rabuske Hendges, intitulada "Tackling genre classification: the case of HTML research articles", foi julgada adequada e aprovada em sua forma final pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção do grau de

DOUTOR EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
Opção: Língua Inglesa e Lingüística Aplicada ao Inglês

_____
Dr. José Luiz Meurer
Coordenador

BANCA EXAMINADORA:

_____
Dr. José Luiz Meurer
Orientador e Presidente

_____
Dra. Barbara Jane Wilcox Hemais
Examinador

_____
Dra. Désirée Motta-Roth
Examinador

_____
Dra. Lêda Maria Braga Tomitch
Examinador

_____
Dra. Vera Lúcia Menezes de Oliveira e Paiva
Examinador

Florianópolis, 20 de dezembro de 2007.

For my parents, Celíria and Leoberto,
and for Vinícius.

ACKNOWLEDGEMENTS

ABSTRACT

TACKLING GENRE CLASSIFICATION:
THE CASE OF HTML RESEARCH ARTICLES

GRACIELA RABUSKE HENDGES

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2007

Supervising Professor: José Luiz Meurer, PhD

Recent studies on scientific communication have revealed that since the late 1990's there has been a shift in journal usage from print journals to online journals (Tenopir, 2002, 2003; Tenopir & King, 2001, 2002) and, as a consequence, it has been predicted that by 2010, 80% of journals will have stopped paper publication (Harnad, 1998). Nevertheless, these studies have also shown that not all disciplines are 'moving' towards the Internet with the same pace. While researchers in the fields of information science, librarianship, electronic publishing, document design, medicine have shown their concern regarding the impact of this major shift in academic exchange, considerations from the perspective of applied linguistics, particularly genre analysis have been less common. In this study, therefore, I explore to what extent and how the (new) medium of the Internet has been affecting the genre research article in its move from the print medium to the digital medium. More specifically, research articles in HTML format from the disciplines of Linguistics and Medicine will be examined. The results will be contrasted with results obtained for an online questionnaire distributed to researchers and professional in the two areas. The methodological approach adopted in this study derived from Askehave and Swales (2001) and Swales (1990, 2004), who claim that communicative purpose is a privileged criteria for genre classification, but it can only be achieved after both textual and contextual analysis is developed. The results from both contextual and textual data analysis show that the HTML research article is a new genre whose communicative purpose is realized by hyperlinks and therefore, it is a genre which is profoundly dependant on the medium.

Key-words: Genre classification – HTML research articles – Medium

Nº de páginas: 191
Nº de palavras: 43.038

RESUMO

CLASSIFICANDO GÊNEROS:
O CASO DO ARTIGO ACADÊMICO EM HTML

GRACIELA RABUSKE HENDGES

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2007

Orientador: Dr. José Luiz Meurer

Pesquisas recentes sobre comunicação científica têm revelado que desde o final dos anos de 1990 o uso de periódicos acadêmicos passou da mídia impressa para o mídia eletrônica (Tenopir, 2002, 2003; Tenopir & King, 2001, 2002) e, conseqüentemente, há previsões de que por volta de 2010 cerca de 80% dos periódicos terão apenas versões online (Harnad, 1998). Todavia, essas pesquisas mostram também que nem todas as disciplinas estão migrando para a Internet com a mesma velocidade. Enquanto que áreas como as Ciências da Informação, Arquivologia, Web design e Medicina têm mostrado interesse e preocupação em entnder e explicar esse fenômeno, em Lingüística Aplicada, particularmente em Análise de Gênero, os estudos ainda são escassos. Neste trabalho, portanto, procuro investigar em que medida o meio eletrônico (Internet) afeta o gênero artigo acadêmico no seu processo de mudança da mídia impressa para a mídia eletrônica. Mais especificamente, examino artigos acadêmicos em HTML nas áreas de Lingüística e Medicina com vistas a verificar se esse hypertexto é um gênero novo ou não. A abordagem metodológica adotada nesta pesquisa deriva da proposta de Askehave e Swales (2001) e de Swales (2004), na qual o critéro predominante para a classificação de um gênero é o *propósito comunicativo*, o qual só pode ser definido com base em uma análise textual tanto quanto em uma análise contextual. Dessa forma, neste estudo foram coletados e analisados dados textuais e contextuais e os resultados de ambas análises revelam que o artigo acadêmico em HTML é um gênero novo, cujo propósito comunicativo é realizado por hiperlinks e portanto, esse gênero é profundamente dependente da mídia eletrônica.

Palavras-chave: Classificação de Gêneros – Artigo Acadêmico em HTML – Mídia Eletrônica

N^o de páginas: 191
N^o de palavras: 43.038

TABLES OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

Chapter 1   Introduction

1.3 Relevance of the study

The language of science has called the attention of researchers from a variety of areas for over a century. As far as Applied Linguistics is concerned, studies about scientific discourse flourished after Swales's (1990) original and notable work. Although other academic genres have been studied – abstracts (Bittencourt, 1995, 1996; Motta-Roth & Hendges, 1996, 1998; Swales, 1996; Hartley & Sydes, 1997; Biasi-Rodrigues, 1998; Cross & Oppenheim, 2006); book reviews (Motta-Roth, 1995, 1998; Araújo, 1996); thesis and dissertations (Paltridge, 2002; Araújo, 2006; Starfield & Ravelli, 2006); conference presentations (Rowley-Jolivet, 2002, 2004; Rowley-Jolivet & Carter-Thomas, 2005); posters (Swales & Feak, 2000; MacIntosh-Murray, 2007) – the print research article has occupied a privileged position in genre analysis. The quintessential reason for this prestige seems to be the role the research article plays in scientific communication, where, apart from being "undoubtedly the most frequent scientific genre" (Salager-Meyer, 1999, p. 292)[1], it is the genre with

---

[1] Salager-Meyer's assertion is more likely to be true for the fields of Medicine and Physics, for example, than for History and Linguistics, where books still play an important role, absorbing over 50% of the total amount of publications. This issue will be considered in more detail later on in Chapter 5.

highest value when it comes to gathering research funding as well as for career upgrade.

Needless to say, the majority of the examinations of research articles has focused on texts written in English, for English continues to be the *lingua franca* of scientific communication since the 1980's (Swales cited in Motta-Roth, 1995; Montgomery, 2004)[2].

Depending on the level of the analysis – i.e., whether the analysis foregrounds macrolevel features (generic features such as schematic structure) or microlevel features (registerial aspects, i.e., lexico-grammatical patterns) – studies about the research article (RA henceforth) can be classified into three groups: 1) macrolevel-centered; 2) microlevel-centered and 3) studies which consider the interplay between macro and micro level patterns. To the first group belong those investigations that consider aspects of the genre's schematic structure as a whole (e.g. Meurer, 1998; Motta-Roth & Hendges, 1996, 1998, 1999; Posteguillo, 1999; Yang & Allison, 2004) or the structural organization of specific sections (e.g. Hopkins & Dudley-Evans, 1988; Brett, 1994; Holmes, 1997; Anthony, 1999; Silva, 1999; Samraj, 2002; Ozturk, 2007). The second groups encompasses research about the citation patterns in RAs (Swales, 1986; Salager-Meyer, 1999; Hemais, 2001), hedging (Salager-Meyer, 1994; Varttala, 1999; Vassileva, 2001), evaluation (Thetela, 1997), verb types (Thompson & Yiyun, 1991; Thomas & Hawes, 1994; Williams, 1996). One example of the third group – studies which explore the interplay between macro and microlevel patterns – is the study by Swales et al.

---

[2] According to Graddol (2006, p. 62), in other contexts but the scientific, English is no longer the "only show in town".

(1998), which investigates the role of imperatives in RAs in relation to the specific sections where they occur. Within each one of the three groups we can find intra-disciplinary as well as inter-disciplinary examinations.

Even though genre studies from the 1970's until the 1990's have been described as being more concerned with the texts than with the contexts in which texts occur (Bhatia, 2004), it is beyond doubt that both the macro and microlevel discussions about the RA above mentioned have offered important insights into our understanding of the relationship between context and text within the scope of scientific discourse. In addition, by examining and explaining how scientists communicate in general as well as how specific disciplinary communities engage in science, these studies have also been decisive in the development of academic literacy materials and courses.

Nevertheless, so far the vast majority of the research on RAs has tended to focus solely on the verbal mode, with a few recent exceptions (Miller, 1998; Roth et al., 1999; Roth & Welzel, 2001; Nascimento, 2002; Nascimento, Motta-Roth & Hendges, 2003). This is not surprising given that essentially, as applied linguists, our primary object of study is the verbal mode of language (while, for example, paintings, drawings and photographs have traditionally been the primary object of the Visual Arts; equations and statistical graphs, of Mathematics; sound, of Music). Another reason that could explain why RA analysts have concentrated attention on the verbal mode in this genre is that it is seen as an essentially *monomodal* genre, i.e. a genre in which one mode carries all the meaning (the written-verbal mode in the case of the RA). Such a view can be illustrated through the definition of the RA

offered in 1990 by Swales, in which the author states that the research article "is taken to be a written text (although often containing non-verbal elements)"[3] (p. 93).

Nevertheless, the reign of logocentrism, of paper and of the written-verbal mode of language – promoted especially by Gutenberg's invention of the printing press – came to an end when changes in our "semiotic landscape" (Kress & van Leeuwen, 1996, p.33) started to take place. The *semiotic landscape* of a given society is, "on the one hand, the range of forms and modes of public communication available in that society, and, on the other hand, their uses and valuations" (Ibid., p. 33). As far as the range of forms of public communication is concerned, cinema, television, computers and the Internet represent new communication alternatives in contemporary society, all of which have significantly contributed for an increase in the importance of non-verbal language. These media have been regarded as responsible for giving birth to a *metaevent* called *visual turn* or "pictorial turn" (Mitchell, 1995, p. 13), a movement which has been particularly "accelerated in the last few years due to the explosion in the use of the Internet" (Thornes, 2004, p. 787).

With the *visual turn* (words giving way to images), the disciplinary pluralism or interdisciplinarity movement, the ease of access to literature from areas other than linguistics through online journal databases, and, more importantly, with the publication of the seminal multimodality theory and toolkit by the linguists Kress and van Leeuwen (1990, 1996) and O'Toole (1994), the meaning potential of the

---

[3] We will se later on, however, that depending on the discipline non-verbal elements can be as obligatory as words in the RA, giving it the status of a *multimodal* genre. Even if there are no figures, graphs, pictures or the like in a text, it has been argued, it would still be multimodal, given that features such as layout and typeface style are non-verbal and meaning-making elements as well (Lemke, 2004, p. 72).

non-verbal language has been unveiled and made analytically accessible. For Kress (2000), "it is now no longer possible to understand language and its uses without understanding the effect of all modes of communication that are copresent in any text" (p. 337). In written scientific discourse, the visual mode seems to carry a particularly "heavy functional load (...), where communication is often nearly impossible without the use of visuals like tables, graphs, or figures" (Tardy, 2005, p. 320). Lemke (1998, p. 89), for example, points out that "there is typically at least one and often more than one graphical display *and* one mathematical expression per page of running text in typical scientific print genres". In 2004, Swales himself seems to be foregrounding this point by stating that throughout the centuries, visuals have assumed increasing prominence in RAs and "in doing so, have themselves been increasingly integrated into the argument" (p. 216).

The functional load of non-verbal modes within scientific communication seems to be further enhanced in the new[4] electronic medium: the *computer medium* and the *Internet*. With the transposition of print journals to the Internet, the meaning potential of these modes has expanded dramatically inasmuch as resources impracticable in print as audiovisual data can now be implemented.

However, to my knowledge, so far all research on RAs has examined print exemplars only[5], i.e., articles that appear in print versions of journals. Other electronic genres, on the other hand, have received a reasonable share of attention from genre analysts, systemic functional linguists as well as from multimodal discourse analysts. That is the case of the *e-mail* (Spooner & Yancey,

---

[4] If we consider that the print medium in science is 340 years old, the electronic medium of the Internet with its 20 years of age can be said to be 'new' or 'young'.
[5] Apart from my own research with Motta-Roth and Nascimento (Hendges & Motta-Roth, 2000; Nascimento, Motta-Roth & Hendges, 2003).

1996; Gains, 1999; Paiva, 2004; Biesenbach-Lucas & Weasenforth, 2001; Chen, 2006), the *blog* (Komesu, 2004; Miller & Shepherd, 2004; Caiado, 2007), the *personal homepage* (Caldas-Coulthard, 2005; Motta-Roth, Marshall & Reis, 2005; Motta-Roth, Reis & Marshall, 2007), *institutional websites* and *homepages* (Kress, 2004a, 2004b, 2005; Lemke, 1999, 2002; Eckkrammer, 2004; Kok, 2004; Djonov, 2005; Bezerra, 2007), *chats* (Araújo, 2004; Araújo & Costa, 2007; Araújo & Biasi-Rodrigues, 2005, 2007; Fontes, 2007; Leal, 2007), *online newspapers* (Ihlström & Åkesson, 2004; Knox, 2007), the *e-forum* (Xavier & Santos, 2005), *FAQs* (Crowston & Williams, 2000), and *netads* (Fortanet, Palmer & Posteguillo, 1998). The contributions of these studies in providing language educators as well as less experienced Internet users with information on how to approach and engage in these genres have been macroscopic. For that reason, if genre analysis wants to maintain its tradition of helping newcomers succeed academically, the urge to examine the *electronic* social practices of academia is imminent insofar as the Internet is currently the primary medium in scientific communication.

Specifically in the case of journals, recent reports have revealed that since the late 1990s there has been a major shift in journal usage from print to online journals (Tenopir, 2003; Tenopir & King, 1998, 2000, 2001, 2002), a phenomenon which eventually could lead to the extinction of print journals, as has been predicted by Harnad (1998), who believes that 80% of all journals will have stopped paper publication by 2010. Another important observation made by these studies and other studies in the fields of information science, librarianship and electronic publishing is that there is considerable disciplinary variation in the

academia-Internet relationship for not all disciplines are 'moving' towards the electronic medium with the same pace.

In an attempt to start shedding some light on the above mentioned issues which have not yet been covered in previous literature, this study considers the research article written in English in its electronic format in order to learn what has been happening to the genre – if anything – when 'transmogrifying' (Swales, 2004, p. 218) to the Internet. To do so, I draw on Swales's (1990, 2004) genre analysis approach and Askehave and Swales's (2001) methodological model, as well as on Martin's (1992) systemic functional notion of register. Kress and van Leeuwen's (2001) understanding of medium and Askehave and Nielsen's (2005a, 2005b) two-dimensional model for the analysis of digital genres will also be considered. Theoretical considerations regarding hypertext and hypermedia theory will also be employed.

Before presenting the general and specific objectives of the present research, it seems important to clarify, at an early juncture, the definitions of *medium* and *genre* adopted in this thesis (section 1.2), as well as the object of study undertaken here (section 1.3).

1.2 *Medium* and *genre* defined

The term *medium* is defined in the present work as the physical tangible technological apparatus by means of which texts are produced, distributed and consumed. Examples of different media are, thus, radio, television, cinema, telephone, printed page, and computer. This definition goes beyond the concept proposed by Jewitt (2004, p. 184), for whom *medium* refers exclusively to "how

7

texts are disseminated, such as printed book, CD-ROM, or computer application". By conceptualizing medium not only as a dissemination artifact but also as part of the production and consumption processes, the general idea conveyed here is that medium does play a role in these processes and ultimately, in why and how genres are used as well as with respect to what genres look like. In this sense, it could be assumed that the same words spoken in a face-to-face encounter, printed on paper or presented on television *would* provide three different kinds of messages. I am obviously not re-inventing the wheel with this suggestion, but it seems that the role of the medium in meaning making has been less commonly acknowledged in genre theory than it has in other areas such as media theory. In the 1960's Marschal McLuhan (1964), whose work is viewed as the cornerstone of medium theory, already (over)emphasized the impact of media on meaning, coining the famous expression *the medium is the message*.

One notable framework in Applied Linguistics which considers the meaning-making potential of media was developed by Kress and Van Leeuwen (2001). According to the authors, in a multimodal theory of communication there are four domains of practice or *strata* in which meanings are dominantly made: *discourse*, *design*, *production* and *distribution*. Discourse and design would account for the content plane of communication while production and distribution for the expression plane. Content, i.e, discourse and design, is realized by the semiotic *modes* while expression, i.e. production and distribution, is realized by media.

Following Kress and Van Leeuwen (2001) and more specifically Askehave and Nielsen (2005a, 2005b), I propose therefore that the notion of *medium* should be incorporated more fully into the notion of genre.

As for the notion of genre, the approach adopted in this study follows Swales (1990) in that genres are viewed "as rather more than texts" (p. 6). As Swales (Ibid.) notes, "while it remains necessary to use texts in order to understand how texts organize themselves informationally, rhetorically and stylistically, textual knowledge remains generally insufficient for a full account of genre". In this perspective, a definition of genre should draw on terminological choices which may less probably lead to the idea that genre is restricted to textual activity. The Bakhtinian (1986, p. 60) definition, for example, that speech genres are "relatively stable types of *utterances*" (my emphasis) may be interpreted as emphasizing text over context rather than text *and* context together as enacting genre. Similarly, Bhatia's (2004, p. 23) description that genre "refers essentially to *language use* in a conventionalized communicative setting" may also lead to the idea that genre is only the textual element in an interaction.

On the other hand, in stating that genre is a communicative *event* (Swales, 1990) or "staged, goal-oriented, purposeful *activity*" (Martin, 1984/2001, p. 155, my emphasis), Swales (Ibid.) and Martin (Ibid.) suggest that a genre goes beyond text. Genres can be defined, in this sense, as relatively typified socio-cultural events/activities/manifestations mediated through a variety of semiotic modes and media and "permeated by a variety of discourses" (Motta-Roth, 2006a, p. 147).

Such a perspective requires a methodology which goes beyond the analysis of texts. For this reason, in this study both textual and contextual analyses were developed, according to the model proposed by Askehave and Swales (2001) (see Chapter 3 for a detailed explanation of the model).

1.3 Object of study

Defining the object of study examined in this investigation was an uncomplicated mission. Naming it, however, entailed tackling the complex issue of genre labeling (Swales, 2004; Askehave & Swales, 2001). From Kress's (2003) point of view, the real issue "is not really to have labels, though they can be useful devices, and it is clear that bad labels can be importantly misleading. […] the real issue is to understand the generic nature of the text – what meanings does the text realize, what social meanings are at issue?" (p. 119). My argument is that in the case of digital genres, and particularly in the case of those which have print counterparts, clear labels are "useful devices".

As stated in section 1.1, the present study investigates the *research article in its electronic format*. In attempt to provide a clear characterization of this object of study, the expression *research article in its electronic format* will be divided into two parts: 1. *research article* and 2. *electronic format*. First, I will consider what is meant by *research article* in this study and, then, move on to the issue of digital genre labeling and finally, referring to how *electronic format* should be interpreted.

Regarding what is meant by *research article*, it is not always clear in the literature whether the term is used to refer specifically to experimental RAs or if it is used as an umbrella term to refer to a group of serially published genres, which, apart from the experimental RA, would include review articles and short communications. Not only has there been an under-specification in relation to the nature of the RA analyzed, but also an unjustified variation in the terminology used to refer to the RA, which has also been called 'scientific article', 'journal article', 'research paper', 'research report', 'academic article', 'scholarly article'; both

shortcomings which could challenge the validity of findings about the RA. In the present study, the term *research article* identifies what Swales (2004) has referred to as the "data-based" (p. 213) or "standard research article" (p. 217), also called "primary research article" elsewhere (Yang & Allison, 2004, p. 265). This genre reports results from experimentation and generally follows the Introduction-Methods-Result-Discussion (IMRD) structural pattern.

As far as the issue of digital genre labeling is concerned, it appears that the trend is to add the words *electronic-* (or simply *e-*), *web-* or *online* to known genre labels. That is the case, for example, of *e-mail*, *e-book*, *web-log* and *online news report*. It could be argued that this labeling process would be logical if the new terms were only calling attention to a change in the distribution media of those genres, but not to a change in the genres themselves. In this sense, an e-book would still be the genre *book*, i.e., it would fulfill the same purpose and be structurally organized as a print book, with the only difference being that of the medium[6]. Notwithstanding, if the new term is referring to a *new* genre, using the same name of an already existing print genre and simply adding an *e-* could generate misidentifications, misclassifications and misuse of the new genre by newcomers. In this sense, the appropriateness of the name *web-log* could be questioned, given that it is has already been identified as a new genre (Miller and Shepherd, 2004).

Adversely to what happens with the terms *e-mail*, *e-book*, *web-log* and *online news report*, there is less labeling consensus when it comes to my object of

---

[6] For some authors (Crowston & Williams, 1999, 2000) genres which only use the Internet as distribution media cannot be classified as digital genres.

study. Before discussing its labeling process, though, it seems necessary to say a few words about its nature.

Frequent ejournal users may be well aware of the fact that in many ejournals research articles are available in – at least[7] – two electronic formats: an HTML (Hypertext Markup Language) format and a PDF (Portable Document Format) format. HTML research articles are hypertextual, multidimensional, and in general quite different from print research articles in terms of layout. They have been adapted to correspond to the software, hardware, operating system constrains and affordances of the electronic medium of the Internet.

PDF research articles, on the other hand, look very much like their print counterparts (in terms of layout, pagination system, schematic structure, for example) and one could even say the PDF format is a "clone" of the paper format (Peters and Smith, 2003, p. 74). The PDF format became famous for its stability in representing documents in a manner that is "independent of the original application software, hardware, and operating system used to create those documents" (The Free On-line Dictionary of Computing, n.d.).

On the assumption that PDF research articles are in the first instance drafted to appear as print documents and given their paper-like configuration, it is acceptable to argue that research mentioned earlier about print RAs is equally valid for PDF articles. In this sense, it seems more urgent to study the research article in its HTML format in order to understand if and how the novelty features of this format have had an impact on the RA as a genre. Hence, the electronic format examined in the present research is the HTML version of research articles.

---

[7] Another format sometimes available is the PDA (Personal Digital Assistant).

Now, how should this hypertext be called if it is not clear yet whether it is or not a new genre? As I noticed before, so far there is no common sense on how it should be labeled. A survey through the tables of contents of several Linguistics and medical ejournals revealed that the terms "Text+Graphics", "HTML" and "Full Text" (with variations such as "Full Text HTML" and Full Text+Links") are all used to identify the HTML format of research articles. In the vast majority of medical ejournals the term that is most frequently used is "Full-Text" for the HTML format in opposition to "PDF". Sometimes, however, the term "Full Text" is also used to refer to the PDF format and for this reason it no longer seems an adequate label for my object of study.

For lack of a better word and convention, for the purpose of this research the object under examination here will be referred to as *HTML research article*. Following the digital genre labeling trend, another option would be to use the term *electronic research article* to refer to my object of study, but then again this label could give margin to confusion in the sense that the PDF format can also be called *electronic research article* because it is an electronic format as much as the HTML format (Kwan-Hoong, 2006).

In future accounts I may choose another name for as LaPorte et al. (1995) observes:

> the nature of the information super highway also allows us to redesign the mode of information transfer – the journal article itself. We are not dealing with journal articles any more – we are dealing with research communications among scientists. Thus the terms "article", "paper", and "publication" should die.

At this point, nonetheless, maintaining the name *research article* and adding a word which identifies the medium (print or electronic) or the mode (written or visual) seems to be the right direction. The ejournal *Journal of Visualized Experiments* (*JoVE*)[8], for example, refers to its innovative multimedia research articles as *video articles*.

Having specified the object of study, in the next section the purposes of the present work will be presented.

1.4 Purpose of the study

The main purpose of this study is to verify whether the HTML research article is a new genre or not. By developing a genre analysis of HTML research articles written in English, I aim at learning to what extent and how the context of the electronic medium of the Internet affects the already well-known research article as a genre. In addition, it is also my objective to find out whether electronic academic genres are also sensitive to disciplinary variations. For this reason, two disciplinary domains will be contrasted: Linguistics and Medicine.

The relevance of considering academic discourse within a specific discipline has been well documented (Bazerman, 1988; Dudley-Evans & Henderson, 1990; Myers, 1990; Salager-Meyer, 1990, 1991, 1992, 1994, 1999, 2000; Salager-Meyer & Alcaraz Ariza, 2003; Salager-Meyer, Alcaraz Ariza & Zambrano, 2003; Henderson, Dudley-Evans & Backhouse, 1993; Berkenkotter & Huckin, 1995; Bhatia, 2004). For the purpose of the present research, the category

---

'Linguistics' is used following the Social Sciences Citation Index (2006) as an overarching category involving Linguistics, Applied Linguistics, Sociolinguistics, Psycholinguistics, Computational Linguistics and Semiotics. Similarly, 'Medicine' covers both clinical and research areas related to the field. The choice of Medicine results from the fact that, immediately after the birth of ejournals, medical scholars revealed to be very enthusiastic about the possibilities of the new medium for their publications (LaPorte et al., 1995; LaPorte & Hibbits, 1996; Delamothe & Smith, 1999, 2001; LaPorte et al. 2002). LaPorte et al. (1995) in particular have expressed their enthusiasm in the following way:

> The latest internet technology uses the web/ hypertext language. An article might have certain terms highlighted in the text: if the Cox proportional hazard model were highlighted, a click on this would bring up information about the Cox model. Citations would not be needed: a click on "Jones 1994" would pull up the complete article. A click on the data in the article would allow direct access to the stored data. We need not be constrained by text alone; a click on "Ron LaPorte" could bring up a video with him saying how important the work is.

Therefore, to verify to what extent predictions such as LaPorte et al.'s have been implemented in medical ejournals I included this field in my analysis. The need for evidence or parameters when discussing if and how HTML research articles reflect disciplinary idiosyncrasies took me to include another discipline into the analysis, Linguistics, whose choice is essentially related to my personal interest as member of this field and regular ejournal user.

The specific objectives that will guide my investigation are:

1. to learn if the communicative purpose of the HTML research article echoes the communicative purpose of the traditionally print RA based on contextual analyses, i.e., on interviews with HTML research article users (producers and consumers);

2. to learn if the communicative purpose of the HTML research article echoes the communicative purpose of the traditionally print RA based on textual analyses, i.e., on an investigation of the HTML research article's schematic structure;

3. to describe and discuss consistency and variation (Hasan, 2004) in HTML research articles at the level of genre across the disciplines of Linguistics and Medicine;

4. to study to what extent and how differently the contexts of Brazil and Australia[9] affect the HTML research article at the level of genre;

5. to investigate and discuss the role of the medium, i.e. of the technology of the Internet, in shaping the HTML research article;

6. to verify to what extent the theoretical approach for analyzing print genres introduced by Swales (1990, 2004) and by Askehave and Swales (2001) applies to the analysis of digital genres.

On a broader basis, in developing this study I also look forward to better informing the research groups ((GRPESq/CNPq) I am a member of, namely: *NUPDiscurso* at UFSC and *Linguagem como Prática Social* at UFSM as well as contributing to the language teaching and learning practices in these universities. Since the work developed at both institutions deals with teacher education, it would be reasonable to expect that some impact on elementary and high school may, as a matter of consequence, be achieved. By the same token, I hope to offer subsides for future genre accounts of academic digital genres.

---

[9] The reasons underlying the choice of the contexts of Brazil and Australia are described in Chapter 3 – Methods.

Chapter 2   Literature Review

2.1 Introduction

The purpose of this chapter is to present an overview of research findings about the genre research article, from its emergence in the seventeenth century until its movement towards the electronic medium of the Internet. In the ensuing section, I refer to the origins, purpose and structure of the print RA. Firstly, aspects of the cultural and social environment in which the genre came into existence will be considered and then an exploration of the changes it has undergone in terms of purpose and form as a result of cultural and social pressures since the seventeenth century to current days will be reviewed. Next, section 2.3 zooms into studies about the print RA as it is used in the two specific disciplines considered in this study, Linguistics and Medicine. Section 2.4 moves away from the print scenario towards the electronic medium by reviewing literature on electronic journals, given that the number of studies which consider the electronic RAs is scarce. Although not focusing specifically on electronic RAs, studies about ejournals can help understand state of the art features of this genre. Finally, section 2.5 also concentrates on the electronic medium, but from the perspective of genre analysis, integrating studies that draw upon digital genres, hypertext and media.

This recollection is essential in helping understand the developments of the RA in terms of its purpose and structural configuration as well as the extent to which the move from paper to screen has had repercussions on the genre.

2.2 The print research article: historical developments

The historical developments of the RA are inseparably related to the history of journals. For that reason, I begin this section looking first at the creation and development of journals in general and then to the emergence and evolution of the RA per se.

2.2.1 The birth of the print journal

A chief assumption within social semiotics is that "[n]ew social, cultural and political needs lead to new ways of communicating and to new communication technologies" (Kress and van Leeuwen, 2001, p. 112). In science, it has been argued that genre conventions can change and "the forces and motives behind these changes may be various, ranging from changing social conditions to changing intellectual conditions, from self-interest to reason" (Bazerman, 1984, p. 165). As a new communication technology, the birth of the scientific journal[10] is no exception to this assumption.

The first journal-like publications were conceived as a response to the need for an improved system for the dissemination of scientific discovery. Just like nowadays, book publishing was not timely, and until the early 1600s, scientists –

---

[10] In the literature several terms have been randomly used to refer to the scientific journal: journal, scholarly journal, academic journal, scientific periodical, serial. Therefore, I will use them interchangeably.

or "natural philosophers" as they were then called (Tenopir & King, 2000, p. 17), basically used face-to-face contact and letters to communicate with each others. According to Tenopir and King (2000, p. 56, footnote 3), these letters were typically written in a "one to many" style and were then copied to be distributed, a communication system which worked well while the number of scientists was small. Nevertheless, as soon as small clusters of scholars evolved into scientific societies and the amount of scientific production increased, written correspondences were no longer a satisfactory alternative: the readership had grown, copyists were very expensive and the whole process became too restricting and time-consuming (Houghton, 1975; Peek & Newby, 1996; Gross, Harmon & Reidy, 2000, 2002; Tenopir & King, 2000), a stimulating scenario for the birth of the scientific journal.

Two titles are generally[11] recognized as the forerunners of the modern scientific journal: *Le Journal des Scavans*, which appeared in France, and *Philosophical Transactions*, in England[12] (Houghton, 1975; Peek & Newby, 1996; Gross, Harmon & Reidy, 2000, 2002; Tenopir & King, 2000). Both appeared in 1665, in January and March, respectively, and exist to the present day. The success of these journals as a new form for *collective* communication soon led to the creation of others so that by the end of that very century, more than 30 titles were being published (Houghton, 1975, p. 18; Tenopir and King, 2000, p. 57).

---

[11] Some authors (Houghton, 1975; Swales, 1990) suggest that *Philosophical Transactions* was the precursor because of the experimental nature of the articles it published, so does the *The Royal Society* of London itself: "*Philosophical Transactions of the Royal Society* has the prestige of being the world's longest-running scientific journal (The Royal Society, 2006).
[12] Early scientific publications were "usually under the patronage of royalty, aristocracy, church or state, such as the Academie Française, established in 1635, or the British Royal Society" (Bostock, 2006, p. 237).

But not all journals created until the end of the seventeenth century thrived. According to Houghton (1975), they only became the accepted medium for scientific communication around the 1750's. Up to 1800, around 700 journals had been produced (Houghton, 1975, p. 18), a number which grew to over one thousand titles by the middle of the nineteenth century (Figure 2.1).

In 1963, Price (cited in Tenopir & King, 2000, p. 57) reported that growth rates indicated that "the number of journals doubles approximately every fifteen years", which would leave us with around one million titles by the year 2000. In 1995, however estimates suggested there were 70,000 to 80,000 journal titles worldwide (Meadows & Singleton, 1995, cited in Tenopir & King, 2000, p. 57). The growth in the number of journals strongly correlates with the growth in the number of scientists (Tenopir & King, 2001, p. 672).
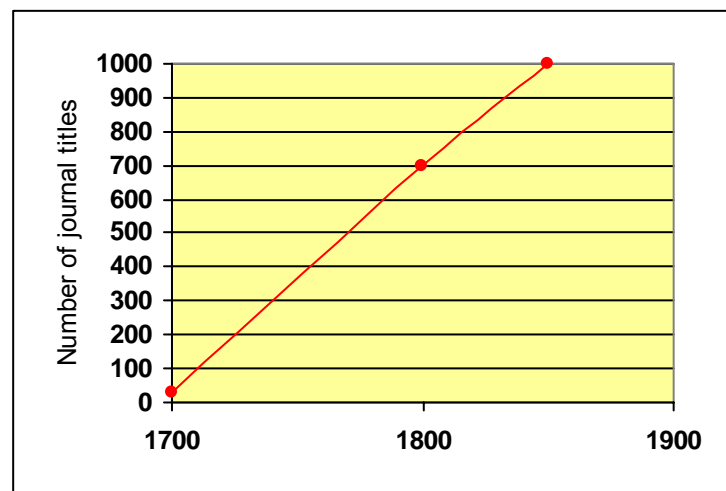
Figure 2.1. Number of journals from 1700 to 1850

The function of journals has also changed over time. They first appeared as a means to allow scientists to communicate in a cheaper and faster manner than

the letter system. Today, thinking of them solely as "distribution outlets" is no longer reasonable (Peek & Newby, 1996; Rowland, 1997; Butler, 1999). Peek and Newby (1996) comment that "[t]he founders of the first scholarly journals […] would no doubt be surprised at how the journal has changed" (p. 4-5).

After becoming more solidly accepted around the 1750s, their role also gained a higher status. Houghton (1975) presents us with a list of the multi-functionality of journals at that time:

1. they provided the scientific community and interested layity with news in the vernacular of work previously reported in foreign language;
2. they provided the means for scientific and literary men to discourse on scientific work without having read the complete original account;
3. they conserved material which would otherwise have been dispersed through publication in individual tracts or pamphlets;
4. they aided scholarship by providing inexpensive channels of communication;
5. they encouraged scientists to plublicise their work;
6. they offered a forum for the continuous critical examination of scientific hypotheses and theories . (p. 19)

As for the role of modern journals, Rowland (1997) believes that their four key functions are 1. to disseminate information, 2. to control quality, 3. to serve as canonical archive and 4. to provide recognition of authors.

Peek and Newby (1996) point out that part of the changes in the role of journals from the 1700s if compared to modern journals

can be credited to technological advancement and structural changes in the academic reward system. Papers in early journals were really works-in-progress, with future monographs being considered the final stage, or the "published work". In modern times the journal is, for many disciplines, the final destiny of a scholar's published work, and thus the fundamental purpose of the journal has significantly changed. (p. 4-5)

The argument that books are no longer "the final destiny of a scholar's published work" has been confirmed in several studies. Bazerman (1984), for example, in examining the changing features of papers published in *Physical Review* from 1890 to 1980, found that "this period marks the virtual disappearance of the book as a way of presenting new results in physics" (p. 66). Likewise, studies in the area of bibliometrics[13] learned that in medical papers, nearly 95% of references are to documents published in journals, meaning that "scientists from the health sciences (biomedical research and clinical medicine) are those for whom journal articles are the most important source of scientific knowledge" (Larivière, Archambault, Gingras & Vignola-Gagné, 2006, p. 1000). In social sciences, this number falls below 50% inasmuch as the book still absorbs a great deal of a scholar's published work in this area (Glänzel & Schoepflin, 1999; Larivière, Archambault, Gingras & Vignola-Gagné, 2006).

As a result of their contrastive bibliometric study between sciences and social sciences, Glänzel and Schoepflin (1999, p. 42) argue that *immunology*, where 94.3% of references are to serials, and *research medicine* (92.1%) are fields where "journal publications have the highest relevance for communicating research results, while in *history and philosophy of science and social sciences* [34.7%] the relevance would be much lower".

In spite of disciplinary variation, journal publications are a critical source of information "inside and outside the academic community" (Tenopir & King, 2001,

---

[13] Characterized by the analysis of the kinds of documents referred to in reference lists of papers across a variety of fields

p. 672). Tenopir and King (1998) argue that journals are preferred over other types of publications "because they cover a range of topics, editors attempt to maintain content quality, the information is relatively current, and the articles serve a variety of purposes". And it is specifically to the articles published in journals that I turn now.

2.2.2 The birth of the print research article

In *The origin of genres* (1976), Todorov argues that "[a] new genre is always the transformation of one or several old genres: by inversion, by displacement, by combination" (p. 15). In answering the question of where genres come from, he remarks: "[q]uite simply from other genres" (Ibid.). If in following Todorov (1976), therefore, we would have to answer the question *from what genre did the modern print RA come from*, the answer would definitely be 'from the personal handwritten letter'[14].

Genre transformations take place because, as discussed in Chapter 2, genres are dynamic institutions undergoing continuous evolution (Swales, 1990; 2004) as a consequence of the dialogic relationship (Bakhtin, 1986) – or *recursive* process (Halliday & Martin, 1993) – between language and context. From this perspective, shifts in the form and function of the set of genres that constitutes a culture (the scientific culture, for instance) are caused by and cause societal shifts. The rate with which these shifts occur varies from one genre to another and may

---

[14] In this regard my supervisor's, Dr. José Luiz Meurer, comment was: "in my own concern about context, I'd ask 'and what caused/gave rise to the origin of the personal handwritten letter'?". Luckily, two weeks later (September, 2007) I had the opportunity to meet the pioneer in the study of the history of the RA, Dr. Charles Bazerman. Dr. Bazerman said that he had not investigated nor accidentally come across any information regarding the origin of the handwritten letter use in scientific communication, but thought it was a very interesting topic to pursue.

be linked to such factors as the genre's age, prestige, frequency of use, and technological developments. It is not surprising therefore that the RA has undergone numerous changes given its age, high status and ubiquitous usage, with millions of instantiations a year.

The prestige of the RA among the scientific community has led an uncountable number of researchers to investigate its features, but diachronic examinations are less common. One substantial book-length work on the evolution of physics RAs was developed by Bazerman (1984). Other studies which have undertaken historical accounts of the RA, though less profoundly than Bazerman, are Houghton (1975), Swales (1990, 2004), and Gross, Harmon and Reidy (2000, 2002)[15]. These, however, have not considered specific disciplines and ideally, according to Bazerman (1984), changes in scientific written conventions should be considered within the scope of particular disciplines because "conventions of prose features are part of the historical developments of a discipline, closely related to its changing intellectual and social structure" (p. 166). Besides Bazerman (1984, 1988), studies of this nature are rare, but significant contributions have been made by Salager-Meyer's (1999, 2000) in Medicine.

Houghton (1975, p. 12) sustains that "the well-structured scientific paper as it is known today did not appear until after 1780 and 1790", i.e., the years after the scientific journal had been accepted as a legitimate channel for scientific exchange. However, as shown by Bazerman (1984, 1988), Swales (1990),

---

[15] For a list and short description of some of these and other studies on the research article's history see Atkinson (1999, p.202-203) and Swales (2004, p. 216).

Salager-Meyer (1999, 2000), and Gross, Harmon and Reidy (2000, 2002), even after 1780 and 1790 the RA metamorphosed significantly.

Deriving from personal handwritten letters exchanged between scientists prior to the foundation of journals, it is not surprising that in their primitive form RAs were written in first person and some even began with the salutation 'Sir' (Swales, 1990, p.110). However, as science continued to evolve, the linguistic and rhetorical means by which it was communicated also changed and soon there emerged a "genre increasingly distinct from its letter-writing origin" (Swales, ibid.). As suggested by Swales (Ibid., p. 114), there are "certain quite marked differences between the eighteenth and the twentieth century RA; among them a decidedly more casual approach to the previous literature and some continuation of the epistolary convention of first person narrative", features that have almost totally disappeared from the contemporary RA.

In her longitudinal study of citation behavior in scientific medical prose from 1810 to 1995, Salager-Meyer (1999, 2000) found that in the 19th century "not only were the cited surgeons'/physicians' surnames mentioned, but so were their first names preceded by a polite and genteel 'Mr' or 'Professor', followed, almost invariably, by their place of work" (Salager-Meyer, 1999, p. 290). According to Valle (cited in Salager-Meyer, 1999, p. 290) "whether this is because the field was still small enough for the writer to assume that his addressees would know the reference or because the apparatus for academic and scholarly documentation was not yet sufficiently developed remains an open question". For Salager-Meyer (ibid.), both factors have contributed for the way citations were at that time.

Bazerman (1984), examining features of experimental articles appearing in the journal *Physical Review* from 1893 until 1980, such as length, reference, sentence length and syntax, word choice and graphics concluded that these features "strongly indicate the increasing abstraction, web of background information, density of knowledge, interpretation and focused argumentation going into the PR [*Physical Review*] article since 1893" (p. 181).

Despite the argument that it is difficult to make any valid claims about the modern RA without considering specific disciplinary idiosyncrasies, in the next sub-section, I will try to describe the purpose and structure of the print RA in general terms as it has been reported in previous literature. In section 2.3, on the other hand, attention to disciplinary conventions will be given, particularly to the areas studied here, Linguistics and Medicine.

2.2.3 The contemporary print research article's purpose

In his 1990 definition of genre, Swales suggested that a genre can serve more than one purpose. According to the author, "a genre comprises a class of communicative events, the members of which share some *set of communicative purposes"* (p. 58) (my emphasis). In 2001, Askehave and Swales explained that identifying these communicative purposes may not be a straightforward task, for these purposes are not a simple enumerable list but "a complexly layered one, wherein some purposes are not likely to be officially 'acknowledged' by the institution, even if they may be 'recognized' – particularly in off-record situations – by some of its expert members" (p. 199).

Linked to the communicative purpose identification issue is the problem of genre labeling (comments on the issue of genre labeling were already introduced in Chapter 1, section 1.3). As far as the name 'research article' is concerned, it has not always been clear in the literature to what genre this label was referring to. Swales (2004) himself acknowledged that his 1990 definition of the *research article*, "was, perhaps predictably, over-generalized" (p. 207). Although the ground-breaking nature of Swales 1990 study of RAs is unquestionable, his definition of the RA can nowadays be refined given the amount of work on RAs that followed from his (Ibid.) work. In 1990, Swales defined the RA as

> a written text (although often containing non-verbal elements), usually limited to a few thousand words, that reports on some investigation carried out by its author or authors. In addition, the RA will usually relate the findings within it to those of others, and may also examine issues of theory and/or methodology. It is to appear or has appeared in a research journal or, less typically, in an edited book-length collection of papers. (p. 93)

In 2004, the author explains that his definition may not be adequate because "the research article per se is not a single genre, but rather divisible into at least theoretical and experimental papers" (p. 208). In a similar fashion, Gross, Harmon and Reidy (2002, p. 11) acknowledge that the term scientific article can refer to all types of articles, i.e. methodological, experimental, observational, theoretical and review articles (the issue of genre labeling was already discussed in Chapter 1, Section 1.3). Hence, Swales (2004) proposes that the traditional research article "needs to be subcategorized into theory pieces, review articles, and the experimental or data-based RA itself" (p. 213). However, he also

recognizes that further subcategories could be added, "such as dividing medical RAs into *clinical* and *empirical* subtypes but at increasing risk of poorly motivated proliferation" (ibid., p. 213). In fact, a look at the table of contents of medical journals will reveal that in this field, the range of genres used to report science goes far beyond the four categories suggested by Swales (Figure 2.2).

Figure 2.2 shows that *The New England Journal of Medicine* indicates nine 'types' of "clinical and research articles" and eight categories of "analysis and opinion articles", and that *The Journal of the American Medical Association* (*JAMA*) offers 13 "Categories of articles". Although I am not in a position to decide whether each of these categories is a different genre, there must be something, some features that make it worth classifying each one of them under a different label. It would be interesting to investigate which of the articles published in *The New England Journal of Medicine* are the same genre of those published by *JAMA* despite label variations. The variety of 'article types' has not always been so big in medical journals and the Internet has certainly opened space for new genres such as "Videos in clinical medicine" and "Audio Interviews", published in the online version of *The New England Journal of Medicine* only as they are impractical in paper. The number of pages of medical articles has also been shrinking while the number of sections in medical journals has been increasing over the years. Current volumes of JAMA have 15 to 19 sections per issue and *The New England Journal of Medicine* contains 10 to 15 section divisions. It could be speculated that soon they'll look more like newspapers or magazines than like traditional print journals.

The NEW ENGLAND JOURNAL of MEDICINE

HOME | SUBSCRIBE | REGISTER | E-MAIL ALERTS | HELP | Search NEJM | GO | Advanced Search

Features Home
Clinical & Research Articles
Analysis & Opinion Articles
Educational Resources
Time-Saving Tools
Delivery Options
Your Access Options

SEE ALL FEATURES

HOME

CLINICAL & RESEARCH ARTICLES
Original Articles
Special Articles
Images in Clinical Medicine
Review Articles
Clinical Practice
Case Records of the MGH
Clinical Problem-Solving
Clinical Implications of Basic Research
Videos in Clinical Medicine

ANALYSIS & OPINION ARTICLES
Perspective
Sounding Board
Legal Issues in Medicine
Health Policy Reports
Editorials
Correspondence
Book Reviews
Audio Interviews

EDUCATIONAL RESOURCES
Medical Mystery
Image Challenge
PowerPoint Image Downloads
PowerPoint Slide Sets
Medical Meetings
NEJM Review CME
NEJM Weekly CME
Videos in Clinical Medicine

TIME-SAVING TOOLS
This Week in the Journal
Personal Archive
CareerCenter
PowerPoint Image Downloads
PowerPoint Slide Sets
Search
Audio Summary

DELIVERY OPTIONS
Early Release
Alerts
Handheld Optimized Site
E-mail TOC
PDA Downloads
RSS

YOUR ACCESS OPTIONS
Individual Subscription
Institutional Subscription
Registration
Visitor
Pay Per View
Site Pass
Access Chart

JAMA&

JAMA®
The Journal of the American Medical Association

HOME | CURRENT ISSUE | PAST ISSUES | COLLECTIONS | CME | CAREERNET | SUBSCRIBE | HELP

Instructions for Authors

JAMA Instructions for Authors

General Information
  Editorial Office Contact Information
Forms
  Authorship Form
  Patient Permission Form
Editorial Policies for Authors
  Authorship Criteria and Contributions and Authorship Form
  Role of the Corresponding Author
  Group Authorship
  Conflicts of Interest and Financial Disclosures
  Funding/Support and Role of Sponsor
  Data Access and Responsibility
  Duplicate/Previous Publication or Submission
  Timeliness of Data
  Survey Research
  Reporting Race/Ethnicity
  Ethical Approval of Studies and Informed Consent
  Identification of Patients in Descriptions, Photographs, and Pedigrees
    Patient Permission Form
  Personal Communications and Unpublished Data
  Previous Presentation or Release of Information
  Embargo Policy
  Unauthorized Use
Editorial Review and Publication
  JAMA-EXPRESS
  Editorial and Peer Review
  Editing

Categories of Articles
  Original Contribution
  Clinical Trial
    Trial Registration
    Protocols
    CONSORT Flow Diagram and Checklist
  Reports of Diagnostic Tests
  Reports of Cost-effectiveness Analyses and Decision Analyses
  Systematic Review (Including Meta-analysis)
    Reports of Meta-analyses
  Caring for the Critically Ill Patient
  Brief Report
  Special Communication
  Commentary
  A Piece of My Mind
  Letter to the Editor
  Research Letter
  Poetry and Medicine
Articles Requiring a Presubmission Inquiry
  Clinical Crossroads
  Clinical Review
  Grand Rounds
  Health Law and Ethics
  On Call
  Patient-Physician Relationship
  Controversies
  Book and Media Reviews

Figure 2.2. Variety of 'types of research articles' published in two medical journals

The relevance of specifying the nature of the article being analyzed derived from studies focusing on specific disciplinary communities as well as cross-disciplinary investigations of RAs. As pointed out earlier in this chapter, the evolution of the RA is closely linked to disciplinary practices, which means that "the style, presentation, and argument in scientific articles continually vary, and these variants selectively reproduce as a consequence of the changing needs of the diverse scientific disciplines" (Gross, Harmon & Reidy, 2002, p. ix).

Given that the term *research article* is used in the present research to identify what Swales (2004, p. 217) calls the "standard research article", let us go back to the issue of the purpose of this genre. In a seminal paper about the function and structure of RAs, Hill, Soppelsa and West (1982, p. 334) state that the purpose of the genre is "to present the results of an experiment or ex post fact study". Assigning a single purpose to a genre could be insufficient because, as pointed out by Swales (1990, 2004), a genre serves *sets of purposes*. In this sense, assuming that the purposes assigned to the RA in previous studies are the result of triangulation, the *set of communicative purposes* of the print RA can be said to be

1. report research experiments;
2. disseminate scientific information;
3. advance the state of knowledge in a field;
4. control quality of research;
5. provide recognition for authors and career upgrade;
6. secure research funding.

It could be argued, following Misser (n/d), that the purposes of the RA depend on who is writing it. In this view, the set of purposes listed above serve for expert writers. For a neophyte student writer, on the other hand, the communicative purposes of the RA would be, in Misser's (Ibid.) point of view, to show that s/he understands her field of study; that s/he can generate ideas and make intelligent observations to support or prove those ideas, and even that s/he can follow a style sheet in her field (Misser, Ibid.). This again shows how complex it is to determine the communicative purpose of a genre. In the present research, however, the purposes that will be taken as point of departure in the analysis of the HTML research articles are the five items listed above. In this sense, this work attempts to verify whether these purposes, i.e., the communicative purposes of the print research article, are the same as the communicative purposes of the HTML research article (the results of this analysis are shown in Chapters 4 and 5).

2.2.4 The contemporary print research article's structure

In terms of structure, one of the first influential accounts on the RA's overall organization is the study by Hill, Soppelsa and West (1982), represented non-verbally in the famous diagram reproduced in Figure 2.3.

In this diagram, what separates one block of information from the other, the authors claim, is the function each one fulfils in the research paper (ibid). The Introduction, in their words, makes "the transaction from the general field or context of the experiment to the specific experiment by describing an inadequacy or inaccuracy in previous research which motivates the present paper" (ibid., p. 335). The Procedure block, according to them, is frequently divided into two

sections – 'methods' and 'results', "[b]ecause there are two distinct steps in doing a research study, namely collecting the data and manipulating the data" (ibid., p. 336). Characterized as a "mirror image of the introduction" (ibid., p. 337), the Discussion "moves from the solution of the problem that motivated the study to the implications of that solution for the larger field" (ibid., p. 337). Though recognizing that "style and content" of RAs vary across fields, Hill et al. (ibid.) argue that this structure is a shared pattern.
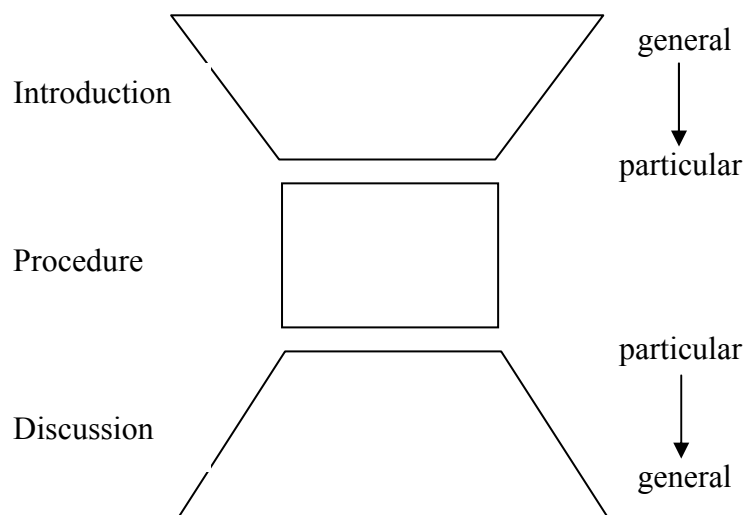
Figure 2.3. Hill, Soppelsa and West's (1982, p. 335) representation of the overall organization of the RA, reproduced with adaptation in Swales and Feak (1994/2004, p. 222)

In a similar fashion, the structural description proposed by Swales (1990) in the IMRD – Introduction-Methods-Results-Discussion – descriptive model, suggests that this pattern is consistent across disciplines.

Over the past 25 years, nonetheless, several examinations of experimental RAs in different disciplines have shown that structural aspects of the RA vary significantly across disciplines. In this respect I believe it is useful to recall Hasan's (1985/1989) Generic Structure Potential principle and argue that the structural disciplinary variation of the RA is due to the 'optional' elements in the structure of a genre, while the core or 'obligatory' elements of the RA structure, i.e., the IMRD, is maintained independently of the discipline. For the purpose of my research, the structure of the print RA used as point of departure for the analysis of the HTML research articles is the IMRD descriptive model (Swales, 1990).

In the next session I will briefly discuss previous literature on the RA in the two disciplines considered here – Linguistics and Medicine – in an attempt to identify state of the art and relevant information about how each field shapes their print RAs. This review may offer important insights into disciplinary practices which could have an impact on the relevance, use, configuration and purpose of HTML research articles in each field.

2.3 Disciplinary idiosyncrasies: the print research article in Linguistics and in Medicine

Unlike the print RA in Medicine, the print RA in Linguistics is relatively under-researched, maybe because of a belief that, as linguists and therefore language/genre experts, we need less instruction in academic literacy than other areas. The only studies to my knowledge that have examined the RA in Linguistics are Yang and Allison (2003, 2004), Hendges (2006) and Ozturk (2007). The medical RA, however, has been approached much more frequently (Salager-

Meyer, 1990, 1992, 1994, 1999, 2000; Thomas & Hawes, 1994; Williams, 1996, 1999; Nwogu, 1997; Hydén & Mishler, 1999; Varttala, 1999; Gledhill, 2000; Luzon Marco, 2000; Rowley-Jolivet, 2002). This wealth of publications dealing with medical RAs is probably due to the fundamental role this genre plays in Medicine[16] (Salager-Meyer, 1990, 1992, 1994, 1999, 2000; Glänzel & Schoepflin, 1999; Larivière, Archambault, Gingras & Vignola-Gagné, 2006), where it is the main scientific publication means, while linguists still rely a lot on books. Helping medical scientists in their RA literacy, therefore, seems to be more urgent that helping linguists.

In a few words, these studies reveal two main contrasting aspects in relation to the RA in each discipline. One aspect refers to the structure of the genre and the other one is related to the relevance of visuals for each area.

As far as the structure of the RA in Linguistics is concerned, Yang and Allison (2003, 2004) and Ozturk (2007) have found that besides the IMRD pattern, three other sections, namely *Theoretical basis*, *Review of the literature*, and *Pedagogical implications*, are core features of the genre in this field. Specifically in relation to the Introduction section in two subdisciplines of Applied Linguistics, second language acquisition research and second language writing research, Ozturk (2007) found that variation occurred not only inter-disciplines but also intra-disciplines, and that a single journal can publish articles with a variety of move structures, which would suggest that in this area "journal editors do not seem to prescribe a particular structural organisation to authors" (p.35). In Medicine, on the

---

[16] Whenever I use the term *Medicine*, it should be read as a 'big' overall term encompassing both research and clinical Medicine.

other hand, structural aspects are less open to flexibility, as Salager-Meyer and her colleagues have found (Salager-Meyer, 1990, 1992, 1994, 1999, 2000; Salager-Meyer & Alcaraz Ariza, 2003; Salager-Meyer, Alcaraz Ariza & Zambrano, 2003). Medical RAs, therefore, seem to follow the IMRD pattern more homogeneously. One exception are the RAs published in the journal *Nature Medicine,* which although presenting all four sections, place the Methods section at the very end of the article, in a structure that looks like this: Introduction-Results-Discussion-Method, or IRDM. According to Ayers (2007)[17], this tendency evidences that RAs are becoming more news-oriented, where the methods are relegated to a secondary position.

Linguistics and medical RAs also reveal variation when it comes to the role of images (figures, tables, charts) in scientific communication. In Linguistics RAs, Hendges (2006) found that there was great variation in the use of images in the genre (from 1 image per RA to 17 images per RA), suggesting there is no established common sense on the relevance of non-verbal elements in the discipline. The importance of visual resources in scientific medical discourse, on the other hand, has been repeatedly acknowledged (Dubois, 1980; Rowley-Jolivet, 2002, 2004; Wang, 2006). In the medical RA, visuals are a valuable reading aid and promotional resource. As a reading aid, they "are the fast track to the new" (Miller, 1998, p. 43) and as a promotional resource, it is the nature and quality of the images that is going to define whether an RA is worth reading or not (Wang, 2006).

---

[17] I would like to thank Professor John Swales for calling my attention to this paper.

Both these aspects of disciplinary variation will be taken into account in the analysis of the HTML research articles. As far as ejournals are concerned, previous studies have already indicated disciplinary idiosyncrasies. These will be discussed in the next section.

2.4 Ejournals: historical developments

The review presented in this section derives from literature produced in various areas other than Applied Linguistics. Although linguists – specially in Brazil – have shown an increasing interest in online discourse and important contributions have been made on the idiosyncrasies of a wide range of digital genres (as listed in Chapter 1, p. 5-6), academic genres have not yet been tackled, with the exception of the studies developed by my colleagues and I on electronic RAs at Universidade Federal de Santa Maria (Hendges & Motta-Roth, 2000; Motta-Roth et al. 2000; Hendges, 2002; Motta-Roth et al., 2004) and of the study about online book reviews reported by Caballero (2007).

Nevertheless, in the areas of librarianship studies, information science, document design and medicine, ejournals have been the focus of attention of a big volume of work, the first dating back to soon after the first ejournals emerged, in the early 1990s. In this literature, the term ejournal has been used to refer to two different types of publication, depending on the distribution medium: 1) electronic-only journals and 2) print journals that moved to the electronic medium. Kling and McKim (1997) explain the difference between these types in the following way:

We define a **pure electronic journal** as an electronic journal which is primarily distributed and accessed through electronic channels. In doing so, we acknowledge that even a pure electronic journal may, and probably will, be printed out for reading, and possibly even stored in libraries in a printed form, for archival purposes. However, pure electronic journals are accessed *primarily* in electronic form. (…) We can then contrast the pure electronic journal with the **hybrid paper-electronic journal** (or **pe-journal**). The pe-journal is a package of peer-reviewed reports available through electronic channels, but whose primary access channels are paper-based. Examples of pe-journals include: *Science Online, Cell, Nature,* the *Journal of Biological Chemistry, Astrophysical Journal,* and the *Journal of Neuroscience.* Since the criteria for distinguishing a pure electronic journal from a pe-journal are anchored in the readership, a pe-journal could certainly become a pure e-journal, if readership changed such that the journal was accessed primarily electronically, and paper copies just were produced for archives or even libraries.

Pure electronic journals or electronic-only journals, which have also been called 'internet journals' or 'online journals' are, thereby, originally produced and published in electronic format. They may be *printed out*, as observed by Kling and McKim (Ibid), but they do not have a *print version* as the hybrid paper-electronic (or shall we say *impure*?) journals have. Although a number of professional groups had started experimenting with pure ejournals in the early 1990s, *Psycoloquy* is usually acknowledged as being the first and most publicized of these journals (Hitchcock, Carr & Hall, 1998; Tenopir & King, 2000); it appeared in 1990 and was published by the American Psychological Association.

The *pe-journals*, then, are those which migrated from the print medium to the electronic medium of the Internet. An *online journal* (purely electronic), in this sense, is different from a *journal online* (print-based).

This terminological distinction is not always acknowledged in the literature. From Llewellyn, Pellack and Shonrock's (2002) point of view, "there is a great deal of ambiguity about the use of the terms electronic or online journals in the

literature" because there are "multiple uses and definitions for the terms 'electronic-only journal', 'electronic journal' and 'ejournal' and because "there is no single agreed upon term for journals published only in electronic/online format versus the electronic/online version of a print publication". For this reason, in their investigation, the authors (Ibid.) specify that their meaning of *ejournal* is "e-only journals", i.e., electronic journals that are originally published only in electronic format.

Another perspective which seems to further complicate this terminological ambiguity refers to the variety of "e-journal models" that are available today, identified by Kwan-Hoong (2006) as:

- e-journal replaces print journal
- e-journal coexists with print journal
- journal is in electronic form only, but individual articles can be ordered in paper form
- e-journal is 'secondary to' the print journal
- electronic version is published several months after the print version
- print version is published several months after the electronic version
- the full print version is not available electronically
- both versions exist but with different pricing arrangements

It could be assumed that the reason pioneering studies about ejournals did not make terminological considerations is the fact that the first journals to appear on the Internet were genuinely electronic and, hence, the labels *online journal* and *ejournal* were primarily coined to define these publications. Only in the mid-1990's, when the first print journals started migrating to the electronic medium, did a terminological distinction become justifiable.

It was argued earlier in Chapter 1 that for the purpose of genre analysis and, more specifically, genre classification, defining the *name* of a genre is a

significant part of the analytical process. Now, what would be the role of terminological accuracy in areas such as electronic publishing, librarianship studies, and information science? Llewellyn, Pellack and Shonrock (2002) explain that each publication model will have a different impact on the economics, access, legitimacy, and interaction resources related to journal publishing and all these issues are of interest in the above mentioned fields. Moreover, those studies which decided to make a distinction between both types of journals (pure or print-based) did so because there was a general expectation that pure electronic journals would be more innovative, ground-breaking and would reveal to be more interesting for investigation and comparison with print journals than print-based electronic journals. Considering that there was already a great amount of literature on traditional print journals and that the first electronic versions of print journals were the exact copy of the print format, only digitalized (scanned), the belief was that there was nothing new to be said about print-based ejournals.

From the perspective of genre analysis, the classification of journals into electronic-only or print-based may help explain how their RAs are used, organized and designed and, hence, offer relevant insights about the evolution of journals and the emergence of new scientific genres. Pure ejournals, in this sense, could be less constrained by the print journal paradigm than print-based ejournals, experimenting with the new medium more freely, without direct threats from the indexing system because there would be no impact factor ranking at stake in the first place.

But the first pure ejournals seem to have failed in exploring the medium's affordances to the degree and with the speed expected by users and investigators

(Hendges, 2001). It soon became worthwhile asking whether pure ejournals had redefined the meaning of a scholarly journal and "exploded it into some new creation" or whether they were "just a continuation of a long tradition of scholarly publication" (Valauskas, 1997), because most of them emulated print journals. As soon as print journals started to evolve online, replacing scanned files by alternative technologies, they also began to impose a new convention for electronic publishing and pure ejournals either vanished or adapted to the new convention. An illustration of this is the e-only journal *Language Learning and Technology* (http://llt.msu.edu/). When it was first released in July 1997, its papers were available in an HTML format only and some made use of resources such as audiovisual data presentation (see journal's January 1999 issue, Volume 2, Number 2, retrievable from http://llt.msu.edu/vol2num2/article1/index.html), but it soon started publishing PDF versions of its articles along with the HTML versions, because that was the convention adopted by print-based ejournals. By the end of the 1990's and beginning of this century, pure ejournals and *impure* ejournals were probably very similar in terms of their overall structure, the genres they published and services they offered. Meanwhile, there was substantial contrast in terms of access and quality control (indexing, peer-reviewing), but these issues will not be addressed in depth in this research.

Quality control seems less of an issue in current pure ejournal publishing. At BioMed Central (http://www.biomedcentral.com), for example, the collection of free pure medical ejournals is being indexed and many titles have been gaining popularity (and rising impact factors) among the medical research community.

The distinction between a pure ejournal and a print-based ejournal is becoming less evident and maybe because of that the term ejournal has evolved to a broader definition including any journal available electronically. In Kwan-Hoong's (2006) words, an ejournal is "a serial publication available in digital format", independently of 1) whether this digital format is distributed via CD-ROMs, e-mail or Internet pages, 2) whether it is presented in HTML format or PDF format or 3) whether it has a print counterpart or not.

If pure ejournals follow the conventions of print-based ejournals and if print-based ejournals follow the conventions of print journals, what exactly are the differences between them?

In this regard, Valauskas (1997) notes that "[t]he medium distinguishes an electronic scholarly journal from its print counterparts but the process of developing content for both print and electronic peer-review scholarly journals is generally the same. By the same token, Jones and Cook (2000) argue that

> an e-journal is a digital periodical that publishes on the Internet or WWW. An e-journal may not be all that different from a print journal in the fundamental editorial process. Articles are submitted by individuals in the academic and practice community, are peer reviewed by editorial board members of the journal to be accepted or rejected, and are subsequently published. It is the digital medium that is different.

Varian (1998) and Sathe, Grady, and Giuse (2002) explain the similarities between ejournals and print journals based on the argument of emulation. According to Varian (1998),

> each new medium has started by emulating the medium it replaced. Eventually the capabilities added by the new medium allow it to evolve in innovative, and often surprising, ways. Alexander Graham Bell thought that

the telephone would be used to broadcast music into homes. Thomas Edison thought that recordings would be mostly of speech rather than music. Marconi thought that radio's most common use would be two-way communication rather than broadcast.

In the case of ejournals, these capabilities are already being acknowledged. Valauskas (1997), for example, recognizes that features such as accelerated peer review, the very quick creation of proofs and final versions, "the sheer interactive nature of digital journals - providing ample opportunities for peers to critically analyze articles - and the ability to access the complete archives of a given title on a server" make ejournals a "significant departure from the long established traditions of print".

Even if pure ejournals emerged before the electronic versions of print journals, at this point there are more of the latter than of the former, inasmuch as almost all print journal titles – 70 to 80 thousand – are online and speculations exist that ejournals will gradually replace print journals (Odlyzko, 1994; LaPorte et al., 1995; Harnad, 1998; Delamothe & Smith, 2001; Bar-Ilan & Fink, 2005). It may not have happened as drastically as Odlyzko (1994, p. 4) predicted, according to whom traditional print journals would be extinct by now and the electronic alternatives would be significantly different from print periodicals. What has been happening is that many medical ejournals have introduced new multimedia sections which only appear in the electronic version, as well as whole articles (*JAMA - Journal of the American Medical Association* for instance) and have also been allowing the online publication of *supplementary data* even for those articles that are also published in the print versions of journals (Figure 2.5). According to Berkenkotter (2007), these extra materials can sometimes be central for the

understanding of an article's argument and, in this sense, the print journal reader would inevitably *have* to access the electronic version of the journal to get the full picture, which means that it is *not at all the same* to use either the print journal *or* the electronic version of a journal.
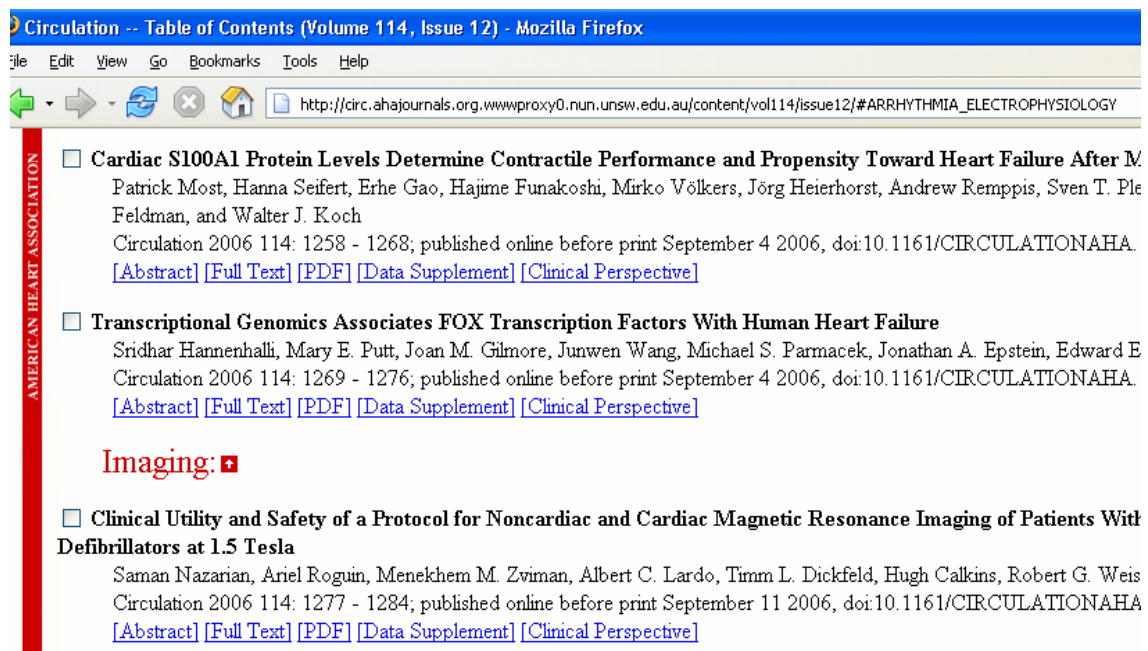
Figure 2.5. Example of Table of contents from the ejournal *Circulation* showing increasing presence of "Data supplements" for online research articles

As argued by Crowston and Williams (2000, p. 203), initially "traditional genres such as the book or academic article have moved intact to the Web" because "in a new situation individuals will typically draw on their existing genre repertoires, reproducing genres they have experienced as members of other communities". For the authors (Ibid.), "these reproductions may be immediately accepted, or there may be a transition period during which the limits of the genre are renegotiated".

Indeed, 15 years after their emergence, ejournals have undergone remarkable changes (Tenopir et al., 2003). Bostock (2006) states that "since 1990, the web has had a profound and still emerging effect on the process of research diffusion" (p. 237).

The reasons underpinning these changes are various. Tenopir et al. (2003) linked them to three evolutionary phases: an *early phase*, which comprises the introductory years of ejournals – from 1990 to 1993; an *evolving phase* "in which a majority of scientific journals are available in electronic format, new features are added to some journals, and some individual articles are made available through preprint archives, author web sites" and an *advanced phase*, "in which searching capabilities, advanced features, and individual articles are integrated in a complete system along with full text of core journals available back to their origin" (Tenopir et al., 2003).

It could be argued that the evolution of ejournals is more directly an effect of the interaction among ejournal users/readers, the electronic medium *per se* and publishers. There has been a great body of studies on usage patterns that support this argument. In short, these studies indicate that in the beginning – early 1990s – not only were ejournals a novelty (for readers as well as for the publishers themselves) but so was the Internet. Once lack of familiarity was overcome, usage increased expressively and users started to identify *problems* with ejournals, to which publishers responded adapting ejournals to users' needs. So, while the first ejournals were pure replicas of print journals, with articles appearing first in print and then being scanned and delivered online as image PDF files (exact photocopy of their print version) and with no added features whatsoever, they soon started to

offer more elaborated PDF versions for their online articles alongside with an HTML format for the same article.

In this regard, Shepherd and Watters (1998, p. 2-3) claim that "when a genre existing in a noncomputer medium migrates to a computer environment, it tends to be faithfully replicated and, initially at least, not to exploit the capabilities of its new medium". Nonetheless, the inherent affordances of the new medium work as a driving force and "these genres typically evolve from an electronic replication of the original, through an initial adaptation to the medium, to genres that fully exploit the new medium and may be recognized only marginally as the original genres" (Ibid., p. 3). That is why current PDFs are no longer replicas or scanned files and their quality is better then the one of a photocopy, showing features such as colorful graphics.

Another unsatisfactory feature of ejournals from the perspective of users was the "breadth or depth of journal coverage" (Tenopir & King, 2002, p. 260-261). As a result, today most journals provide or are on their way to provide back files which, many times, include articles published up to almost two centuries ago, which is the case of *The Lancet* whose articles are available since the journals' foundation in 1823. Finally, the first ejournals users also indicated hyperlinked references that led them to the full-texts of cited articles as an indispensable feature of ejournals (Bostock, 2006; Kwan-Hoong, 2006) and today ejournals include links to most of the cited literature, a practice which has been facilitated by

services such as CrossRef[18] and by the elementary fact that an increasing number of publications exist in digital format (back-file projects).

There are three further practical obstacles that ejournals are to overcome, according to their users: the "poor graphics" used in electronic RAs, i.e., the lack of use of non-printable features – videos, 3-D graphics (Tenopir & King, 2002, p. 260-261); the readability problem which is related to the visual strain caused by computer screens (Pullinger, 1997; Tenopir & King, 2001; Curran, 2002; Bar-Ilan & Fink, 2005); and ejournal portability. All three aspects seem to go beyond publishers' jurisdiction, as the first is dependant on RA authors and the latter two on hardware technology. There has been much incentive on the part of publishers for authors to use non-printable features in their papers but the pace for this to happen is slow. As pointed out before, what many medical journals have introduced are extra multimedia sections – which are not part of research articles but to a certain extent begin to answer these expectations.

Regarding the discomfort caused by continuous screen reading, monitor technology has already improved a lot (compare the quality of LCD – Liquid-Crystal Display – screens with the previous CRT – Cathode-Ray Tube – models) and new technologies are constantly being developed, such as the E Ink (Coburn,

---

[18] "CrossRef is an independent membership association, founded and directed by publishers. CrossRef's mandate is to connect users to primary research content, by enabling publishers to work collectively. CrossRef is also the official DOI® link registration agency for scholarly and professional publications. It operates a cross-publisher citation linking system that allows a researcher to click on a reference citation on one publisher's platform and link directly to the cited content on another publisher's platform, subject to the target publisher's access control practices" (Retrieved on July 2, 2007, from http://www.crossref.org/).

Burrows, Loi & Wilkins, 2001, p. 157)[19], which supposedly produces a "paper-like display" (Sony, 2007).

Sony used E Ink technology on a portable device which they have called *Sony Reader* (Figure 2.6). The development of portable devices such as *Sony Reader* for on-screen reading has also been moving rapidly but it is still in its infancy. Personal Digital Assistants (PDAs), though more popular among the Health sciences (Groote & Doranski, 2004), are gradually "becoming part of the research process" (Sathe, Gravy & Giuse, 2002, p. 242). Medical ejournal publishers are increasingly offering PDA files of RAs, but software and hardware limitations (short battery life, for example) as well as technological literacy deficiencies still need to be sorted out.

It looks like ejournals will continue to meet the demands of its users. The emergence of ejournals has brought many changes and issues for the scholarly community, publishers, and libraries. When ejournals first became available to users, they found it hard to adapt their research habits to this new technology. Now that ejournals have been available for more than a decade[20], it seems that users have reached "a mature stage of usage in which e-journals are recognized as essential tools in research activities" (Jeon-Slaughter, Herkovic & Keller, 2005). And it is reasonable to think that there has been an increase in the acceptance

---

[19] Coburn, Burrows, Loi and Wilkins (2001, p. 158) explain that "a sheet of E Ink is composed of millions of tiny microcapsules, each about the diameter of a human hair". According to Sony (2007), the result of E Ink technology "is a reading experience that's similar to paper – high contrast, high resolution, viewable in direct sunlight and at a nearly 180-degree angle, and requiring no power to maintain the image".
[20] It will be discussed in Chapter 4 that in some contexts, in Brazil for example, ejournals have been accessible for less than a decade (since 2001, when the *Portal Periódicos Capes* project was created).

and frequency of ejournal use just because the traditional print format is not easily available anymore (Bar-Ilan & Fink, 2005, p. 365).
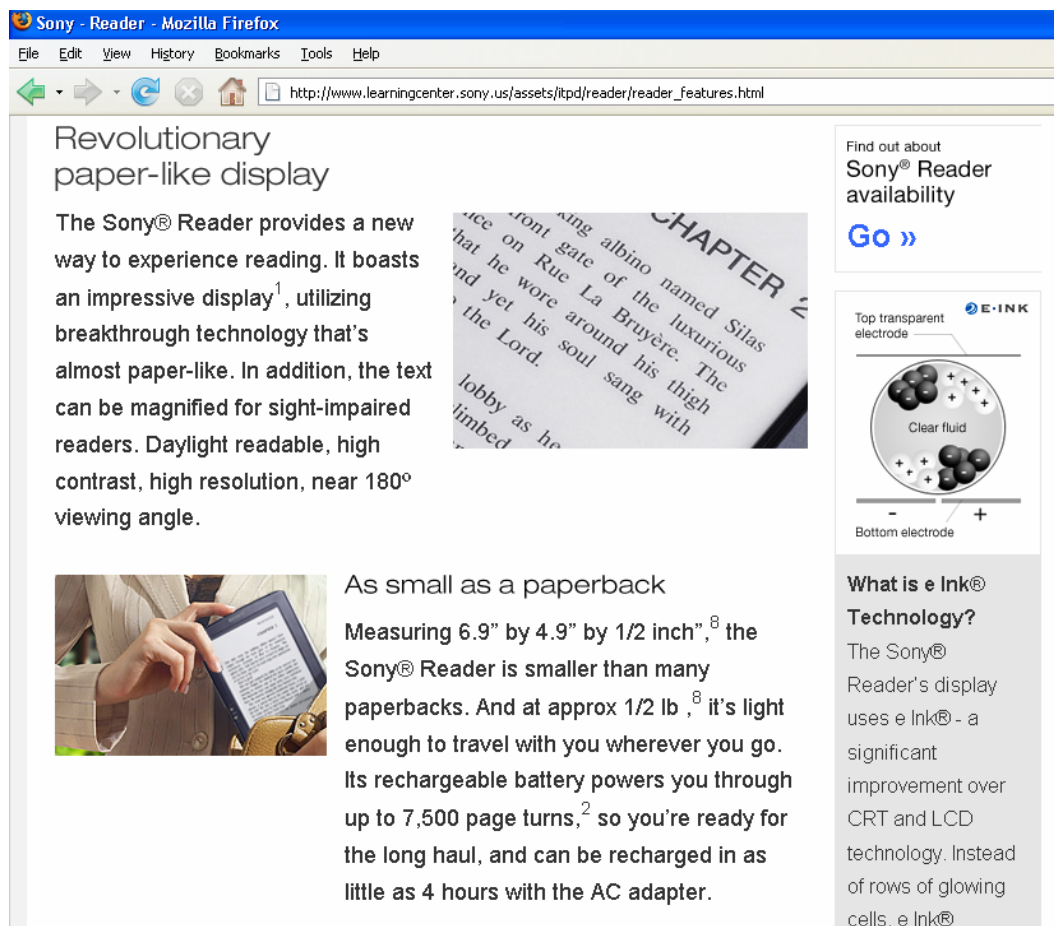
Figure 2.6. Sony Reader: example of technological advance which could help solve the readability and portability problem of electronic research articles

In light of these investigations, it becomes evident that the emergence, evolution and characteristics of ejournals have been widely discussed. It also became evident that ejournals have gained little attention from genre analysts, with a few exceptions (Hendges, 2001, 2002; Motta-Roth & Hendges, 2000; Motta-Roth et. al. 2000; Motta-Roth et al., 2004; Berkenkotter, 2007; Ayers, 2007)

In the following section, therefore, I will discuss studies in the area of genre analysis which have looked at digital genres in order to verify how these genres have been approached and to what extent genre theory has been challenged by digital discourses, given that it was primarily developed based on oral and print written texts.

2.5 Genre analysis and digital genres

The majority of the studies about digital genres cited earlier (beginning of section 2.4) have focused on the analysis of specific linguistic and medium features rather than on the issue of genre classification. In this sense, they have either assumed that the genres they were examining were new genres without undergoing a closer examination or ignored the issue of genre classification for lack of relevance for their particular investigations.

In the case of those analyses of commercial websites, personal homepages, and blogs, for example, questioning genre classification seems less of an issue because these genres do not maintain print counterparts as they were essentially born with the Internet (although having antecedents in print). In this respect, Askehave and Nielsen (2005a, p. 121), in their analysis of homepages, point out that their reasons for choosing this genre were:

> it is a web-generated genre in the sense that it came into existence with the advent of the World WideWeb (WWW) and has no direct parallel outside the web (as opposed to other texts on web sites such as extracts from annual reports, corporate brochures, etc.); and it is among the first web-generated texts to have reached genre status which means that the form and content of the homepage is now becoming conventionalised after more than ten years of 'rhetorical anarchy'.

On the other hand, when it comes to genres which have a print version along with the electronic version, as in the case of newspapers, magazines and journals, the matters of genre labeling and genre classification seem more significant and intriguing. Crowston and Williams (2000, p. 203) believe that

> because the definition of genre relies on social acceptance, it is impossible to define the exact point at which a new genre emerges from the old one. Acceptance may take many years. However, after some period of coexistence, the new combination of form and purpose may become generally recognized and named as a separate genre.

In another article where they examine the effects of linking on genres of web documents, Crowston and Williams (1999) argue that

> only linking that affects the purpose of the document changes the genre of the document; merely dividing a document into pages does not, any more than routine repagination affects the genre of a paper document. Stated alternately, we argue that there is a class of changes to form, namely those related to pagination rather purpose, that do not affect genre.

Apparently what Crowston and Williams (1999) are suggesting is that *purpose* should serve as a guiding criterion for the generic classification of online documents. This proposal had already been presented by Swales in 1990, but was more recently revisited by Askehave and Swales (2001) and Swales (2004) with the argument that determining the purpose of genres is not a straightforward and easy endeavor and that contextual data should be gathered if the communicative purpose of a genre is to be established.

As far as textual analysis of genres is concerned, Askehave and Nielsen (2005a, 2005b) point out that Swales (1990) has indicated three constituent

elements that "capture the essence" (Askehave & Nielsen, 2005a, p. 122) of a genre and allow genre classification: communicative purpose; move structure, and rhetorical strategies (choices of content and style). These would be the criteria of *traditional* genre theory for the analysis of *traditional* genres (Ibid., p. 124). For Askehave and Nielsen (2005a), however, these criteria are not appropriate for the analysis of digital genres inasmuch as they do not account for unique properties of the electronic medium. In their words, "this co-existence of genre and medium, which seems to be ignored in traditional genre theory, is fundamental to web communication and must not be overlooked when trying to determine the genre characteristics of the homepage" (p. 124). In spite of the relevance of medium for genre analysis, the authors warn that they do not mean to suggest that

> the medium always plays a role in genre identification. An uploaded pdf file of an annual report, or an audio-taped novel cannot be said to constitute a new genre. They are still annual reports and novels because the medium does not significantly change the status and function of the genres. However, if the annual report or the novel are published as hypertexts on a web site – and therefore not only provide information about the financial year or tell a story – but also contain links which provide access to other sites on the web, their genre status and function change and thus the media characteristics should be included in their identification.

As a solution that considers the interplay between genre and medium, Askehave and Nielsen (Ibid.) propose a two-dimensional genre model (reproduced in Figure 2.7).
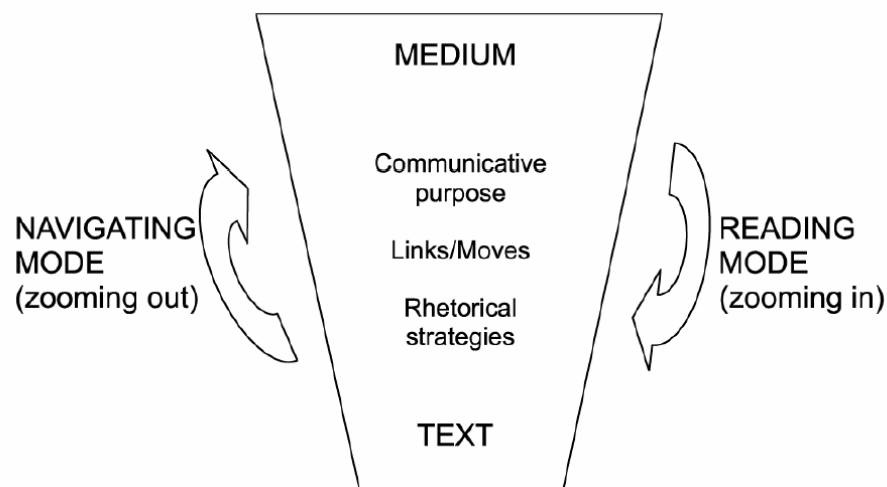
Figure 2.7. Two-dimensional genre model proposed by Askehave and Nielsen (2005a, p. 127)

In this model, the authors introduce the notion of modal-shifts, which consist of analyzing a digital genre in two separate modes: the *reading mode* and the *navigating mode*. The *reading mode* corresponds to the traditional genre analysis methodology in which the digital genre is examined as a print-out in terms of the known criteria of communicative purpose, moves and steps, and content and style. In the *navigating mode*, communicative purpose, structure, and content and style are established based on hyperlinking patterns.

What is not clear in this study is how the results of the two analyses – reading mode analysis and navigating mode analysis – should be integrated so that consistent conclusions about the purpose of the novel digital genre could be drawn and genre classification could be accomplished. Therefore, although Askehave and Nielsen (Ibid.) offer important insights in relation to the need of incorporating the notion of medium into genre analysis, the issue of the appropriateness of traditional genre theory for the analysis of digital genres remains to be clarified. In this thesis I attempt to diminish this gap.

Chapter **3   Methods**

3.1 Introduction

Genre studies have been traditionally classified into 'three schools' (Hyon, 1996; Breure, 2001; Johns, 2004; Hyland, 2002; Paltridge, 2001): the 1) *English for Specific Purposes (ESP)* tradition (inspired by Swales, 1981/1990), the 2) *New Rhetoric* tradition (inspired by Miller, 1984) and the 3) *Systemic Functional Linguistics (SFL)* tradition (inspired by Halliday, 1978, 1985/1989, 1994; and by Martin, 1984/2001). The idea of the three schools was first introduced by Hyon (1996), but there has been some debate on how clear cut the boundaries among them are. For Hyland (2002, p. 114), the three genre schools are "essentially overlapping" approaches; Johns (2004) believes "there's considerable cross-over" among them, especially between the ESP and the New Rhetoric groups; and in Bekenkotter's (2006) point of view, "Hyon's categories have stuck"[21]:

> genre theory (outside of literary theory) is commonly divided into the above three camps--; however, in actuality there is considerable overlap, since no researcher or theorist ever stays in the same place, and many who have been categorized as being in one group or the other have moved on to new interests, and new issues and problems which call for new approaches to genre studies.

---

[21] I would like to thank John Swales for calling my attention to this quote by Carol Berkenkotter (Retrieved August 24, 2007, from http://blog.lib.umn.edu/wardx278/genre/2006/01/north_american_genre_theorists.html) during his plenary talk at the IV SIGET in Tubarão, SC, 2007.

Considering the Brazilian context, Hyon's (Ibid.) categories are certainly outdated for, as pointed out by Motta-Roth (2007), we can identify at least one further very active genre school in Brazil, the socio-discursive interactionism proposed by Bronckart (1999, cited in Motta-Roth, 2007). In this sense, a contemporary map of genre theories would have to contemplate *at least* these *four* schools, not to talk about work developed along the lines inspired by the seminal writings of Bkahtin.

As far as Hyon's (Ibid.) categories are concerned, although being overlapping in relation to some aspects, as in the way they define genre, they are simultaneously divergent in relation to others, as in the methodology they use to analyze genres. This divergence, however, rather than making the approaches contradictory, seems to suggest that they complement each other. For example, it has been argued there is a difference in the discourses they are interested in investigating, in the emphasis each school gives to either context or text, and in the tools each one uses to analyze genres (Hyon, Ibid.; Breure, Ibid.; Hyland, Ibid.). According to Hyon (Ibid.), most of ESP and New Rhetoric genre analysts have been dedicated to academic discourse, while systemicists have examined preferably the genres used in educational – elementary and high school – and professional settings.

In relation to the emphasis each school gives to either context or text, Motta-Roth (2006a, p. 146) explains that Swales's (1990) genre analysis model "was initially received as an essentially linguistic-textual analysis, in which the majority of those who used it tried to replicate, amplify or adapt the already famous

CARS [Creating A Research Space] model". As a result, genre analysts working within this perspective have been accused of being prescriptive formalists who pay little attention to the context, only to linguistic features and structure of texts (Hyon, 1996). Nevertheless, according to Motta-Roth (Ibid.), this tendency of focusing on text as *product* may have changed within the past 15 years to an approach that also considers the context of the text and, therefore, text as *process.*

Similarly, it has been argued that the SFL genre school's interest is in "textual features using schemes of linguistic analysis" (Breure, Ibid.). Ventola (1995) explains that within this tradition, which has its foundations in Firthian linguistics, the procedure of a linguistic analysis suggested by Firth has not been fully followed. According to Firth, a linguistic analysis encompasses three steps: "(a) first the instances of language use as social behavior must be studied, (b) then the theoretical abstractions from the data are made, and (c) finally the theory will be applied to another set of instances for validation" (Ventola, 1995, p.5). She (Ibid.) contends that in SFL steps (a) and (b) have been followed extensively for "instances of language use in social contexts have been studied, and certain theoretical abstractions (...) have been reached" (p. 5), but that step (c) has usually been neglected. Personally I find it ironic that systemicists are more concerned with "text form" and "linguistic features" (Hyon, 1996, p.697) than with context, given the centrality of notion of *context* in Systemic Functional theory, evident in such concepts as Context of Culture and Context of Situation. Even more ironic is the fact that these two concepts were proposed based on *ethnographic* research developed by the anthropologist Bronislaw Malinowski (Halliday, 1985/1989, p. 5) after he realized that texts themselves were not self-

evident. Maybe the systemic functional premise that it is possible to predict the context based on the text and vice-versa (Halliday & Hasan, 1985/1989) has led analysts working within this perspective to be overconfident about the power of texts in revealing social purposes[22].

By contrast, an emphasis on contextual analysis rather than textual has been the characteristic of the New Rhetoricians (Hyon, Ibid.) and, therefore, their genre analysis methods also vary in relation to the ESP and SFL schools:

> in line with their focus on the functional and contextual aspects of genres, a number of scholars in New Rhetoric fields have used ethnographic rather than linguistic methods for analyzing texts, offering thick descriptions of academic and professional contexts surrounding genres and the actions texts perform within these situations. (p. 696)

In this study I try to look at texts in their contexts, following a methodology which I believe to be more closely related to the more recent tendency of ESP genre studies above mentioned by Motta-Roth (2006a). The specific problem I am addressing is genre classification: how can the genre analyst decide whether a text is an exemplar of genre X or Y? How can the genre analyst decide whether a text is a new genre? What criteria can or should be used to classify texts as exemplars of a (new) genre? My main objective is to examine whether the HTML research article can be considered a new genre.

---

[22] Meurer (2004, 2006) criticized the lack of theory in Systemic Functional Linguistics regarding the Context of Culture and in initial effort fulfill the gap he offers a broad scheme to deal with this level of context, based on Giddens's (1984) *structuration theory*.

3.2 Considerations on genre classification

The issue of genre classification[23] has been meticulously discussed by Askehave and Swales (2001). In this remarkable account, the authors (Ibid.) acknowledge that determining whether a text is an exemplar of genre X or genre Y is not a simple and easy task. They (Ibid., p. 196) point out that the key element for determining genre membership in most major contemporary approaches to genre has been *communicative purpose*[24]. Far from being a solution, however, this criterion only further complicates the identification process, for communicative purposes are "complex, multiple, and evasive" (Swales, 2004, p. 69) because they can "evolve, and they can also expand or shrink" (Ibid, p. 73). Maybe this complexity is the reason why not rarely genre studies have taken purpose for granted or based their examinations on "preliminary or received interpretations" (Swales, 2004, p. 74).

For Askehave and Swales (Ibid.),

> most of the important work following the early publications in this field has, in various ways, established that the purposes, goals, or public outcomes are more evasive, multiple, layered, and complex than originally envisaged. How then, as Askehave (1998, 1999) argues, can communicative purpose be used to decide whether a particular text qualifies for a membership in one particular genre as opposed to another? If communicative purpose is typically ineffable at the outset, or only establishable after considerable research, or can lead to disagreements between 'inside' experts and 'outside' genre analysts, or indeed among the experts themselves, how can it be retained as a 'privileged' guiding criterion (Swales 1990)? (p. 197)

---

[23] In Brazil, where the confusion about genre classification still lies between the concept of genre and text type (Motta-Roth, 2006b), a very useful reading is Meurer, J. L. (2003). Genre as diversity, and rhetorical mode as unity in language use. *Ilha do Desterro*, *43*(2), p. 61-82.
[24] In Brazil, Askehave and Swales's (2001) discussion about communicative purpose has been most directly addressed by Hemais and Biasi-Rodrigues (2005) and Biasi-Rodrigues (2007).

The authors (Ibid., p. 207) then suggest two types of procedures for "sorting discourses into generic categories": (a) a "text-first or 'linguistic' approach" or (b) a "context-first or 'ethnographic'" approach (reproduced in Figure 3.1). The novel element in both procedures – which "may of course be combined", according to Swales (2004, p. 73) – is the idea of genre *repurposing* which "acknowledges that sets of texts or transcripts may not be doing what they seem, or not doing what they have traditionally been assumed to have been doing" (Ibid, p. 73).

For that reason, in regards to the text-driven approach, Askehave and Swales (Ibid., p. 207) explain that "the first or earliest uses of *purpose* (in step 1) and *genre* (in step 2) [Figure 3.1 (a)] have been placed in scare-quotes to indicate their provisional status at these stages in the procedure".

(a)

1	Structure + style + content + 'purpose'
	↓
2	'genre'
	↓
3	context
	↓
4	repurposing the genre
	↓
5	reviewing genre status

(b)

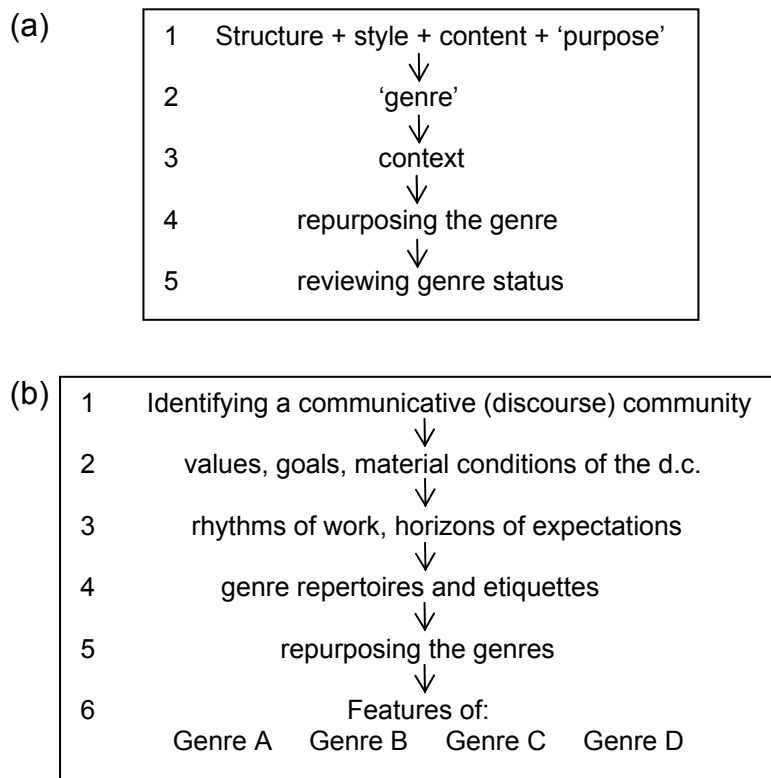|	
1	Identifying a communicative (discourse) community
	↓
2	values, goals, material conditions of the d.c.
	↓
3	rhythms of work, horizons of expectations
	↓
4	genre repertoires and etiquettes
	↓
5	repurposing the genres
	↓
6	Features of:
	Genre A Genre B Genre C Genre D

Figure 3.1. Procedures for genre analysis proposed by Askehave and Swales (2001, p. 207, 208): (a) text-first and (b) context-first

What is interesting to notice from the authors' suggestions is that even though only the context-first approach is explicitly acknowledged as being ethnographic, step 3 - *context* in the text-first approach (Figure 3.1 (a)) clearly predicts that whoever follows this approach will also have to look into the context where the text unfolds. This unavoidable consideration of the context is what is going to help determine if and to what extend a *repurposing of the genre* (step 4) is necessary, because, as Swales (2004, p. 74) indicates, "there will be times when the received interpretation, after investigation, will actually need no modification, but that too is an insight of a valuable kind" . What the authors seem to imply is that when it comes to identifying the purpose of a text, the ethnography-inspired element cannot be eluded.

Regarding the two options offered by Askehave and Swales (Ibid.), the methodology I implemented in this study seems to be more closely related to the text-first approach. I say 'seems to be' because what has really happened in my research is more like a constant back-and-forth movement from text to context and back to text and again back to context than a single direction from text to context. In this way, my procedure would be better represented by a bidirectional arrow (↔) instead of the unidirectional arrow (↓) chosen by Askehave and Swales (Ibid.)[25]. In this sense, my come-and-go analytical process is better captured by the cycle for critical genre analysis suggested by Motta-Roth (2006a, p. 157), which illustrates that it might be necessary during the research to go from interviews (context) to previous literature, from previous literature to text, then

---

[25] Askehave and Swales (2001) did warn the readers that their "schematic depictions" were "idealized and sanitized" (p. 208).

back to the interviews (Figure 3.2). It should be noticed, of course, that while Askehave and Swales's (2001) model is concerned with describing methodologies for genre analysis, Motta-Roth's (2006a) cycle encompasses the whole research process, i.e., theoretical and methodological aspects (that is why the 'theory' element is included in her proposal). What I am trying to say is that, traditionally, when one describes the research method in an experimental paper, one is not expected to mention the literature review as one of the steps of the analysis, because that would be a taken-for-granted part of the process.
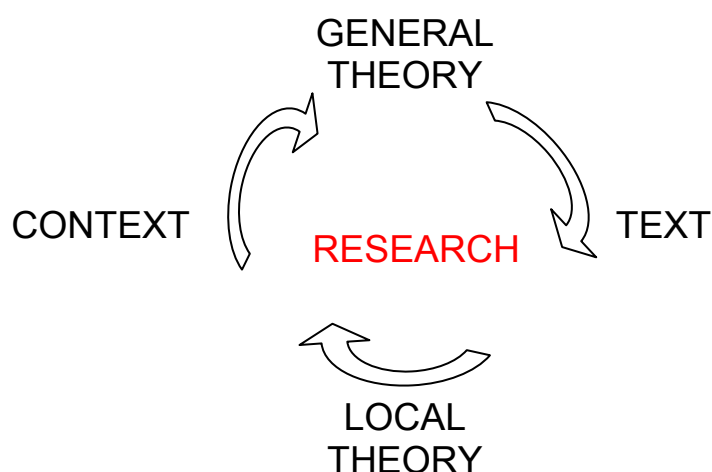
Figure 3.2. "Research cycle for Critical Genre Analysis" proposed by Motta-Roth (2006a, p. 157)

In spite of the fact that my research process followed a back-and-forth directionality, it is still possible to identify four basic stages in which it was conducted: 1. textual survey, 2. contextual data collection and analysis, 3. textual data collection and analysis, 4. genre repurposing. Each one of these stages is detailed in the following sections.

3.3 Textual survey

The first stage of my research consisted of surveying ejournals and specifically HTML research articles to identify how different they were from what we already know about print journals and print RAs as a consequence of being distributed by different communication media – electronic and print.

In practical terms, it involved 'visiting' ejournals in Medicine and Linguistics at random, including current issues and archives, and making unsystematic observations about the configuration and medium-related features of their HTML research articles. The main purpose of this survey was to get familiarized with ejournals and HTML research articles and to learn from the texts what questions were relevant to be asked to researchers in these disciplines.

3.4 Contextual data collection

It was pointed out earlier in this chapter that over the years it has become clear in genre analysis that textual analysis alone may not suffice in the investigation of a genre if we want to have an accurate picture of how that genre functions. In other words, data from the context in which a text is used should be included into the analysis.

In their model (Figure 3.1), Askehave and Swales (2001, p. 208) noted that what counts as *context* "is best left as a 'black box' which can then be operationalized by individual investigators according to their circumstances". In Systemic Functional Linguistics the context that counts for genre analysis is the Context of Culture, where *context* is not equivalent to setting but "a theoretical

construct for explaining how a text relates to the social processes within which it is located" (Halliday, 1991, p. 10) and the meaning of *culture* goes beyond "the traditional life styles, beliefs and value systems of a language community" (Ibid., p. 18). Halliday (*Ibid.*) exemplifies the concept of Context of Culture in the following manner:

> Many years ago, when I was still a language teacher, teaching Chinese at Cambridge University in England, I used to teach a class of scientific Chinese to a group of Cambridge scientists. They wanted to read scientific texts written by Chinese scholars (…). Now, they had no interest in the Chinese culture in the traditional sense of the term; it wasn't necessary for me to teach them anything about Chinese history or family life, or about (…) Confucian values. Did this mean, however, that there was no "context of culture" for my teaching? Of course not. There certainly was a context of culture; and you couldn't hope to learn scientific Chinese without knowing quite a lot about it. But 'culture' here does not mean the traditional culture of China. It means the culture of modern science. (p. 17-18)

Based on Halliday's explanation, it seems that, although terminologically speaking the expression *Context of Culture* may convey the idea that there is only one context for each text, as long as we can identify the effects on a genre's purpose and structure, there can be a countless number of contexts of culture for each text. In the case of my object of study, HTML research articles written in English, for example, I can identify at least four interconnected layers of Context of Culture (Figure 3.3).

A very useful term which captures this interconnectedness of contexts was suggested by Meurer (2004): *intercontextuality*. According to the author (Ibid., p. 86), the term *intercontextuality* is an analogy to intertextuality and interdiscursivity and refers "to the various contexts that intermesh to influence or determine and are influenced or determined by, text, discourses, and other social practices".
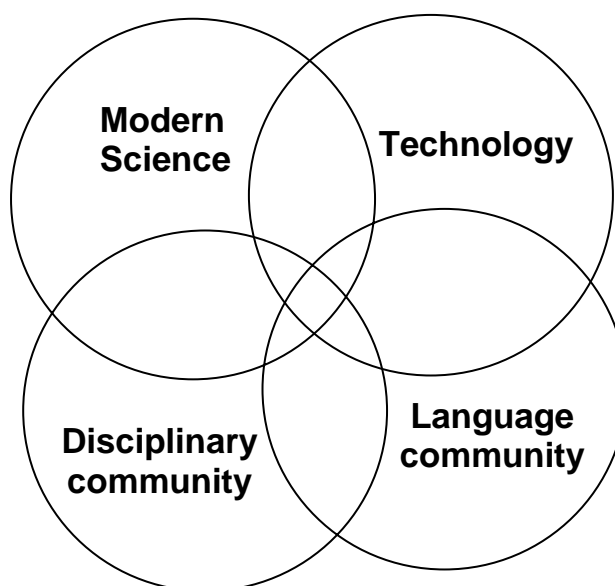
Figure 3.3. Several interconnected Contexts of Culture affecting and being affected by HTML research articles

Following this perspective and in an attempt to collect data from all four contexts mentioned in Figure 3.3, a document-based data collection and analysis and a subject-based data collection and analysis were implemented here. Both stages constitute the *contextual data collection* and are described in the ensuing sections.

3.4.1 Document-based data collection

The document-based data were gathered from a variety of written documents: editorials, instructions for authors and news updates published in ejournals' homepages, documents published in ejournal aggregator platform[26]

---

[26] Aggregator platforms (or aggregator databases) are electronic/online publishers, such as *ProQuest* and *Science Direct*. For a more detailed explanation, see Section 3.5.

websites, and institutional websites, namely, *Australian Bureau of Statistics* (http://www.abs.gov.au), *Comitê Gestor da Internet no Brasil* (http://www.cgi.br), *Internet World Stats* (http://www.internetworldstats.com), *Portal Periódicos Capes* (http://www.periodicos.capes.gov.br/portugues/index.jsp) and *Thompson Scientific* (http://scientific.thomson.com/products/solutions/acad/).

Data obtained from these documents as well as from previous literature were imperative for the elaboration of the online questionnaire, through which the subject-based contextual data were collected.

3.4.2 Subject-based data collection

In the subject-based data collection the central step was the elaboration and distribution of an online questionnaire to researchers in Linguistics and Medicine in two different countries, Brazil and Australia. Furthermore, subject-based data were also obtained from 1. a face-to-face interview with a medical researcher in Australia; 2. email consultations and personal communications with library personnel in Australia[27]; and 3. email consultations with the information technology support teams of two major aggregator platforms in Medicine and Linguistics, respectively: *ProQuest* and *Science Direct*.

The online questionnaires were rendered in two versions, one version in English and one in Portuguese (Appendix A)[28]. For researchers in Brazil, the

---

[27] Face-to-face interview and personal communications with researchers and professionals working in Australia were made possible during my 11 months' stay in Sydney in 2006, as part of my doctoral research (I was awarded by CAPES with a *Programa de Estágio de Doutorado no Exterior (PDEE)* scholarship and enrolled in the *University of New South Wales* as a *Full-Time Practicum Research Student* under the supervision of Dr. Louise Ravelli).
[28] I would like to thank Suzana Cristina dos Reis for developing the online versions of both questionnaires.

Portuguese version was distributed, and for researchers in Australia, the English version. The electronic mail addresses (henceforth *emails*) of linguists in Brazil were obtained from the abstract proceedings of the *III SIGET* (*Simpósio Internacional de Estudos de Gêneros Textuais*, 2005), while the emails of Brazilian medical researchers were obtained at medical departments' websites from different Brazilian universities as well as at the *Lattes* platform (http://lattes.cnpq.br/index.htm). I first tried to collect the emails of Brazilian medical researchers only at medical departments' websites, but I soon realized that many times only the names of the lecturers/researchers were provided in those websites. Their emails (around 150 names) were then collected, one by one, at their online CVs available at the *Lattes Platform*, which ended up being a very labor-intensive process.

Luckily, the faculty and department websites of major Australian universities provide staff's emails and hence all addresses of medical researchers and most emails of linguists in Australia were obtained from these websites. Around 20 of the 72 emails sent to linguists in Australia were also obtained from a private email list which had been created for a course I was enrolled in when in Australia, which were therefore the emails of my colleagues and lecturer. This list was only used after participants allowed me to do so. In fact, in Australia, before distributing the online questionnaire, it had to be approved by the *Ethics Panel* of the University of New South Wales.

The face-to-face interview with a medical researcher in Australia was conducted in 2006 with the then head of the Centre for Vascular Research/Vascular Redox Processes Group of the Faculty of Medicine at the

University of New South Wales. For ethical reasons, I will refer to the interviewee hereafter using the code [MI] (Medical Interviewee). The interview was recorded so that relevant information could be properly quoted. The questions that were used to guide the interview can be retrieved from Appendix B.

Once the contextual data were collected, they were assembled in tables and charts. Simultaneously, the corpus selection process started to take place, given that information obtained in the contextual data collection was used to guide the corpus selection.

3.5 Corpus selection

Before selecting the HTML research articles per se, it was necessary to establish the ejournals from which they were going to be extracted. Both the ejournal selection process as well as the article selection process are described next in separate sections.

3.5.1 Ejournal selection

The HTML research articles that constitute the corpus of the present research were selected from 14 ejournals, seven in Linguistics and seven in Medicine. For the selection of these ejournals three criteria were used: a) status, b) indication, and c) nature of their electronic research articles.

The first criterion, status, was used to select those ejournals with the highest prestige in Linguistics and Medicine. For the Linguistics ejournals, prestige was estimated based on *impact factor* rankings, while for Medicine, prestige was mostly estimated based on previous literature and common sense rankings,

although the *impact factor* figures also contributed for the establishment of ejournal status. The impact factor is a tool for measuring journal quality and reputation and is based on citation frequency. Despite its flaws, it has been used for decades in science (Garfield, 2005, 2006). In this study, therefore, the *2006 Journal Citation Report Social Sciences Edition* was used to help establish which Linguistics ejournals were going to be considered, because there seems to be less consensus in Linguistics than in Medicine regarding the most relevant titles for the area. What happens is that each 'sub-area' linked to Linguistics has its preferred set of titles and, in an attempt to avoid a biased selection, I chose to follow the indication provided by the *Journal Citation Report*. If I had chosen titles which are relevant to my own research interests, I would necessarily have included, for instance, *Applied Linguistics* and *English for Specific Purposes* in my corpus. Unfortunately, they occupy the 18th and 39th places, respectively, in the *2006 Journal Citation Report Social Sciences Edition*.

The 10 journals in Linguistics with the highest impact factors according to the *2006 Journal Citation Report Social Sciences Edition* are: 1. *Journal of Memory and Language*, 2. *Studies in Second Language Acquisition*, 3. *Brain and Language*, 4. *Journals of Speech, Language and Hearing Research*, 5. *Computational Linguistics*, 6. *Language*, 7. *Language and Cognitive Processes*, 8. *Journal of Phonetics*, 9. *Journal of Fluency Disorders*, and 10. *Language Speech and Hearing Services in Schools* (Figure 3.4). It seems important to mention that for the impact factor criterion in Linguistics only those titles classified under the category "Linguistics" in the *2006 Journal Citation Report Social Sciences Edition*

were considered, given that there is no such category as 'Applied Linguistics' in this document.

In relation to the medical ejournals selection, it was not necessary to use the impact factor as a guiding criterion because, according to previous literature, there is already common sense in Medicine about which titles are the most relevant in this area. These titles have been referred to as "the big five" (Kwan-Hoong, 2006), namely, *New England Journal of Medicine*, *The Lancet*, *Journal of the American Medical Association*, *Annals of Internal Medicine*, and *British Medical Journal*. The fact is that this classification as "the big five" is a result of years of high impact factor rankings that these titles have obtained and an examination the *2006 Journal Citation Report Science Edition* (Figure 3.5) will offer evidence that these ejournals have been maintaining their ranking, they are among the 10 medical titles with the highest impact factor. In this sense, the *impact factor* criteria seems to be a more powerful tool in determining journal importance in Medicine than in Linguistics, because such common sense does not prevail in Linguistics.

Nevertheless, during the face-to-face interview with [MI], he called my attention to the fact that, apart from the *New England Journal of Medicine*, the remaining 'big four' do not depict Medicine fully, as they cover only the *clinical* side of the area, leaving *research* Medicine unrepresented. He then cited *Nature*

Journal Citation Reports | GO

Journal Citation Reports®

WELCOME | ? HELP

2006 JCR Social Science Edition

Journal Summary List

Journal Title Changes

Journals from: **subject categories LINGUISTICS** | VIEW CATEGORY SUMMARY LIST

Sorted by: Impact Factor | SORT AGAIN

Journals 1 - 20 (of 47) | [1 | 2 | 3] | Page 1 of 3

MARK ALL | UPDATE MARKED LIST

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to journal information)	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-life
☐	1	J MEM LANG	0749-596X	4186	2.827	0.574	68	8.9
☐	2	STUD SECOND LANG ACQ	0272-2631	742	2.417	0.529	17	8.0
☐	3	BRAIN LANG	0093-934X	4339	2.317	0.455	88	8.7
☐	4	J SPEECH LANG HEAR R	1092-4388	2423	1.801	0.163	98	6.0
☐	5	COMPUT LINGUIST	0891-2017	834	1.800	0.188	16	>10.0
☐	6	LANGUAGE	0097-8507	1467	1.786	0.300	20	>10.0
☐	7	LANG COGNITIVE PROC	0169-0965	1000	1.500	0.706	34	9.2
☐	8	J PHONETICS	0095-4470	851	1.487	0.304	23	>10.0
☐	9	J FLUENCY DISORD	0094-730X	359	1.387	0.059	17	8.3
☐	10	LANG SPEECH HEAR SER	0161-1461	539	1.314	0.483	29	7.6

Figure 3.4. Highest impact factor journals in Linguistics according to the *2006 Journal Citation Report Social Sciences Edition*

Journal Citation Reports ▾ | GO

Journal Citation Reports®

🏠 WELCOME | ❓ HELP

2006 JCR Science Edition

📖 **Journal Summary List**

Journal Title Changes

Journals from: subject categories MEDICINE, GENERAL & INTERNAL 🔲 VIEW CATEGORY SUMMARY LIST

Sorted by: Impact Factor ▾ SORT AGAIN

Journals 1 - 20 (of 103)

⏮ ⏪ ◀ [*1* | 2 | 3 | 4 | 5 | 6] ▶ ⏩ ⏭

Page 1 of 6

MARK ALL | UPDATE MARKED LIST

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to journal information)	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-life
☐	1	NEW ENGL J MED	0028-4793	177505	51.296	12.743	303	6.9
☐	2	LANCET	0140-6736	133932	25.800	7.419	301	7.4
☐	3	JAMA-J AM MED ASSOC	0098-7484	100317	23.175	7.781	270	6.7
☐	4	ANN INTERN MED	0003-4819	39609	14.780	3.720	164	8.6
☐	5	PLOS MED	1549-1277	1998	13.750	3.419	148	1.4
☐	6	ANNU REV MED	0066-4219	3632	13.237	4.343	35	6.0
☐	7	BRIT MED J	0959-8146	61517	9.245	4.412	413	7.8
☐	8	ARCH INTERN MED	0003-9926	29480	7.920	1.846	286	7.0
☐	9	CAN MED ASSOC J	0820-3946	7724	6.862	3.044	113	6.2
☐	10	MEDICINE	0025-7974	4524	5.167	0.281	32	>10.0

Figure 3.5. Highest impact factor journals in "Medicine, General & Internal" according to the *2006 Journal Citation Report Science Edition*

70

Medicine as one of the "biggest" journals in *research* Medicine, and *Circulation* as one of the most relevant research medical journals in his specific sub-field, Vascular Medicine. According to [MI], the differences between *clinical* Medicine and *research* Medicine can have an impact on how ejournals are configured as well as on ejournal usage practices.

In fact, as shown in the *2006 Journal Citation Report Science Edition*, there are six categories which explicitly refer to Medicine (Figure 3.6) in their labels. Two of them are particularly related to the issue pointed out by [MI]: "Medicine, General & Internal", encompassing *clinical* Medicine, and "Medicine, Research & Experimental", encompassing *research* Medicine.

Figure 3.6. Different journal categories in *2006 Journal Citation Report Science Edition* referring to Medicine

In "Medicine, General & Internal", the big five can be found (Figure 3.5), and in "Medicine, Research & Experimental", *Nature Medicine* figures as the highest impact factor journal (Figure 3.7). *Nature Medicine* and *Circulation* were then considered for the corpus of medical journals based on the second criterion for selecting the ejournals: indication.

71

Figure 3.7. Highest impact factor journals in "Medicine, Research & Experimental" according to the *2006 Journal Citation Report Science Edition*

The third and last criterion used for ejournal selection is the nature of the electronic research articles provided. This criterion was used to eliminate from the sample those ejournals which did not present an HTML version of their articles, only a PDF version. Considering that it is the HTML version that is the object being investigated here, the titles necessarily had to offer this version. In Medicine all seven journals under consideration presented at least two versions of their articles, an HTML and a PDF version. In Linguistics, however, two out of the 10 highest impact factor titles (Figure 3.4) included a PDF article version only: *Computational Linguistics* and *Language*. These were then excluded from the sample, which was then composed of eight titles.

Given that seven titles had already been established for Medicine, in search for a balance between the areas I decided to eliminate one more title from the Linguistics sample, *Language Speech and Hearing Services in Schools*, which occupied the eighth rank in the impact factor list. Based on these criteria, the final list of ejournals from which the HTML research articles in the corpus were selected includes 14 titles, enumerated in Table 3.1 with their respective impact factors.

Once the ejournals were delimitated, it was necessary to establish criteria for selecting the HTML research articles.

3.5.2 HTML research article selection

The main aspect taken into consideration for the selection of the articles in the corpus was the year of their publication, i.e., only those published between 2002 and 2006 were selected.

Table 3.1
List of the ejournals in Linguistics and Medicine considered in the investigation

		Title	Impact Factor	HTML Version
Linguistics	1	*Journal of Memory and language*	2.827	+
	2	*Studies in Second Language Acquisition*	2.417	+
	3	*Brain and Language*	2.317	+
	4	*Journal of Speech Language and Hearing Research*	1.801	+
	5	*Language and Cognitive Processes*	1.500	+
	6	*Journal of Phonetics*	1.487	+
	7	*Journal of Fluency Disorders*	1.387	+
Medicine	1	*New England Journal of Medicine*	51.296	+
	2	*Nature Medicine*	28.588	+
	3	*The Lancet*	25.800	+
	4	*Journal of the American Medical Association*	23.175	+
	5	*Annals of Internal Medicine*	14.780	+
	6	*Circulation*	10.940	+
	7	*British Medical Journal*	9.245	+

The initial plan was to select one HTML research article per year from each one of the 14 ejournals, with a final amount of 70 articles, 35 from Linguistics and 35 from Medicine. The total number of HTML research articles that constitute the corpus, however, is 112, because the count could not be made by the number of ejournal titles, but by the number of *aggregator databases* which publish each ejournal. Some ejournals in the corpus are published and distributed online by more that one *aggregator database*. An aggregator database (or *aggregator platform*, or simply *aggregator*) is a web service hired by one or more print journal publishers via contractual agreements to transmit journal content on the Internet. Examples of aggregator databases are *Gale Group, Ovid, ProQuest, Science Direct,* and *Wilson*. If we take the *Journal of Phonetics,* for example, in print it is published by *Elsevier* but the online publication is made by *Science Direct*, an aggregator database hired by *Elsevier*. The advantage for a print publisher in allowing its journals to be delivered through aggregator databases is the "increased exposure or greater

market penetration for their journals" (Chambers & SooYoung, 2004, p. 188). In this sense, the greater the number of aggregator databases by which an ejournal is published online, the higher the chances of it reaching a greater amount of users/readers. According to Chambers and SooYoung (Ibid., p. 189), today aggregator databases have become "central gateways to the world of journal literature".

The kind of access one has to an ejournal will thus depend on the kind of deals an institution made with aggregator databases[49]. In this study, ejournals were accessed via the *University of New South Wales*, Australia. The number of aggregators for each ejournal in the corpus provided by the *University of New South Wales* is shown in Table 3.2.

Each aggregator database has its particular 'online style sheet' for publishing its HTML research articles.

Table 3.2
Number of aggregators per ejournal accessible via the University of New South Wales

		Title	No. of Aggregators
Linguistics	1	*Journal of Memory and language*	2
	2	*Studies in Second Language Acquisition*	1
	3	*Brain and Language*	1
	4	*Journal of Speech Language and Hearing Research*	4
	5	*Language and Cognitive Processes*	2
	6	*Journal of Phonetics*	1
	7	*Journal of Fluency Disorders*	1
Medicine	1	*New England Journal of Medicine*	2
	2	*Nature Medicine*	2
	3	*The Lancet*	4
	4	*Journal of the American Medical Association*	2
	5	*Annals of Internal Medicine*	3
	6	*Circulation*	2
	7	*British Medical Journal*	3

---

[49] This issue will be discussed and exemplified in Chapter 5 – Contextual data analysis and discussion.

Therefore, the higher the number of aggregator databases for one ejournal, the bigger the number of versions of HTML research articles published in that journal. For example, the medical ejournal *Annals of Internal Medicine* is available through three different aggregators (Table 3.2) if accessed via the *University of New South Wales*, namely, *American College of Physicians*, *Ovid* and *ProQuest*. Given that each aggregator database follows its own online style sheet for publishing its HTML research article, the same article can be available in three different HTML designs (Figure 3.8).

In this sense, the number of HTML research articles collected for the corpus was dependant on the number of HTML designs available for each article between the years 2002 and 2006 via the University of New South Wales, because all HTML versions were taken into account in my analysis.

a) *American College of Physicians-ASIM*

b) *Ovid*

c) *ProQuest*

Figure 3.8. Three different HTML designs for the same research article

Another factor affecting the article collection was that some Linguistics ejournals considered for corpus selection have not had HTML versions for the

research articles along the same period as they have had PDF versions. That is the case of *Studies in Second Language Acquisition*, which started HTML publication only in 2004, while PDF versions of articles date back to 1997; *Language and Cognitive Processes*, which started HTML publication only in 2005, with PDFs available back to 1996; and *Journal of Phonetics*, which started HTML publication in 2003 and PDFs in 1995.

Within these constraints and multiplicity of ejournal publication practices, the total number of HTML research articles selected from each one of the 14 ejournal previously established is considerably variable (Table 3.3), ranging from 2 (*Language and Cognitive Processes*) to 15 (*Journal of Speech Language and Hearing Research* and *The Lancet*) HTML research articles per ejournal.

Table 3.3
Number of HTML RAs collected from each ejournal

		Title	No. of HTML articles
Linguistics	1	*Journal of Memory and language*	5
	2	*Studies in Second Language Acquisition*	3
	3	*Brain and Language*	5
	4	*Journal of Speech Language and Hearing Research*	15
	5	*Language and Cognitive Processes*	2
	6	*Journal of Phonetics*	4
	7	*Journal of Fluency Disorders*	5
	Total		**39**
Medicine	1	*New England Journal of Medicine*	10
	2	*Nature Medicine*	5
	3	*The Lancet*	15
	4	*Journal of the American Medical Association*	8
	5	*Annals of Internal Medicine*	13
	6	*Circulation*	10
	7	*British Medical Journal*	10
	Total		**71**

Once the corpus was collected, the HTML research articles were examined in terms of their schematic structure, hyperlink usage, and communicative purpose. The analytical criteria for the textual examination are explained in the next section.

3.6 Textual data collection and analysis

In their text-first methodological approach (discussed at the beginning of this chapter), Askehave and Swales (2001, p. 207, 208) suggest that the first step in textual analysis is to examine the "Structure + style + content + 'purpose'" of the genre. Drawing on Systemic Functional Linguistics (SFL), I propose *style* and *content* to be seen here as features of register (Martin, 1992), realized by interpersonal and ideational meanings, respectively. In this perspective, the aspects that remain to be analyzed are *structure* and *purpose*. Considering that "what is immediately manifest to the genre analyst is not purpose but form" (Askehave & Swales, 2001, p. 200), the 112 HTML research articles in the corpus were initially analyzed in terms of their schematic structure. For this analysis, Swales's (1990) *moves and steps* model was used. It was expected that the results obtained in the structural analysis would contribute for the identification of the purpose of the HTML research articles and, therefore, help to determine whether they can be considered a new genre or not, Step 2 - *genre* in Askehave and Swales's (Ibid.) proposal (Figure 3.1 (a)).

It soon became evident that the *moves and steps* analysis alone did not suffice because it could not capture the hypertextual nature of the HTML

research articles, such as the presence of hyperlinks. As part of the second step in the textual analysis, therefore, the nature, function and frequency of the hyperlinks used in the HTML research articles was examined. The point of departure for the hyperlink analysis was the hyperlink classification offered by Landow (1997) as well as the two-dimensional digital genre analysis model proposed by Askehave and Nielsen (2005a, 2005b).

The results obtained in this textual analysis were described and discussed in relation to each one of the two disciplines being investigated – Linguistics and Medicine – as well as a whole. In this sense, disciplinary similarities and differences were highlighted.

Once the results of the textual analysis were available, a tentative set of purposes for the HTML research article was proposed and then tested based on the data from the analysis of the context. This testing corresponds to the *genre repurposing* (Step 4) process proposed by Askehave and Swales (Ibid.) (Figure 3.1).

3.7 Genre repurposing

At this level of the investigation, the results obtained in the textual analysis were cross-checked with the data collected in the analysis of the context in an attempt to verify whether it was necessary to repurpose the genre or not (Step 4 in Askehave & Swales, 2001, p. 207) (Figure 3.1). At this point in the examination, therefore, the communicative purpose(s) of the HTML research article was (were) established and, hence, its generic classification could be concluded.

During this stage of the analysis, disciplinary idiosyncrasies were also taken into consideration, as well contextual differences in terms of the countries from where the contextual data were collected, Brazil and Australia.

Chapter 4 Contextual analysis

4.1 Introduction

As described earlier in Chapter 3 – Methods, the analysis of the context in this study consisted of two core moments: 1) document-based data collection and analysis and 2) subject-based data collection and analysis. In addition, it was also pointed out that previously to the contextual data collection a textual survey was carried out so that basic parameters for the contextual data collection could be established.

Given that it consisted of an unsystematic and randomized analysis of HTML research articles, the textual survey does not seem to deserve a chapter on its own and therefore, the main outcomes of this stage are presented here as follows.

The survey considered articles published in different years, from 1997 until 2006 and revealed that HTML research articles:

- had changed significantly over time;
- could appear in a variety of layouts;
- were more commonly available in medical ejournals than in Linguistics ejournals;
- presented a growing number of hyperlinks from 1997 to 2006;
- presented hyperlinks of various natures;
- never showed images (figures, tables, charts, etc.) in the homepage of the articles, i.e., the main entrance page when HTML research articles are accessed via ejournals' tables of contents;

It was also noticed the same RA in print could be available online in a variety of HTML configurations, depending on the number of aggregator databases that publish it online (see section 3.5 in Chapter 3 for details on this matter). In light of these observations, an online data collection instrument containing 10 questions was developed and distributed to researchers in the areas of Linguistics and Medicine. I have called this stage *subject-based data collection* and the results obtained from the analysis of these data constitute the essence of the present chapter. However, before introducing and discussing the subject-based data, results from the document-based data analysis will be shown.

4.2 Document-based data analysis

The document-based data collection comprised the examination of documents available at a variety of institutional websites (e.g., *Australian Bureau of Statistics - *http://www.abs.gov.au*, Instituto Brasileiro de Geografia e Estatística*, http://www.ibge.gov.br/home*, Comitê Gestor da Internet no Brasil -* http://www.cgi.br*, Internet World Stats -* http://www.internetworldstats.com*, Portal Periódicos Capes -* http://www.periodicos.capes.gov.br/portugues/index.jsp*, Thompson Scientific -* http://scientific.thomson.com/products/solutions/acad/*) and it was developed as a means of obtaining information about elements of the context that affect ejournal publication practices, which in turn could have an impact on how and why eRAs are produced and consumed as well as on what eRAs look like and why they look like that.

The contexts under consideration in this research are the countries of Brazil and Australia and the disciplinary contexts of Linguistics and Medicine. Regarding Brazil and Australia, it seemed relevant to consider technology-related factors, such as Internet access and ejournal access, which could have an impact in eRA writing and reading practices. Particularly with regards to the latter, my hypothesis is that the more familiarized a user is with the Internet, the more developed his/her online reading skills are and the more s/he will benefit from the medium's affordances.

As far as the disciplinary contexts are concerned, it has already been pointed out in the previous chapters that the role print journals have played in Linguistics is less central than their role in Medicine. Specifically in the case of HTML research article availability, in Chapter 3 – Methods it was observed that more HTML formats for the same articles are provided in Medicine than in Linguistics. I expect, therefore, that the document-based search will offer evidence that explains this and further differences between the disciplines, particularly in relation to how HTML research articles are instantiated in each area in terms of structure and hypertextuality.

4.2.1 Internet access in Brazil and in Australia

In March of the current year, *Instituto Brasileiro de Geografia e Estatística* (IBGE, 2007), Brazil's bureau of statistics, published a set of data about Internet access in Brazil during 2005. Expectedly, the data show how inseparably related Internet access and income rates are in our country. Out of the 21% of Brazilians who reported having access to the Internet, 69,5% earn more than 5 minimum wages (IBGE, 2007) and 50% have access to it from home.

The data also reveal that for the majority (37,2%) of the 78,6% of the people who do not use the Internet, it is the very lack of access that impedes them to use it. Internet access costs and free internet access are indeed still troublesome matters in Brazil. This is no secret for those working in public educational institutions, where insufficient and inadequate equipment as well as digitally illiterate teachers have been an obstacle to increase access. The issue does not affect primary and secondary schools only, but also universities. At least as far as teacher education degrees[50] are concerned, it is very common to have undergraduates reporting lack of Internet access, given that these are the degrees that mostly absorb students coming from lower income households who do not have access to the Internet from home. Although access on campus is provided, most of the times the number of computers available does not suffice and, apart from that, the fact that many of the computers available are obsolete and malfunctioning (no USB ports, damaged CD and floppy disk drivers) and that most of them are not connected to a printer can make the whole Internet access process counterproductive. According to *Comitê Gestor da Internet no Brasil* (2006), the governmental committee which collects, organizes and disseminates Internet-related statistics, the number of Brazilian households with computers in 2006 was 19,6%, out of which only 14,5% had Internet access (Comitê Gestor da Internet no Brasil, 2006).

---

[50] Particularly in the case of the *Universidade Federal de Santa Maria*, many of the students that take teacher education (e.g., Geography, History, Languages, Mathematics, Philosophy) are those students which did not have the financial conditions to afford preparatory courses for the university's entrance examination – Vestibular, which would allow them to apply for degrees such as Medicine, Electrical Engineering, Dentistry, Architecture, for which the entrance competition rates are much higher than for teacher education degrees.

All these factors, lack of access, high costs of band width, local of access, obsolete equipment, generate a very contrasting scenario between Brazil and Australia as far as Internet use is concerned (Figure 4.1).

While the Internet penetration rate in Brazil is 21% (IBGE, 2007; Internet World Stats, 2007a), which renders it the 62nd place in the world, in Australia the same rate is 71,9% (Internet World Stats, 2007b; Brendler, 2007). This figure has rendered Australia the 6th place in a list of the top 36 countries worldwide where over 50% of the population uses the Internet (Internet World Stats, 2007c). As reported by the Australian Bureau of Statistics (2006), Australians depend less on outside home access than Brazilians, because "in 2005-06, 60% of Australian households had home Internet access and 70% had access to a home computer".
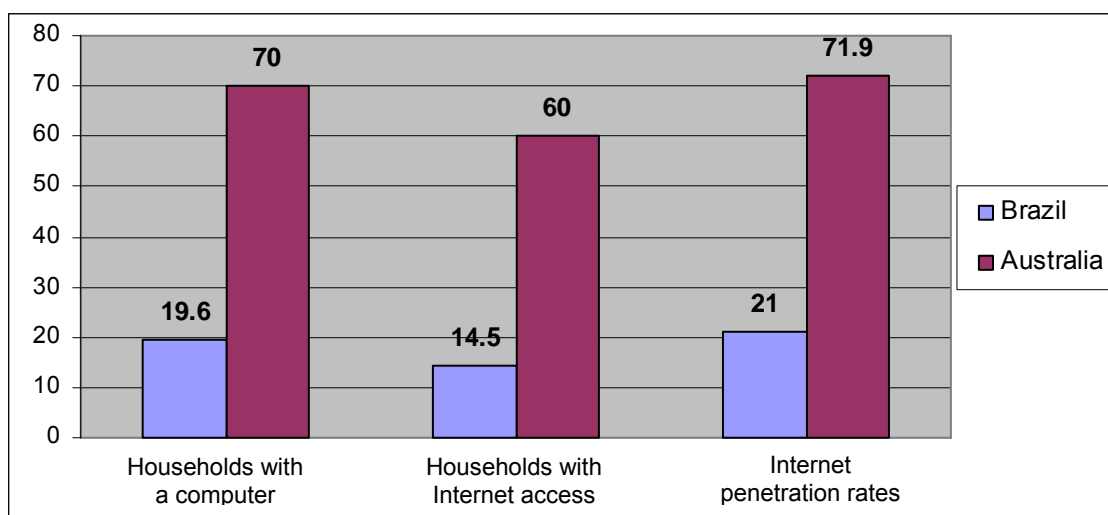
Figure 4.1. Internet access rates in Brazil and Australia

As far as universities are concerned, since there are no state universities in Australia, part of their budget is dedicated to information technology. At the

beginning of 2006 while I was at the UNSW, all *old* computers, though working well, were replaced by new models.

My recent experience with using *Portal Periódicos Capes* showed that undergraduate students are widely unaware of the service and have difficulties locating information they are required to. My informal observations have also shown that the less familiar the student is with the Internet, the more difficulties s/he has in navigating through the web.

Access issues as well as ejournal access factors seem to be influencing these practices. Regarding ejournal access, the figures across the two countries undertaken here are equally discrepant, as shown in the next sub-section.

4.2.2 Ejournal access in Brazil and in Australia

Two of the major differences between Australia and Brazil regarding ejournal access are the amount of time they have been available in each country and the number of titles available. In both cases, Australia seems to have an advantage over Brazil.

In Australian universities, access to ejournals was available considerably earlier (hence, longer) than in Brazilian universities. The former implemented ejournal collections in the late 1995's[51] and early 1996's, i.e., when paid journals were giving their first steps towards the online medium. In Brazil, however, ejournals became available only five years later, in November 2000, when the *Portal Periódicos CAPES*[52] was created. To be more precise, before that, since 1998, we actually had another electronic ejournal collection

---

[51] Specifically at University of New South Wales the e-journals collection started in 1995 (Snell, 2006, personal communication).
[52] Retrievable via the URL (Uniform Resource Locator) http://www.periodicos.capes.gov.br.

operating in Brazil, the *Scientific Electronic Library Online – SciELO*[53]. Nevertheless, the central objective of this platform was to give visibility and encourage access to national publications only, and thus, most of the ejournals available through *SciELO* are national titles and some are from other Latin American and Caribbean countries.

Besides this discrepancy in the amount of time ejournals have been available in Brazil as opposed to Australia, the other element which could be affecting the way researchers in Brazil use ejournals as opposed to Australian researchers is the number of titles available in each context. In November 2000, *Portal Periódicos CAPES* offered access to 1,400 ejournals (Reis, 2005, p. 38). Today, the number of titles has increased significantly to 11,419 titles (CAPES, 2007). However, when compared to the number of ejournals Australian universities offer access to, the Brazilian figures still seem very low, because at the UNSW, there are around 23,000 ejournals available (in 2006); at the University of Sydney, 68,130 (in 2005); at the University of Technology of Sydney, 28,000 (in 2005); and at Macquarie university, over 30,000 (2007), an average of 37,000 titles, more than three times as much as major Brazilian universities have access to via *Portal Periódicos CAPES*.

Australian universities, therefore, made agreements with more ejournal aggregators than CAPES and the aftermath of that, besides the higher number of titles available for researchers in Australia, is that the same journal when accessed via the UNSW, for example, can be accessed through different aggregators, allowing UNSW ejournal users more choices than researchers in

---

[53] Retrievable via the URL http://www.scielo.br.

Brazil. This issue can be more clearly explained using the medical journal *The Lancet* as an example.

It was stated in the Chapter 3 – Methods that a publisher can use one or more aggregators to deliver a journal online. That is the case of the medical journal *The Lancet*, which is published in print by Elsevier, but whose online version can be accessed through four different aggregators (if accessed via the UNSW): *Science Direct*, *GaleGroup*, *Ovid*, and *ProQuest*. This number of aggregators, however, can vary from one institution to another, because it depends on the kind of deals that were made with aggregators, i.e., which kind of consortia has been established. If *The Lancet*, for instance, is accessed via the *Portal Periódicos CAPES*, it will be available through two aggregator databases (Figure 4.2(a)); if accessed through Argentina's *Biblioteca Electrónica de Ciencia y Tecnología*[54], only one option is available (Figure 4.2(b)), contrasting with the four alternatives offered by both the UNSW electronic library and Harvard University library, for instance (Figures 4.2(c) and (d)). As can be noticed based on Figure 4.2, while in Australia and in the United States of America each university has its own electronic collection accessible through their own libraries, in Argentina 52 institutions use the same collection and in Brazil 163 institutions do so. These access differences can influence the way researchers in these countries conceive and use electronic journals, specifically their use of HTML research articles.

---

[54] The 'Argentinean version' of *Portal Periódicos CAPES*. It was created in 2002 as the result of an agreement between Argentina and Brazil and is hosted at the *Secretaría de Ciencia, Tecnología e Innovación Productiva – SECYT*'s webapage (retrievable via the URL http://www.biblioteca.secyt.gov.ar/), which is part of the *Ministerio de Educación, Ciencia y Tecnología de la República Argentina*.

(a) *Portal Periódicos CAPES*

(b) *Biblioteca Electrónica de Ciencia y Tecnología*

(c) Harvard University

(d) *UNSW Library*

Figure 4.2. Contrast among the number of aggregator databases available via different institutions for the e-journal *The Lancet*

90

The reason why I am calling attention to these differences is that these elements, namely type of access, duration of access, navigation facilities, aggregator website interface design, and availability are factors which can affect ejournal usage practices. I believe that these aspects can affect the degree of familiarity researchers in each country have with ejournals, which, if low, could have a negative impact on the scientific productivity of a country.

In the following section, when considering how linguists and medical researchers in Brazil and in Australia use ejournals, I will keep these differences in mind.

4.3 Subject-based data analysis: how ejournals are used by linguists and medical researchers in Brazil and in Australia

The total number of emails – with the research questionnaire – sent to linguists and medical researchers in Brazil and in Australia was 760, out of which 97 (13%) were not delivered because the target email address was no longer operative. The total number of questionnaires *delivered* was, therefore, 663, but only 53 (8%) were answered, which was no surprise to me given that, at least in Linguistics, data collection of this nature usually obtains low feedback rates. The number of answers per country and per discipline is shown in Table 4.1.

The percentage of answers could have been even lower were many of the Linguists not my colleagues and acquaintances. As for Medicine in Australia, at least 50% of answers (Table 4.1) were obtained thanks to the generous interference of my interviewee [MI], who asked his research team to answer the questions. Considering that the ethnographic component has not

always been acknowledged as an unavoidable step in genre analysis, genre analysts (including myself) may need some training in terms of strategies for successful ethnographic data collection.

Table 4.1
Total number of respondents per country and per discipline

	Australia			*Brazil*			*Total*		
	Delivered	Answered		Delivered	Answered		Delivered	Answered	
	N	N	%	N	N	%	N	N	%
Linguistics	63	11	17	197	17	9	260	28	11
Medicine	160	10	6	243	15	6	403	25	6
Total	223	21	9	440	32	7	663	53	8

In the present research, therefore, data discussion and conclusions will be drawn based on the answers of 53 subjects from two specific disciplines and from two countries. According to Swales (2007), this sample can be considered sufficiently representative for making generalizations about HTML research articles. Still, it seems important to keep these aspects in mind during the analysis, discussion and conclusions achieved during this investigation.

Table 4.1 shows that the feedback rate from linguists was significantly higher (17%) in Australia when compared to the feedback rate from Australian medical researchers (6%). Similarly, in Brazil more linguists (9%) participated in the data collection, although here the difference is not that striking when opposed to Medicine (6%). As pointed out above, solidarity and *compassion* could be reasons why more linguists answered the questionnaire, because many of them are either acquaintances or friends of mine and, in knowing that this kind of data collection method usually gets low response rates, they

decided to literally help me out before considering that their contribution could be as beneficial for science in general. That being the case, it is crucial to highlight that none of them was involved with the research process to the extent to produce biased answers.

Regarding age, the questionnaire also revealed a disciplinary contrast (Table 4.2). As much as 50% of respondents in Linguistics are younger than 35 years old (14 out of 28 are between 25 and 34 years old), while most participants from the field of Medicine range from 45 to 54 years old, precisely 56% (14 out of 25). I know the age range of most of the linguists to whom I sent my questionnaire and therefore I know that the majority of them are over 35 years old. Still the highest feedback rate was from the younger *linguists* and, therefore, I am again inclined to think these numbers are indeed a result of that feeling of partnership and solidarity already mentioned, because that is the age group I fit in and some of them had been my classmates in Brazil and in Australia. This group has also been involved with research more recently and is, thus, more aware of the need to contribute with data collections of this nature.

Table 4.2
Questionnaire feedback according to age group and discipline

	Australia		*Brazil*		*Total*		Total
	L	M	L	M	L	M	
Under 25	−	−	1	−	1	−	1
25-29	4	−	4	−	**8**	−	**8**
30-34	4	1	2	2	**6**	3	**9**
35-39	1	−	4	3	5	3	8
40-44	−	−	1	2	1	2	3
45-49	1	3	2	4	3	**7**	**10**
50-54	1	4	2	3	3	**7**	**10**
55 or older	−	2	1	2	1	4	5
Total	11	10	17	15	28	25	**53**

As for the medical researchers to whom the questionnaire was sent I would not be able to tell whether most of the questionnaires were sent to subjects over 45 years old or not, but it could be speculated that the majority of respondents are older because most of the questionnaires were sent to older researchers only.

Whether this disciplinary variation in terms of age will have an impact on HTML research articles' usage practices or not will only become clear after the other answers to the questionnaire have been analyzed. This will be done in the following sections, where the results obtained in 5 of the 10 questions are presented and discussed. Questions 6, 7, 8, 9 and 10 (Appendix A) will not be discussed in the present work because of lack of time and because the data obtained in these questions seem to be less vital for the argument presented in this thesis, which aims at determining the communicative purpose of HTML research articles from the perspective of their users.

4.3.1 Question 1 - Where do you access academic journal articles?

In question 1, participants were asked to indicate how they access journal articles and to do so they were given six options: personal subscription of journals, group subscription of journals, with peers, library (print journal collections), electronic library (online journals), and one final alternative which was left open so that they could indicate further ways through which they access articles. The purpose of this question was twofold. First, it aimed at verifying to what extent the technological contrasts between Brazil and Australia, pointed out at the beginning of this chapter, may interfere with the

way researchers access journals in each country. Secondly, it seeks to uncover if there is disciplinary variation when it comes to ejournals, given that in Medicine journals are a much more important source for the dissemination of science than in Linguistics and have, as a consequence, developed much faster and moved to the Internet earlier than Linguistics journals, which could cause medical researchers to be more familiar with ejournals and use them more frequently than linguists.

Table 4.3 shows the numbers and percentages for the most frequent way to access journal articles used by respondents in each country and in each discipline.

Table 4.3
Most frequent way to access journal articles: electronic library (online journal)

	Australia				*Brazil*				*Total*			
	L		M		L		M		L		M	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
e-collection	9	82	9	90	6*	35	12	80	15	54	21	84
Total	*N*		*%*		*N*		*%*		*N*		*%*	
	18		86		18		56		36		68	

*In Linguistics-Brazil electronic collection was the preferred way alongside with library print collections, which was also indicated by 6 (35%) of respondents.

The results show that the majority of respondents (68%) access academic journals through electronic collections. This is especially true for linguists in Australia (82%) and Medicine researchers in Australia (90%) and Brazil (80%). Linguists in Brazil also use online collections (35%) but equally rely on library print collections (35%). As pointed out by one participant (L#12), there are 'young' private universities in Brazil which do not have access to

Portal Periódicos CAPES because they do not fulfill the requirements made by CAPES to join the consortia, a fact that can help explain why linguists in Brazil still rely strongly on print journals.

> LB#12 – *Universidades como a X[55] (que estão começando sua atividade de pesquisa) não têm acesso às Bases da Capes. Não tendo acesso a tais meios, não temos como desenvolver essas práticas e familiaridade com tais gêneros 'emergentes'. O contato que tenho com revistas digitais é bastante esporádico e acidental. O contato com boas bases me permitiria, claro, um trabalho mais sistemático nesse sentido.[56]*

Brazilian linguists were also the group that most intensely (5 - 29%) made use of the "Other" option in Question 1, indicating two alternative paths as frequent ways they access academic articles: 3 of them use free ejournals, *Google*, and *Google Scholar* and 2 indicated that the most frequent way they obtain articles is from their lecturers.

If we compare each country as a whole concerning the total figures of electronic collection usage, it is possible to notice a significant difference: 86% of researchers in Australia use ejournals, considerably more than the 56% rate of ejournals use in Brazil, which could be a consequence of the above cited lack of access to this service in some Brazilian universities, but also of the latter's lower familiarity with ejournals due to the fact that they have been available for a shorter time in Brazil compared to Australia. Nevertheless, as highlighted above, it is particularly Linguistics in Brazil that does not use ejournals frequently. More importantly than indicating a cross-country contrast, the percentages in Table 4.3 reveal a disciplinary divergence in Brazil: only 35% of

---

[55] The name of the university was substituted by a code for ethical reasons.
[56] The original answers provided by the participants were transcribed here without undergoing grammar and spell checking.

linguists use ejournals, contrasting with 80% of respondents in Medicine who do so, which clearly illustrates how important ejournals are for Medicine, whose researchers have no way out but adapting to the new technological requirements if they want to follow the pace of the discipline.

In relation to the least frequent way linguists and medical researchers access journals articles, linguists indicated personal subscriptions (19 or 68%), which is also the alternative indicated by the majority of Brazilian medical researchers (7 or 47%). Australian medical researchers, however, rarely obtain articles from peers (5 or 50%). During the interview with [MI], he pointed out that most medical researchers (in Australia) have personal subscriptions of print journals through their membership in medical societies, i.e., membership fees include getting a hard copy (or print copy) of journal issues.

Finally, one last aspect that deserves to be commented is that, conversely to what could be expected, there is no interdependency between the favorite way of accessing journals and age. As shown in Table 4.2, the average participant from Linguistics is relatively younger than the average participant from Medicine and still fewer linguists use ejournals, indicating that disciplinary pressures play a much bigger role than age in ejournal usage practices.

4.3.2 Question 2 - Whenever you find a relevant article in an online journal, do you print it out to read it or do you read it on the screen? Why?

In Question 2 subjects were asked to comment on their reading habits when it comes to articles published in ejournals, specifically in relation to the contrast print technology versus screen technology. My goal with this question was to learn if ejournals are seen simply as article delivery devices, which

would mean that they do not affect traditional reading practices, or if users have not been taking advantage of the affordances of the electronic medium when reading HTML research article, such as the immediate access to cited references provided in HTML research articles. Based on the percentages displayed in Figure 4.3, it seems that whatever the advantages of the electronic medium might be, they do not beat the kind of interaction, comfort, and portability that paper allows, given that the majority of respondents (44 or 83%) prefers to read articles printouts instead of screen reading.
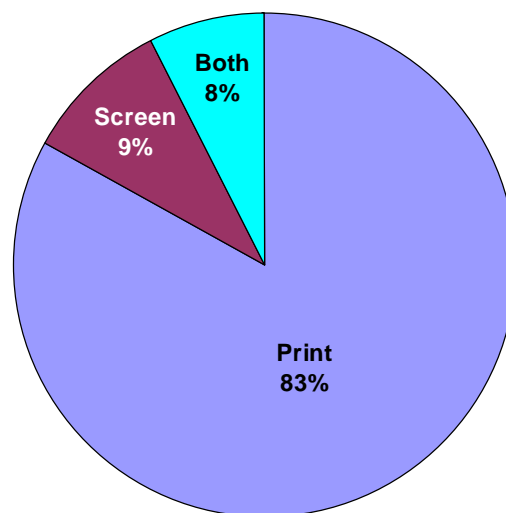
Figure 4.3. Preferred medium for reading journal articles: print vs. screen

This tendency toward the print medium predominated throughout the answers collected from both Brazilian and Australian researchers, being uniform across disciplines as well (Table 4.4). As far as the countries are concerned, out of the 21 participants from Australia, 18 (86%) pointed out that they prefer to read articles in print and out of 32 Brazilian participants, 26 (81%) did so, which means no significant difference. Similarly, no significant variation was observed in relation to the reading preference in the disciplinary contexts, because 23

(82%) of the linguists and 21 (84%) medical researchers who responded prefer to read articles in print.

Figure 4.3 also shows that a small number (4 or 8%) of respondents use both formats indistinctively, but just 5 (9%) prefer screen reading. Consonantly to the results obtained in Question 1, age doesn't seem to affect reading preferences, inasmuch as 4 (LB#10, LB#16, MA#7, MA#10) out of the 5 participants in the sample who refer screen reading belong to the older group: LB#10 is between 45-49, MA#7 is between 50-54 and both LB#16 and MA#10 are 55 or older.

Table 4.4
Cross-country and cross-disciplinary reading preferences regarding the print or screen medium

	Australia		*Brazil*		*Total*	
	L	M	L	M	L	M
Print	10	8	13	13	23	21
Screen	1	2	2	–	3	2
Both	–	–	2	2	2	2

The reasons why respondents prefer to read print articles instead of reading them on the screen are summarized in order of frequency in Table 4.5. Clearly, there are three core aspects of print reading that influenced researchers' choice: 1) the handling possibilities of paper, which can be grasped and manipulated, allowing easier highlighting and comment insertion, 2) the harmless nature of paper when it comes to ophthalmic issues, and 3) the portability of paper. If we compare the incidence of each one of these aspects across the disciplines and the countries, it can be observed that the numbers in

Table 4.5 are quite evenly distributed, although a slight variation can be noticed in relation to the disciplinary contexts: the great majority of linguists prefers a printout because it allows them to add notes and highlight important parts, while for most respondents in medicine the major advantage of paper is that is it more comfortable for the eyes.

Table 4.5
Reasons why respondents prefer to read articles in print format

Reasons why print	**Australia**		**Brazil**		*Total*
	L	M	L	M	
Easier to scribble on and highlight	4	3	7	2	16
Easier (more comfortable, less strenuous for the eyes) to read on paper, better to concentrate on paper	4	4	3	4	15
To read elsewhere	1	3	4	3	11
For use in lectures and courses, sharing with students and staff	−	1	3	1	4
To use when writing papers	1	2	−	1	4
Easier to archive/file	−	1	1	1	3
Easier to retrieve (not necessary to turn on computer)	1	−	2	−	3
Dislike reading on screen	1	−	1	1	3
Habit	−	−	−	2	2
Print is clearer	1	−	−	1	2
Like to keep a hard copy	1	1	−	−	2
Easier to refer to tables and figures	−	1	−	−	1

It is important to notice that the favoritism for the print format is straightly connected to a specific 'type' of reading: careful and time consuming. As indicated in the body of the question, subjects were asked to identify which format they prefer when reading *relevant* articles only, suggesting that the kind of reading being referred to involves attention, concentration and time. In fact, respondents clearly signaled that for quicker activities such as skimming they also read articles on the screen, as illustrated by LA#2, LA#8, LB#7, MB#3, MB#5, MB#7, and MA#2.

LA#2 – *I read it on screen first* and if there is a lot of important information relevant to my research I then print it out and highlight all the bits I want, and write lots of margin notes all over the document.

LA#8 – *skim online* and print if i intend to use it. i dislike extensive reading on screen, and i pile up hardcopies of books and journal articles around me when I write. i do sometimes use pdfs in writing, but generally if i can't find my hardcopy.

LB#7 – Normalmente, *faço uma primeira leitura na tela do computador* e salvo o arquivo em uma pasta. Dessa forma, fica mais fácil a própria socialização do texto. Para utilizar em cursos ou para não perder o texto, faço também uma cópia impressa e guardo em pastas.

MB#3 – Imprimo os mais importantes, que requerem leitura detalhada, pois ler no papel é muito mais confortável. *Leio na tela os que me interessam mais superficialmente*.

MB#5 – Geralmente *faço leitura rápida na tela do computador* e só imprimo se quiser usá-lo em alguma aula ou se o considero excelente e se quero lê-lo com calma em outros locais.

MA#2 – If it's worth reading carefully (i.e. *I'm not just checking on methodology*) then I will always print it out: I find it much easier to read on paper and to think about an article when I can scribble notes on the hard copy.

Of the 5 (9%) respondents who prefer to read online research articles exclusively on screen, only two justified their choice (MA#7 and LA#9):

MA#7 – *Read on screen becAuse of links*. Print out if I'm travelling or otherwise don't have computer access.

LA#9 – read onscreen. *I can make notes using acrobat professional. If I need to take it somewhere, I either transfer it to my phone or laptop*, or, rarely, print it out.

Just as if s/he was anticipating that his/her colleagues would argue that they prefer print reading because of note taking and portability facilities of

paper, the Australian linguist (LA#9) explained that these facilities are also available on the screen if the adequate software and equipment are available.

As was pointed out above, although the print format is the favorite medium for article reading, a number of respondents (LA#2, LA#8, LB#7, MB#3, MB#5, MB#7, MA#2) also mentioned that they read on the screen when it comes to article skimming. For two participants, however, one further reason motivates them to reading on the screen: lower cost. Interestingly, both are Brazilian linguists (LB#3 and LB#6) and according to them, it becomes cheaper to read articles on the screen than to print them out:

> LB#3 – *Depende da importância do artigo. Na maioria das vezes, os que julgo importantes são impressos, já os que julgo para mero conhecimento/curiosidade, os leio online. Também levo em consideração o número de páginas a serem impressas. Se forem muitas, seleciono as mais importantes. Se o artigo tiver gráficos animados, etc, prefiro lê-los online.*

> LB#6 – *Geralmente eu leio na tela, imprimo somente quando acho discussões relevantes para minha pesquisa. Eu já acostumei ler textos na web, acho que pela agilidade e tb pq custa mto caro imprimir tudo em papel. A internet facilidade o acesso e tb diminuir o custo com gastos desnecessários.*

One explanation for this concern with financial factors could be that both LB#3 and LB#6 are young researchers and, therefore, may not have a job or a stable income yet.

As for those who use the screen for superficial reading, it did not become clear how this process is developed, which article format is used, PDF or HTML, and what the role of hyperlinks and other medium-related features is in this

process. The analysis of the answers to Questions 3 and 4 below should offer more information on this matter.

4.3.3 Question 3 - Some articles in online journals are available in HTML as well as in PDF format. Which format do you prefer? Why?

Question 3 is one of the most important questions in the subject-based data collection because it is directly oriented towards the main objective of the present research: to determine whether HTML research articles can be considered a new genre or not. Answers to this question should indicate the purpose(s) with which ejournals users read HTML research articles and, therefore, reveal whether the communicative purpose of this format is different from the communicative purpose of print RAs or not. If HTML research articles are used for new purposes, the initial step towards classifying them as a new genre can be taken. As pointed out earlier in Chapter 3 – Methods, Askehave and Swales (2001, p. 195) indicate that in genre analysis "an important and often primary criterion for deciding whether a particular discourse falls within a particular generic category" is *communicative purpose*. Notwithstanding, the ways by which the communicative purpose of a socio-discursive event can be identified are neither always simple nor quick, if we consider that "communicative purpose can be sufficiently elusive to be largely unavailable for the initial or early identification and categorization of discourses as belonging to certain genres" (Ibid., p. 204). In this sense, answers to this question were important but not sufficient for classifying HTML research articles as a new genre, mainly due to the fact that they account for only one of the participants in this social practice, the consumers, leaving HTML research article producers'

purposes uncovered. For this reason, two major linguistic and medical HTML research article producers, *ProQuest* and *Science Direct*, were also interviewed and their answers will be discussed in this section.

First, ejournal readers' answers will be discussed. Figure 4.4 shows which article format they prefer, HTML or PDF. The majority (35 or 67%) of respondents prefers the PDF format, two of whom (LB#5 and LA#2) actually never used the HTML format. This result is coherent with the preference for reading articles in print reported by respondents in their answers to Question 2, inasmuch as the PDF version of articles is considered the *printer-friendly* format (Tenopir, 2003, p. 45-46).

Figure 4.4. Favorite ejournal article format: HTML vs. PDF
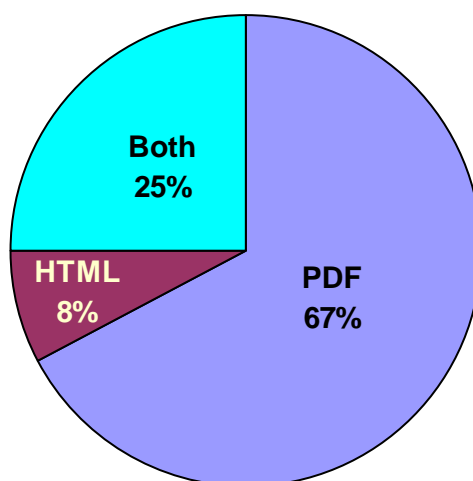
In fact, one of the two major reasons reported by researchers in the sample for their preference for PDFs over HTMLs is printing (Table 4.6). The other reason is the layout of PDF research articles, which respondents like better than the design of HTML articles. The resemblance between the PDF format and the print journal format of RAs could help explain why participants

are more attracted towards the PDFs, because that is the design they are more accustomed to after centuries of printed page predominance in scientific communication. As pointed out by LA#5 and LB#9, familiarity can be an influential factor in online reading practices.

> LA#5 – *I prefer the PDF format because it is in more of a Print format, and I am very familiar with the Adobe Acrobat program which opens it (and this familiarity attracts my attention).*

> LB# 9 – *Prefiro PDF. Porque já estou mais familiarizada com este formato. Nada me impede de usar o outro formato, se necessário.*

In spite of that, scientific communication style sheets still demand that page numbers must be mentioned whenever direct quotations are made, page numbers being only available in the print journal article or in PDF articles. This citation aspect was also indicated by respondents for their preference of PDF over HTML articles (LA#8, LB#16, MA#2).

> LA#8 – *pdf - i can quote the page numbers and find them easier to save / print / etc.*

> LB#16 – *PDF , pois é mais fácil para citar as páginas.*

> MA#2 – *PDF every time: because of the need for pagination for citation, and because I like a printout that looks like a real journal article, not a web page.*

The paper-like pagination system, overall configuration, and stability of the PDF format of research article, therefore, make it look much more like the traditional well established print journal research article than the HTML version.

In this sense, the PDF version seems to be more valuable that the HTML version because it looks more formal and *real*, as suggested by MB#13 and MA#2:

> MB#13 – *Prefiro o formato PDF, acho que pelo fato deste formato manter o estilo mais formal e assim imprimi-lo (motivos da questÃ£o anterior)*

> MA#2 – *PDF every time: because of the need for pagination for citation, and because I like a printout that looks like a real journal article, not a web page.* (my emphasis)

Further reasons presented by the participants on why they prefer the PDF format are summarized in Table 4.6, as well as the frequency rates by which each reason was indicated across disciplines and countries.

Table 4.6
Reasons why participants prefer the PDF format of RAs

Reasons - PDF	Australia		Brazil		*Total*
	L	M	L	M	
Better/clearer design/layout/format	6	6	3	6	21
Printing	4	4	7	4	19
Easier to archive	4	2	1	1	8
Habit	–	–	3	1	4
Page numbers/citation	1	1	1	–	3
Doesn't require internet access	–	–	1	–	1
To "text search"	1	–	–	–	1
Familiairty with *Adobe Acrobat*	1	–	–	–	1
Take notes on pages	1	–	–	–	1
"mais econômico"	–	–	–	1	1

The 13 (25%) respondents who use both PDF and HTML formats did not praise PDF's design, but equally see it as the best format for printing. As a printout, the PDF version of RAs fulfills the same communicative purpose of its

print journal counterpart. What about the communicative purpose of the HTML format? Why would ejournals present two formats with the same function?

For the purpose of the present research, although the majority of respondents prefers the PDF format, it is the HTML's use and function that is being investigated. To this end, the answers of the 4 participants which prefer the HTML format as well as of the 13 participants who use both formats will also be carefully examined.

First of all, it was found that only 1 (MA#7) of the 4 respondents who prefer the HTML version of online articles justified her/his answer.

MA#7 – *HTML because of links (when they're there)*

Similarly, only 7 of the 13 respondents who use both formats indicated a reason for using the HTML version or RAs. The remaining 6 just mentioned that they use the HTML as well as the PDF. Consequently, as far as HTML consumers are concerned, interpretations about the purpose of this format will be drawn based on the answers of 8 subjects, 5 from Medicine (MB#6, MB#14, MA#4, MA#7, MA#9) and 3 from Linguistics (LB#3, LB#6, LB#10). It seems important to underline this fact because it could be one of the limitations of the present research which in turn could jeopardize any generalizations about HTML research articles suggested in this study.

Table 4.7 summarizes the reasons why these 8 subjects use the HTML format.

Table 4.7
Reasons why participants use the HTML format of RAs

Reasons – HTML	Australia		Brazil		*Total*
	L	M	L	M	
Immediate access to previous literature	–	1	2	1	4
Presence of hyperlinks	–	1	1	1	3
Access to quality graphics	–	2	–	–	2
Rapid scaning	–	1	–	–	1
Immediate access to supplementary information	–	1	–	–	1

Table 4.7 shows that the main advantages associated with the HTML article format are related to *speed* and *access* affordances. The *speed* factor is evidenced in the answers through the explicit lexemes "rapidamente" (LB#10), "rapid" and "immediate" (MA#9).

> LB#10 – *formato html pela possibilidade de ir para outras referências <u>rapidamente</u>. no formato pdf é melhor para impressão.*

> MA#9 – *HTML for <u>rapid</u> scanning of papers for specific technical information - <u>immediate</u> links to supplementary info & literature and ability to resize figures are useful. PDF files used for printing and more careful reading archiving*

The *access* element refers to the different types of information that can be reached through the HTML version: previous literature, higher quality graphics and supplementary information. Regarding the first aspect, access to previous literature refers to the fact that the entire texts of most of the studies cited in a given HTML research article can be reached through the links available in the reference list (Figure 4.5)
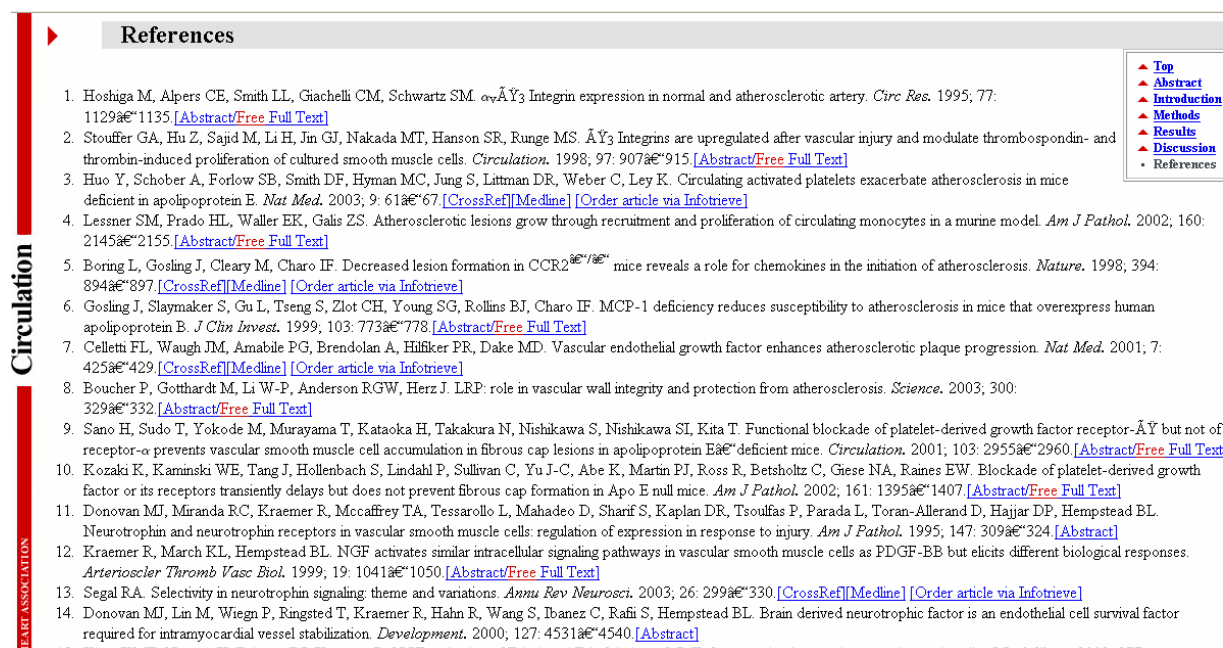
Figure 4.5. Immediate access to previous literature in the HTML format of research articles

This characteristic of HTML research articles was pointed out as the main reason why participants use this format (Table 4.7), as illustrated by LB#6, LB#10, MB#14.

> LB#6 – *Para imprimir eu prefiro o PDF, mas <u>qdo estou lendo o texto e tem hiperlink eu prefiro em HTML, pq assim posso ir de um texto para outro ou pelo menos verificar que relações são estabelecidas entre os artigos citados</u>.*

> LB#10 – <u>*formato html pela possibilidade de ir para outras referências rapidamente*</u>. *no formato pdf é melhor para impressão.*

> MB#14 – *Inicialmente <u>eu vejo no formato HTML (consultar referencias e seus links</u>) e em seguida no formato PDF para impressão.*

> MA#9 – *HTML for rapid scanning of papers for specific technical information - <u>immediate links to</u> supplementary info & <u>literature</u> and ability to resize figures are useful. PDF files used for printing and more careful reading archiving*

Another reason related to access refers to high quality images offered in HTML articles, which are not available in the print journal version of articles neither in the PDF version.

> MA#4 – *It depends on the purpose. I use HTML if I need a graphic/figure/table for a particular purpose (e.g., to incorporate into a PowerPoint presentation).*
>
> MA#9 – *HTML for rapid scanning of papers for specific technical information - immediate links to supplementary info & literature and ability to resize figures are useful. PDF files used for printing and more careful reading archiving*

Similarly, for [MI], the interviewee from the area of Medicine, images are also the main cause why he choose the HTML version of articles:

> [MI] – 90% of the papers I read are in PDF. I would use the HTML if I wanna reproduce something, if I wanna grab images and do something with them because there is more flexibility.

Finally, the *access* advantage of the HTML format is also due to the fact that supplementary information can be readily retrieved. This component was also indicated by MA#9.

> MA#9 – *HTML for rapid scanning of papers for specific technical information - immediate links to supplementary info & literature and ability to resize figures are useful. PDF files used for printing and more careful reading archiving*

Nevertheless, what MA#9 does not seem to know is that supplementary information is no longer uniquely accessible *inside* the HTML article. Recently,

online publishers have started to place a link to supplements in the *Table of contents* of ejournals (Figure 4.6) and, hence, this may not be the driving force why researchers would use HTML research articles in the future.
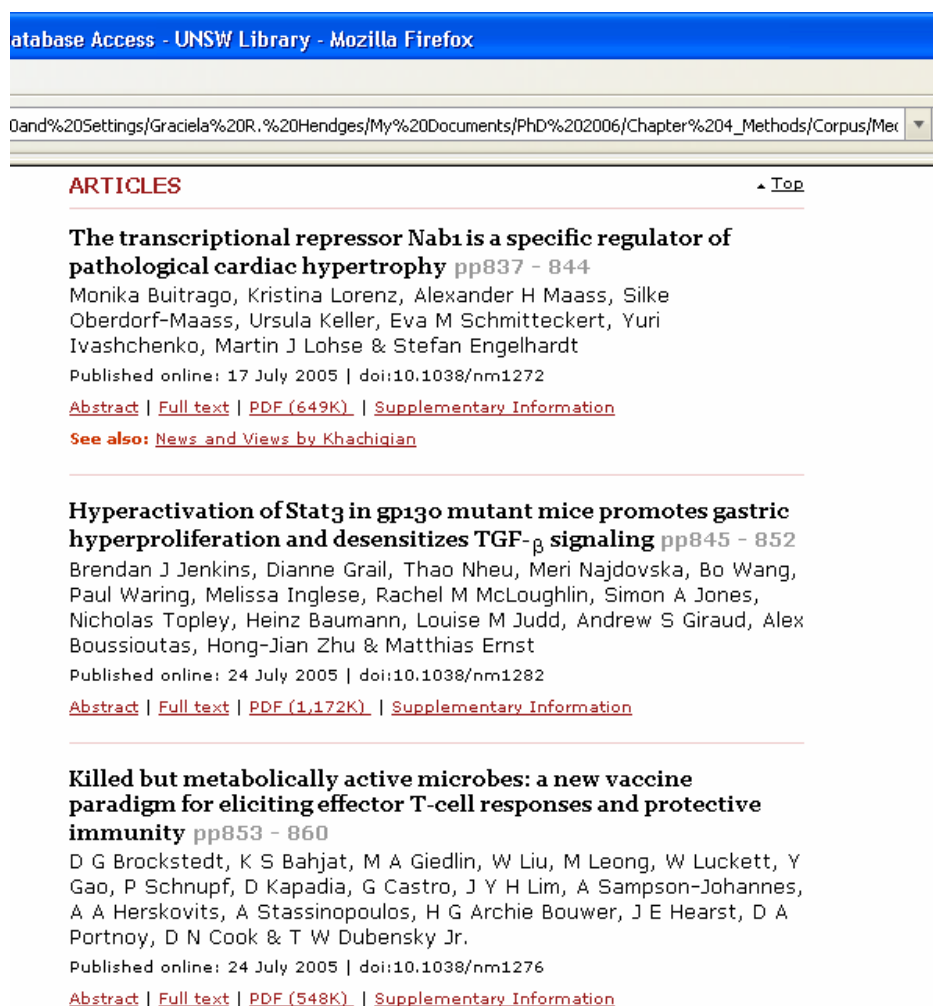

Figure 4.6. Links to *Supplementary Information* available in the *Table of contents* of ejournals (example from *Nature Medicine*)

For Tenopir (2003, p. 45-46), while the role of the PDF format of documents is to be *printer-friendly*, the role of the HTML format is to be *viewer-friendly*. Although it is not clear what the author means by viewer-friendliness, based on some responses to the questionnaire one could claim that for medical

researchers the HTML version can be considerably unfriendly to the viewer rather than friendly. This unfriendliness refers particularly to the fact that in HTML research articles, figures, tables, graphs and the like are generally not readily visible at the *homepage* of the articles. The homepage of an HTML research article is the first page that is accessed via the table of contents of the ejournal. In this homepage, non-verbal elements are *hidden* behind hyperlinks (as with the table in Figure 4.7) and/or hyperlinked thumbnails (in the case of the figures depicted in Figure 4.7)[57], if not totally omitted. These hyperlinks must then be activated by the HTML research article reader if s/he wants to see the images properly.

Medical researchers (MA#8 and MA#10) have reported to be dissatisfied with this configuration:

> MA# 8 – *PDF because I print it and it is all readily visible where in html I need to remeber to separately print all the figures/tables. Even <u>if I read on line I find it annoying to have to click hyperlinks to be able to see all the parts I want</u>.*
>
> MA# 10 – <u>*PDF - more compact,* *figs easier to read*</u>.

This issue of image accessibility probably does not affect all disciplines with the same intensity. It has been well documented how crucial tables, figures and other visuals are in Medicine and science in general (Bazerman, 1988; Lemke, 1998; Gross, Harmon & Reidy, 2000, 2002) and that, when reading RAs, many scientists "may well read the tables or graphs first, and then their

---

[57] This configuration will also be discussed in Chapter 5.

captions, and only then the main text" (Lemke, 1998). This is specially the case

in the field of Medicine, as acknowledged by Wang (2006):

> radiologists read the literature in a way that is probably a bit different to most other disciplines: basically there is a huge mass of literature, as you know, there is more than we can possibly read in a year, so what most people do is they skim the title, then look at the pictures, *maybe* read the abstract, and then occasionally read the paper, so essentially if the pictures don't grab your attention, there is a low likelihood that the radiologists will read the paper.

Figure 4.7. Images placed behind hyperlinks in HTML academic articles

In this sense, skimming a medical HTML research article, which should

be a quick activity, means skimming the images in this article. In this sense, to

be user-friendly, it seems that the HTML format of medical articles should

present all images in the homepage of the HTML research article, right after it is

accessed in the table of contents of medical ejournals. This would enhance rather than decelerate the skimming process and maybe the HTML format would become more popular among researchers in general[58].

In spite of their dissatisfaction with the configuration of HTML research articles, medical researchers in the sample still use this format more frequently than linguists (Figure 4.8).
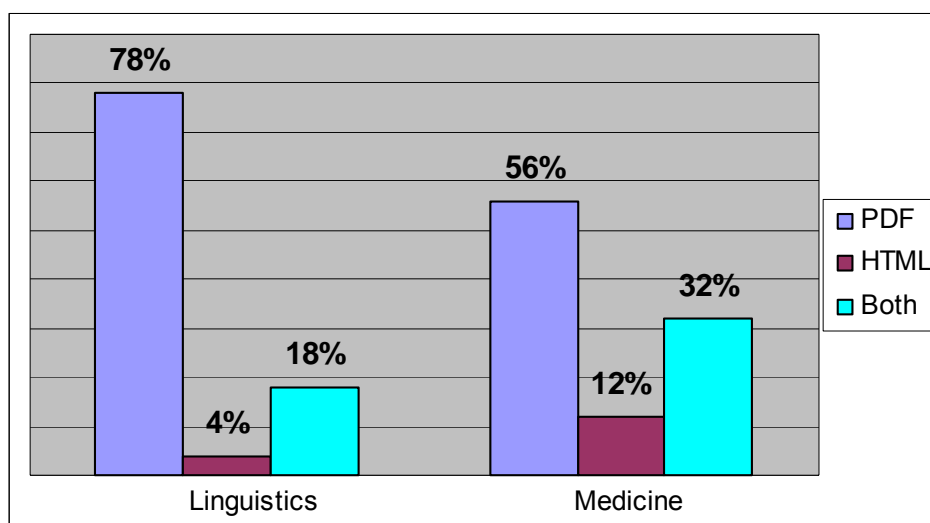
Figure 4.8. Ejournal article format preferences across disciplines

As reported in Chapter 3 – Methods, the presence of the HTML article format in Linguistics ejournals is less common than in medical ejournals, where both the PDF and the HTML format for each RA is always available in the ejournals that constitute the corpus. In this sense, Question 3 - "Some articles in online journals are available in HTML as well as in PDF format. Which format do you prefer? Why?" also aims at verifying if this difference could be attributed to

---

[58] Not unlike the outcome obtained here, informal surveys with the audiences of several conferences where I have presented preliminary results of the present research revealed that most of them never use the HTML format.

usage practices, i.e., if the HTMLs are more often available in medical ejournals because they are used more frequently by medical researchers than linguists.

Based on the percentages shown in Figure 4.8, it could be speculated that there is some relationship between the larger presence of HTML research articles in medical journals and the higher frequency with which medical researchers make use of this format when compared to linguists. It was mentioned above that 2 linguists even indicated never having used the HTML format.

If HTML research article are rarely accessed or even unknown by ejournal consumers, why are they available? In search for an answer to this question, I interviewed ejournal producers.

Conversely to the traditional three-component model of linguistic interactions characterized by [author/producer + text + reader/consumer], in the case of HTML research articles a forth component is added to the picture: [author/producer + *information architect/aggregator* + text + reader/consumer]. In previous studies about the print RA, the roles of editors, designers, publishers and other similar actors have rarely been fully acknowledged[59], maybe because their roles have limited impact on the genre and therefore were not worth considering, or maybe because genre analysis approaches and methodologies in general paid less attention to context and the tools for analyzing the roles of these actors were not straightforwardly available. The actors in print RA analysis that have traditionally been considered are *author*, *text* and *reader* (Figure 4.9).

---

[59] Obviously, there are exceptions such as Motta-Roth's (1995) study about academic book reviews in which she considers the role of journal editors in the constitution of the genre.
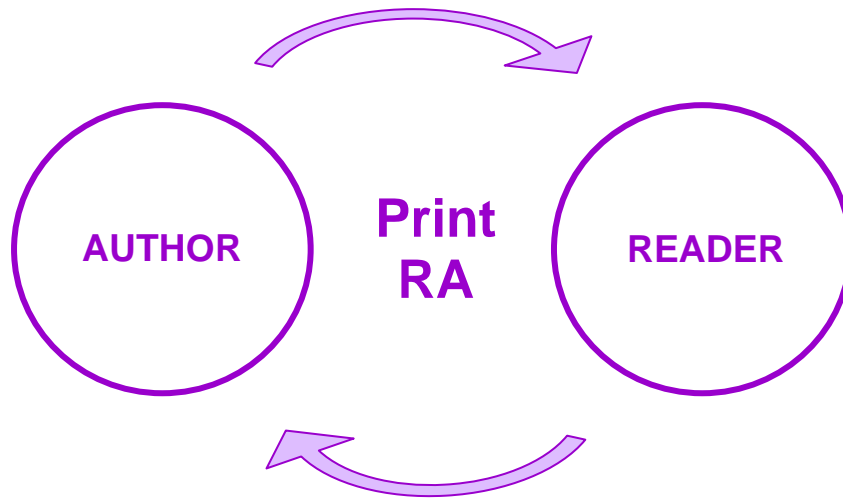
Figure 4.9. Three-component communication model in previous studies of print RAs

Technological advancements such as the Internet, however, seem to have increased the power of these actors in generic activity, pressing genre analysts to consider their roles.

In the case of HTML research articles, the *creators* of these texts are information architects working in online publishing companies or aggregator databases. It is they – and not the scientist responsible for the *propositional content* in the HTML research article – who take the decisions about the layout, design, organization of HTML research articles, and, even more importantly, about the presence, nature and placement of hyperlinks in HTML research articles as well as about the nature, presence, placement, and quality of images in these articles. I suggest, therefore, that a model for analyzing HTML research articles should consider the role of aggregators in shaping these texts (Figure 4.10).
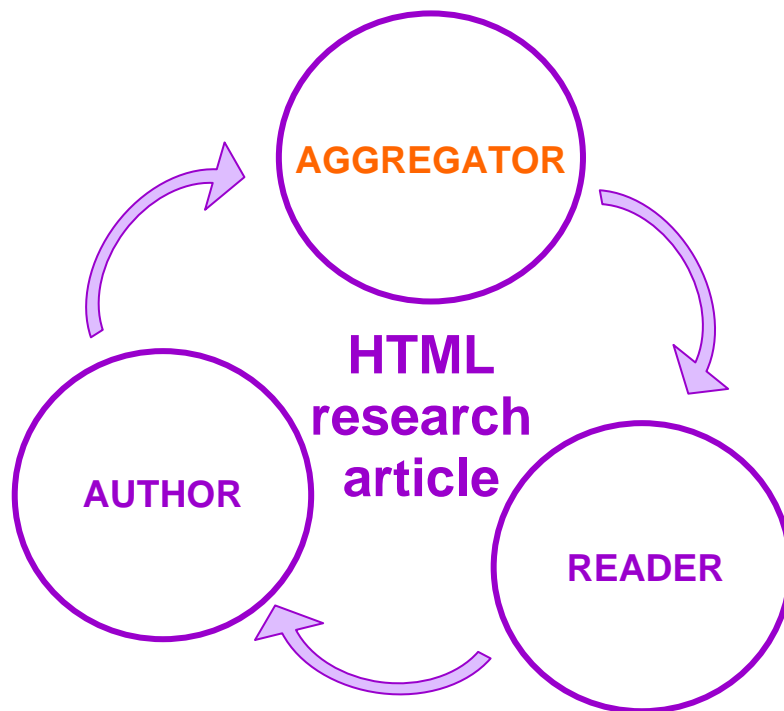
Figure 4.10. Four-component communication model for the study of HTML research articles

In the present research, in order to understand the purpose of HTML research articles from the perspective of producers, two aggregators were interviewed: *ProQuest* and *Science Direct*. These are the two most productive databases in my sample of 13 aggregators, each one being responsible for the online publications of 5 out of the 14 ejournals examined here. *ProQuest* alongside with *Ovid* is the leading aggregator in the publication of the medical ejournals in the corpus and *Science Direct* publishes most of the Linguistics ejournals in the corpus (4 out of 7) (refer to Table 5.1 in Chapter 5 for a detailed visualization of the distribution of the ejournals per aggregator).

Based on ProQuest's[60] answer, it seems that whatever the role of HTML research articles is, it is secondary in importance to the role of PDF research articles, given that HTML articles work as *substitutes* for PDF articles, when technological resources for viewing PDF files are not available.

> ProQuest: *I am not 100 % sure on this so most of this is supposition on my part. I believe as PDF has become more popular over the years this format has become the standard and <u>HTML has declined to me more of a legacy format</u>. The reason therefore I suspect you will have both for the Medical journals is because that we have published them for longer and initially they probably only provided them in a non PDF format so we had to produce them in HTML. <u>We have then subsequently maintained the HTML format because we always used to provide it and because I would have thought there are a few people who still may not have Acrobat installed and hence cannot read the PDF files</u>.*

The lower status of HTML articles is also indicated by the choice of the verb "declined". A similar concern with the availability of software technology was expressed by the second aggregator interviewed here, *Science Direct*.

> Science Direct: *In response to your question, <u>I believe that the design in mind was to have the HTML version available to users who either do not have Adobe Acrobat installed on their computer or for users who have problems opening them up, such as slower systems</u>.*

Once more, the role of the HTML format is relegated to that of *substitute*. Nevertheless, for those ejournal users who have no obstacles in accessing PDFs, *Science Direct* pointed out that another purpose of the HTML version is to offer quick access to the content of an article.

---

[60] The answer was provided by one of the aggregator's IT (Information Technology) support team members.

Science Direct: *The HTML version is used quite often for users who want to see a copy of the article relatively faster, as opposed to accessing a separate PDF file version.*

This content, however, is essentially verbiage of RAs, i.e., the written or verbal part of the texts, because, as pointed out previously in this chapter, when an HTML article is accessed from the table of contents of ejournals the only content that can be readily viewed is the verbiage and not the images. Because the download process of images takes longer than the downloading of verbiage, aggregators eliminate images from the homepage of HTML research articles, placing them behind hyperlinks, as mentioned earlier. For some disciplines, however, seeing "a copy of the article relatively faster" would be worth if images could be skimmed right away. In these contexts, if images are not presented at the homepage of HTML research articles, the 'fast seeing' process is delayed and this purpose of the HTML format is undermined.

In spite of the fact that the perspectives of producers and readers about the functions of HTML research articles do not fully coincide, a summary of the set of purposes as manifested by both is discriminated in Table 4.8.

Table 4.8
Purposes of the HTML research article according to producers and readers

Purposes of HTML articles	Producers	Consumers
Allow rapid viewing of article content	+	+
Offer immediate access to previous literature	−	+
Offer access to quality graphics	−	+
Offer access to supplementary information	−	+
Be a substitute for the PDF format	+	−

These purposes are all very technology oriented in the sense that they are medium-dependent. Hence, the responsibility of different media in determining the purpose of genres must not be overlooked.

Three of the purposes listed in Table 4.8 are realized in HTML research articles through hyperlinks, namely, 1) immediate access to previous literature; 2) access to quality graphics and 3) access to supplementary information. If ejournal readers see these as purposes of HTML research articles, question 4 seeks to explore whether they know that this purposes are realized by hyperlinks and if they do, how and with what frequency they use these hyperlinks.

4.3.4 Question 4 - On hyperlinks

In "Question 4 – On hyperlinks", subjects were asked about their 1) expectations, 2) usage practices and 3) feelings in relation to the use of hyperlinks in HTML research articles. The first question was "When you read online academic articles, do you expect to find hyperlinks?"; the second question was "Do you always 'click' on the hyperlinks?", and the third question was "If at the first 'page' of an online article you are faced with a hyperlinked table of contents, do you feel freer to choose where to start reading the article?". The responses per discipline and per country are shown in Table 4.9.

Table 4.9
Hyperlink usage as reported by participants

	Linguistics Brazil			Linguistics Australia			Medicine Brazil			Medicine Australia			Total		
	yes	no	sometimes	yes	no	sometimes	yes	no	sometimes	yes	no	sometimes	yes	no	sometimes
a. expect to find hyperlinks	12	3	2	3	5	3	10	3	2	7	2	1	32	13	8
b. click on hyperlinks	5	3	9	0	2	9	6	3	6	3	2	5	14	10	29
c. feel freer because of hyperlinks	11	2	4	5	5	1	11	3	1	5	5	0	32	15	6

A contrast between the disciplines revealed that hyperlinks are more expected, used and have a higher effect on medical HTML article readers' than on readers in the area of Linguistics (Figure 4.11). Figure 4.11 shows the percentage of participants from each field who responded *yes* in each one of three sub-questions in Question 4.
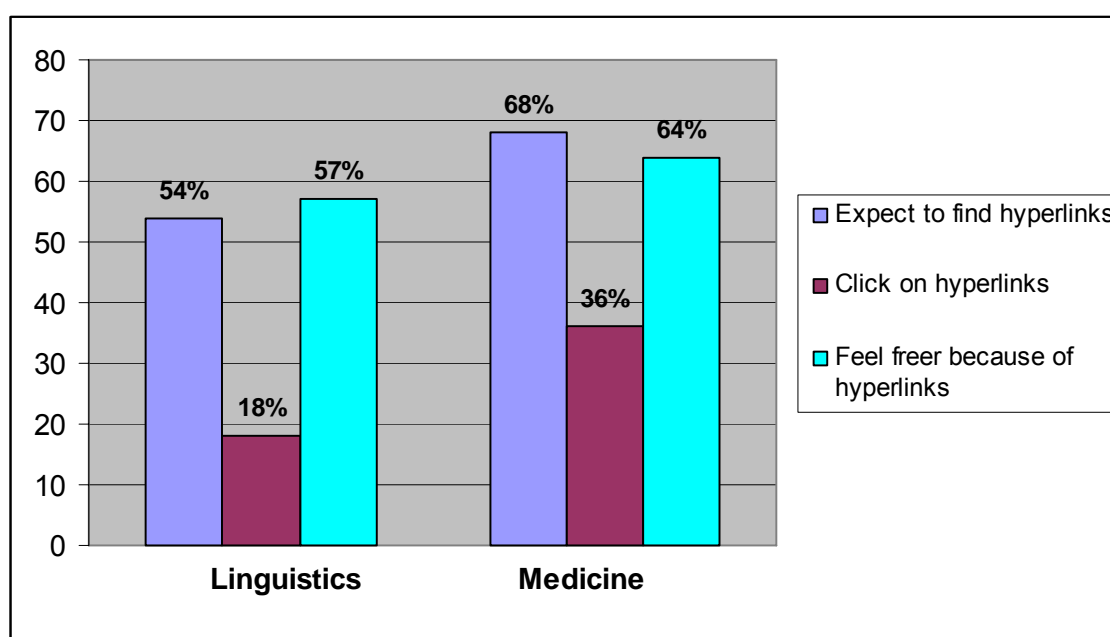

Figure 4.11. Percentages of respondents in each discipline who expect to find hyperlinks, use hyperlinks and have their reading of HTML research articles affected by hyperlinks

The disciplinary difference seems more significant particularly in relation to hyperlink usage. As can be observed in Figure 4.11, twice as many medical researchers (36%) *click* on HTML research article hyperlinks than linguists (18%). This result seems coherent with the fact that more medical researchers use HTML research articles than linguists (Figure 4.8) and, therefore, medical researchers may be more literate in the genre. Another reason for the higher *clicking activity* among medical researchers could be the fact that they are

obliged to click on hyperlinks if they want to read the images in HTML articles, because, as pointed out above, these are only readable when accessed via hyperlinks.

Since the majority of respondents expect to find hyperlinks in HTML research articles, in Question 5 the reasons and types of hyperlinks expected will be addressed.

4.3.5 Question 5. If you expect to find hyperlinks in online research articles, <u>why</u> do you expect to find them and <u>what kind of hyperlinks</u> do you expect to find (e.g. in references to previous research, to the research's stored data, to figures and tables, etc.)?

The results obtained from Question 5, regarding the reasons why ejournal users expect to find hyperlinks and the types of hyperlinks they expect to find are similar to those shown in Table 4.7 above: hyperlinks offer immediate access to cited literature and, therefore, it is this type of links they expect to find (Table 4.10). The term *expect* was used in the questionnaire elaboration process to mean *anticipate*. However, some of the answers seem to indicate that it might also have been interpreted as *hope*, as in "what kind of hyperlinks do you *hope* to find". Nevertheless, given that there is no safe way to be certain about which interpretation was given, the answers to question 5 will be analyzed conjointly.

Table 4.10
Reasons why respondents expect to find hyperlinks in ejournals

Reasons	LA	MA	LB	MB	Total
offer access to more information	1	−	4	2	7
are fashionable	−	1	1	−	2
are medium-inherent	−	1	4	−	5
speed up access	−	−	1	1	2
enhance reading and searching processes	−	−	3	1	3

The second most frequent reason why participants anticipate that they will encounter hyperlinks is their medium-inherent nature (MA#8, LB#6, LB#13).

MA#8 I guess I expect to see them just because they are there. It would not worry me if they were not.

LB#6 Pq o contexto é digital, espera-se que os textos estejam adequados Para que se tornem mais atrativos ao leitor na tela.

LB#13 Espero encontrar hiperlinks em artigos acadêmicos online porque a interface tecnológica possibilita isso, portanto espero que essa vantagem seja aproveitada.

Directly related to these reasons are the types of hyperlinks participants in both disciplines expect to find, with those offering access to previous literature being the most popular among researchers in both disciplinary domains (Table 4.11). Of the 53 respondents, 37 indicated some kind of hyperlink that they expect to find.

Table 4.11
Types of hyperlinks researchers in each field and country expect to find in HTML research articles

Types of hyperlinks	LA	MA	LB	MB	Total
References	5	5	12	9	31
Data supplements	3	6	3	4	16
Figures and tables	1	4	2	5	12
Contact details	–	1	1	–	2
Author's homepage	1	–	1	–	2
"Data that can not be presented in traditional printout format"	1	–	–	–	1

Hyperlinks to cited references were the first kind of hyperlink to appear in HTML research articles (Hendges, 2001). It is interesting to notice that some

researchers pointed out that they expect to find this type of links because they

help them to judge the quality of the article they are reading (MB#11, MB#13).

> MB# 11 Hiperlinks proporcionam acesso a maior número de informações. <u>Isso me permite uma análise mais criteriosa daquilo que é apresentado no texto originalmente acessado</u>. Habitualmente procuro por hiperlinks relacionados a referências e citações.

> MB#13 Os hiperlinks passam a idéia de outras fontes de pesquisa, de disponibilidade e de extensão/profundidade do artigo em questão.

This function of hyperlinks to cited literature has been highlighted in

previous studies (Hendges, 2001; Hendges & Motta-Roth, 2000; Motta-Roth et

al., 2000), where it was characterized as intensifying the interaction process

between reader and cited author. In this perspective, while in print this

interaction is mediated by the writer/author of the print article that is being read,

in online articles the reader has direct access to the original argument. Hence,

the mediation power of the writer/author would be reduced and the importance

of his/her role, reduced.

Another type of hyperlink which researchers reported that they expected

to find are those which offer access to supplementary data. This kind of link,

however, was indicated more often by medical researchers (59%) than linguists

(30%) (Table 4.12). One possible explanation for this contrast is that the

presentation of data supplements is not yet a common practice of Linguistics

ejournals, as acknowledged by LB#6:

> LB#6 Pq o contexto é digital, espera-se que os textos estejam adequados Para que se tornem mais atrativos ao leitor na tela.

Realmente um hiperlink pode levar a informações que um texto normal não levaria, por exemplo, aos dados da pesquisa, no formato de videos, etc. No entanto, ainda são poucos os textos que fazem uso desses recursos. Acho que ainda os hiperlinks são mais usados para relacionar textos publicados online, nas referências, citações. <u>É dificil encontrar um texto que apresente o corpus da pesquisa e tenha link voltado para ele</u>. Acho que isso ainda é um avanço que vai acontecer.

Considerable disciplinary variation was also observed in regards to the hyperlinks to images (tables and figures). Inasmuch as the role of images in medical RAs is central and considering that medical researchers are more familiarized with the HTML research articles than linguists and, thus, with the hyperlinking practices of ejournals, it is not surprising that they anticipate to find hyperlinks to tables and figures more frequently (53%) than linguists (15%) (Table 4.12).

Table 4.12
Types of hyperlinks researchers expect to find in HTML research articles discriminated per discipline

	L		M		Total	
	N	**%**	**N**	**%**	**N**	**%**
References	17	85	14	82	31	84
Data supplement	6	30	10	59	16	43
Figures and tables	3	15	9	53	12	32
Contact details	1	5	1	6	2	5
Personal homepage	2	10	–	–	2	5
"Data that can not be presented in traditional printout format"	1	5	–	–	1	3

In light of these findings, it seems that a temporary communicative purpose or *set of* communicate purposes for the HTML research articles from the point of view of its users – producers and consumers – can be outlined.

This purpose is presented in the next section, preceded by a summary of the

main results obtained in the contextual data analysis.

4.4 Summary of contextual data analysis

The main results obtained in this chapter are:

- there are significant differences between the Internet penetration rates between Brazil and Australia, with Australia being in a advantageous position;

- there are considerable differences between the nature of ejournal access offered in Brazil and in Australia in terms of the total number of journals researchers in each geographical context have access to and in terms of the numbers of aggregators researchers in each geographical context have access to;

- as a consequence of this technological scenario, fewer Brazilian researchers have access to and use ejournal collections than in Australian researchers;

- there is no direct correlation between age and ejournal usage rate when it comes to the use of ejournals;

- the preference for reading print-outs of RAs instead of screen reading prevailed in both disciplines and geographical contexts;

- medical researchers use HTML research articles more often than linguists, independently of the geographical context;

- HTML research articles are not viewer-friendly for medical researchers, at least as far as the presentation of images is concerned;

- The type of hyperlinks linguists as well as medical researchers expect to find more frequently in HTML research articles are hyperlinks to cited references;

- In the analysis of digital genres, whenever the roles of author and information architect do not conflate, i.e., whenever the author and the designer of the text are not the same person, the meanings created by both actors should be taken into account in a multimodal and multimedia perspective. Rather than the tripartite communication model traditionally invoked in the analysis of print

genres – author+text+reader – a four-component model should be considered: author+information architect+text+reader.

Nevertheless, these results do not address the main objective of the present research directly, namely, to verify *whether the HTML research article is a new genre or not.*

Considering that, following Askehave and Swales (2001) and Swales (2004), the privileged criterion for classifying a genre is *communicative purpose* and that to identify the communicative purpose both contextual and textual analysis should be carried out, the main result of this chapter is related to the purpose of HTML research articles as expressed by its users.

The contextual data analysis revealed that researches in the domains of Linguistics and Medicine do not use HTML research articles with the same purpose they use print or PDF RAs. They use HTML research articles with the purpose of taking advantage of resources which are exclusive of the electronic medium of the Internet, namely, search and browsing affordances and access to previous literature, quality images and data supplements.

As far as HTML research article producers are concerned, they also *write* HTML research articles having in mind technology related matters.

Based on these preliminary results, it could be argued that *the HTML research article is a new genre.* Its *primary communicative purpose* is not to report a research experiment; nor to advance the state of knowledge in a field, neither control quality of research, which are the purposes of print RAs or PDFs, but to serve as a technological aid by implementing features which are impractical in print. It may be used for those other purposes by a few or in those

cases when PDF files are not available, but these situations are not representative of what generally happens.

It may be the case that, in 2000, as viewed by Crowston and Williams, it could be argued that "the electronically distributed journal article is still in transition" […]. It is being used, but this adapted genre is not yet completely accepted or considered legitimate for all purposes (e.g., as evidence for a tenure case) by the academic community as a whole" (p. 203). The technological clock, however, tics much faster than the regular clock and after seven years the HTML research article has changed considerably if compared to the print RA.

These changes were to a certain extent expected by Crowston and Williams (2000, p. 203):

> the journal article will likely change as it moves onto the Web to take advantage of the possibilities of linking or embedding information; the eventual form may bear only passing resemblance to the self-contained 20- to 25-page articles of today. Also, modifications of genres that are parts of genre systems may require corresponding changes to the rest of the system. For example, changes in citation habits will be necessary before page numbers can be dropped from the technical paper genre. Such interdependencies between genres will tend to slow the adoption of a new genre.

To account for these "citation habits", academia has created the PDF version of electronic research articles. The accepted genre, therefore, i.e., the valid and "serious" format, that looks like a "real journal article", as indicated by some of the respondents in the questionnaire (LA#3, MA#2), is the PDF, which is the format that is used for the same reasons of the print RA: to report research experiments; to disseminate scientific information; to advance the

state of knowledge in a field; to control quality of research; to provide recognition for authors and career upgrade; and to help obtain funding.

The classification of the HTML research article as a new genre from the perspective of the context suggests that the medium has the power to impact and alter genre status, corroborating the argument presented by Crowston and Williams (1999, 2000) and Askehave and Nielsen (2005a, 2005b).

In the following chapter, the textual analysis will be presented and this set of purposes pointed out by HTML research article users will be put to the test, for it is textual analysis *combined* with contextual analysis that reveals the communicative purpose of a genre.

Chapter 5   Textual analysis

"The easy adoption of definitions can prevent us from seeing newly explored or newly emergent genres for what they really are." (Swales, 2004, p. 61)

5.1 Introduction

The Swalesian genre approach conceptualizes textual analysis as three-step procedure: 1. analysis of structure + 2. analysis of style + content and 3. analysis of purpose. All three constituents appear in Swales's 1990 definition of genre:

> A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These *purposes* are recognized by the expert members of the parent discourse community and thereby constitute the rationale for the genre. This rationale shapes the *schematic structure* of the discourse and influences and constrains choice of *content* and *style*. (p. 58). (My emphasis)

The interdependency among the constituents can be represented as in Figure 5.1, where *rhetorical strategies* mean *content* and *style*:

Communicative purpose
Realised by
Move structure
Realised by
Rhetorical strategies

Figure 5.1. Three-level genre model (reproduced from Askehave & Nielsen, 2005b, p. 2)

Because there is a large amount of studies that show how to analyze a print genre in terms of schematic structure (particularly within the so called ESP approach where the *moves and steps* framework is largely used; see Chapter 1 for a list of those), that seems to be less on an issue than analyzing style and content. Although aspects of style and content have been covered in genre-based publications, the procedures followed in these examinations are less systematic than studies about generic structure in the sense that it is not always clear what counts as style and what counts as content and how these features contribute for the classification of a text (the material entity of a genre) as genre X or Y. Within the Systemic Functional Linguistics (SFL), style and content would probably count as categories to be analyzed at the level of register, that is to say, one level below generic analysis. Thus, what counts in genre analysis in the systemic framework at the textual level is *form* and *purpose*.

Both in SFL and in English for Specific Purposes – Swales's approach –, the role of structure (form) in determining purpose is a crucial one inasmuch as genres are "highly structured and conventionalized in the sense that genres represent or lay down the way to go about accomplishing particular communicative purposes" (Askehave & Nielsen, 2005a, p. 122). The structure of a genre is composed by a number of functional units, the *moves*, which together realize its communicative purpose.

Despite the wide use of Swales's *moves and steps* model, Askehave and Nielsen (2005a, p. 123) argue that

The notion of moves is particularly useful for displaying generic conventions in terms of text organization. However, the 'utility value' of the concept is somewhat hampered by the fact that strong disagreement exists as to which criteria should be used for identifying move structure (see Paltridge, 1994). [Thus] Swales (1990) appears to base his criteria on two different systems namely lexis and grammar and rhetorical function (though most often the latter), Eggins (1994) relies solely on lexis and grammar, Martin (1992) focuses on the layout of a text, suggesting that 'titles, sub-titles, headings and subheadings are commonly deployed to keep track of the composition structure [of the texts]' (p. 443), while Bhatia (1993) concludes that 'the ultimate criteria for assigning discourse values to various moves is functional rather than formal' (p. 87).

The issue of move identification criteria seems to be further complicated when it comes to digital genres, for analytical categories of the existing genre analysis models (e.g. tense, voice, metadiscourse markers, hedges) seem to be insufficient to fully "capture the essence of genres mediated through the net" (Askehave & Nielsen, 2005a, p. 121). Specifically for genre classification, medium-related properties such as hyperlinks should not be overlooked because they may affect genre status (Shepherd & Watters, 1999; Crowston & Williams, 1999, 2000; Ihlström & Åkesson, 2004; Askehave & Nielsen, 2005a, 2005b). In the following sections I attempt to develop a genre analysis of HTML research articles which considers these properties.

5.2 The HTML research article: a textual genre analysis

At first sight, it seems that the HTML research article *looks* very much like the print RA: title, author and author affiliations, abstract, IMRD, references, endnotes, everything is there just like in the print RA.

But not so fast. Swales (2004, p. 76), quoting Dewitt, points out that "we know genres by what they are not as well as by what they are". Swales (Ibid.) agrees with Beebee (1994, quoted in Swales, 2004, p. 77) that it is a valuable perspective for genre theory to consider that "any generic reading of a text is based equally on what is not there, on what the text does not say, and ultimately on what cannot be done with it". From this viewpoint, when comparing print and HTML research articles it also becomes clear that there are several features of the HTML version that are not in the print version and several things that cannot be done with print RAs, all of which end up by making the HTML article quite a different text in terms of structural organization and layout (what the genre looks like) and functionality (what the genre can do and what can be done with it).

The main element that triggers the appearance, structural organization and functionality differences between print RAs and HTML research articles is the *hypertext link*. The hypertext link, or *hyperlink* or simply *link*, is the main feature of a hypertext (Landow, 1997; Crowston & Williams, 1999; Bar-Ilan & Fink, 2005) and is responsible for implementing the connection between two *nodes*: the "origin node", the node from which the connection starts, and the "destination node", the node where the connection ends (Agosti, Crestani, Melucci, 1997, p. 134). Another way of putting it is to say that, following Bar-Ilan and Fink (2005, p. 977), a hyperlink "connects two pages, the source and the target". *Node* and *page*, therefore, are used here as synonyms. Another term that has often been used to identify this kind of text chunk is *lexia* (Landow,

1997), which, according to Landow (Ibid.), was originally coined by Barthes in 1974 to identify units of reading in print texts.

For Crowston and Williams (1999), the form of a genre "is defined in part by the pattern of links it exhibits", and hence, links *could* affect the purpose of a genre. In this sense, it could be considered that in the case of a known print genre such as the RA, moving it to the electronic medium and adding links to it could change the purpose of this genre. But as pointed out in Chapter 2 (section 2.5), not all linking patterns affect purpose (Crowston & Williams, 1999, 2000; Askehave & Nielsen, 2005a). In the following sections, therefore, I will describe and discuss the kinds of hyperlinks found in the HTML research articles in the corpus. If structure realizes purpose and if the purposes of HTML articles are different from the purposes of print RAs, to what degree will hyperlinks contribute for the realization of the new purposes, namely: 1. allow rapid viewing of article content; 2. offer immediate access to previous literature; 3. offer access to quality graphics, 4. offer access to supplementary information, and 5. be a substitute for the PDF format?

5.2.1 Hyperlinking patterns in HTML research articles: the role of aggregator platforms

The analysis of hyperlinking patterns in HTML research articles cannot follow traditional methodologies employed in genre analysis of print RAs. In such methodologies, in general, a representative sample of RAs – from one or more than one discipline – is selected and then scrutinized in an attempt to find quantitative patterns, which are then qualitatively interpreted. While disciplinary constraints have been highly valued in such interpretations, journal editorial

policies are not always taken into account and yet these interpretations have proven to be legitimate.

When it comes to the analysis of HTML research articles, overlooking the role of publishers, aggregator databases, information architects and web designers, to say the least, in shaping this genre could jeopardize the validity of results and interpretations about the form and purpose of the HTML article[61]. In the print medium, each journal has a particular layout, which is different from other journals, whether they are or not published by the same publisher. Page layout is not only a result of editorial policies, but it can be influenced by disciplinary conventions (Nascimento, 2002). In the electronic medium, that is not quite as things work. In Chapter 3 – Methods I have tried to explain how an ejournal is published online by a web service called aggregator database, a service which can congregate journals from one or more publishers. For example, the medical journal *Annals of Internal Medicine* is published in print by the American College of Physicians. Online, one of its publishers is the aggregator database *ProQuest*, Similarly, the Linguistics *Journal of speech, language, and hearing research* is published in print by the *American Speech-Language-Hearing Association* (ASHA), but online it is also *ProQuest* that is responsible for its publication. As a result of being published online by the same aggregator database, these two ejournals look exactly alike, despite being from two different disciplines and from two different publishers (Figure 5.2).

---

[61] The reader may want to refer to Figure 4.10, page 117 in Chapter 4.

Figure 5.2. Two journals from different publishers and different disciplines conveyed online by the same aggregator (*ProQuest*): same design

This happens because ProQuest uses the same *template* for publishing

all its ejournals, whether they are from different publishers or not, or whether

they are from different disciplines or not. A template is a standardized electronic file with a predesigned, customized format that works as a master page ready to be filled in[62] and templates are used when rapid generation and mass production of web pages is necessary.

It is also the template that defines where and what kind of hyperlinks are going to be used in a web document. In the case of ejournals, then, this means that as long as two ejournals are published by the same aggregator database, the hyperlinking patterns in both will be the same, whether they are from different publishers or not, or whether they are from different disciplines or not. This is what happens with the above mentioned *Annals of Internal Medicine* and *Journal of Speech, Language, and Hearing Research* at *ProQuest*. Another example are the ejournals *The Lancet*, from Medicine, and *Journal of Memory and Language*, from Linguistics: both are published online by the aggregator platform *Science Direct* and therefore, though being two titles from very different disciplines, their HTML articles are equal in hyperlinking patterns and layout, including color patterns and typeface settings (Figure 5.3).

Therefore, for the analysis of hyperlinking patterns in the corpus, it will not be necessary to analyze each HTML research article in particular, but one exemplar of each one of the aggregator databases covered in the present research.

---

[62] According Dictionary.com, retrieved May 23, 2007, from http://dictionary.reference.com/browse/template, and Wikipedia, retrieved May 23, 2007, from http://en.wikipedia.org/wiki/Template.

ScienceDirect   UNSWLIBRARY   Login:
Regist

Home | Browse | Search | My Settings | Alerts | Help | Live Chat

Quick Search   Title, abstract, keywords [   ]   Author [   ]   e.g. j s smith
search tips   Journal/book title [   ]   Volume [ ] Issue [ ] Page [ ]   Clear Go

THE LANCET
Volume 366, Issue 9498, 12 November 2005-18 November 2005, Pages 1711-1717

SummaryPlus | Full Text + Links | PDF (664 K)   View thumbnail images | View full size images
Add to my Quick Links Cited By E-mail Article Save as Citation Alert Export Citation Citation Feed

S·F·X

Request Permission
View Record in Scopus
Cited By in Scopus (12)

Articles

Early changes in the skin microcirculation and muscle metabolism of the diabetic foot

Robert L Greenman PhD[a], Svetlana Panasyuk PhD[c], Xiaoen Wang MD[a], Thomas E Lyons DPM[b], Thanh Dinh DPM[b], Lydia Longoria BA[b], John M Giurini DPM[b], Jenny Freeman MD[c], Lalita Khaodhiar MD[b] and DrAristidis Veves MD[b],

[a]Department of Radiology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[b]Joslin-Beth Israel Deaconess Foot Center and Microcirculation Laboratory, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[c]HyperMed Inc, Watertown, MA, USA

Available online 10 November 2005.

Summary

ScienceDirect   UNSWLIBRARY   Login:
Regist

Home | Browse | Search | My Settings | Alerts | Help | Live Chat

Quick Search   Title, abstract, keywords [   ]   Author [   ]   e.g. j s smith
search tips   Journal/book title [   ]   Volume [ ] Issue [ ] Page [ ]   Clear Go

SummaryPlus | Full Text + Links | PDF (616 K)   View thumbnail images | View full size images
Add to my Quick Links Cited By E-mail Article Save as Citation Alert Export Citation Citation Feed

S·F·X

Request Permission
View Record in Scopus
Cited By in Scopus (1)

Disruption of attention by irrelevant stimuli in serial recall

Elke B. Lange,
University of Postdam, Allgemeine Psychologie 1, PO Box 60 15 53, 14415 Potsdam, Germany
Received 2 March 2005; revised 14 July 2005. Available online 29 August 2005.

Abstract

In four experiments the behavioral consequences of an involuntary attentional distraction concerning memory performance was investigated. The working memory model of Cowan (1995) predicts a performance deficit for memory representations that are held in an active state when the focus of attention is distracted by a change in physical properties. In the first experiment, the distraction was realized by an irrelevant tone, which was repeatedly presented and exchanged by another tone unexpectedly. Further experiments explored an analogous effect of irrelevant visual–spatial stimuli. The experiments demonstrated a domain-specific distraction effect: an irrelevant tone change impaired performance in verbal but not in spatial serial recall, whereas a change of position of an irrelevant object disrupted spatial but not verbal memory performance. The

Figure 5.3. Two ejournals from different disciplines published online by the same aggregator (*Science Direct*): same hyperlinking patterns

For the 14 ejournals under consideration, there are 13 different aggregators, which means that there are 13 distinct HTML article layouts. The 13 aggregators are:

1. *Gale Group*
2. *Informaworld*
3. *Ovid*
4. *ProQuest*
5. *Science Direct*
6. *Wilson*
7. *American College of Physicians*
8. *American Medical Association*
9. *Cambridge Journals*
10. *Highwire Press American Heart Association*
11. *Highwire Press British Medical Journal Publishing Group*
12. *Highwire Press Massachusetts Medical Society*
13. *Nature*

Numbers 1 to 6 are what I call *general aggregators*, because they are not connected to any specific publisher, differently from the *home aggregators* presented from number 7 to 13, which are platforms developed by specific publishers for the online publication of their journals only. The distribution of the 14 ejournals across these aggregators is shown in Table 5.1. Based on Table 5.1, it can be observed that the average number of aggregators by which the Linguistics ejournals in the corpus are published is smaller (1.3) than the average in Medicine (2.5)[63]. This disparity in the choice range for linguists when compared to medical researchers can have negative consequences for Linguistics ejournals users, who will have no way out but consulting the HTML article they are interested in via the aggregator that is available to them, even if that particular aggregator does not meet their needs and/or does not fulfill the purpose(s) of the genre (those manifested by the respondents in the contextual based   data   collection).

---

[63] As discussed in Chapter 3 – Methods, these numbers are calculated based on the ejournal access alternatives provided by the University of New South Wales.

Table 5.1
Distribution of the 14 ejournals in the corpus across their aggregator platforms

		Linguistics							Medicine							Total
		JML	SSA	BL	JSLHR	LCP	JP	JFD	NEJM	NM	L	JAMA	AIM	C	BMJ	
General Aggregator	*Gale Group*				+						+				+	3
	Informaworld					+										1
	Ovid										+	+	+	+		4
	ProQuest				+				+		+		+		+	5
	Science Direct	+		+			+	+			+					5
	Wilson				+											1
Home Aggregator	*American College of Physicians*												+			1
	American Medical Association											+				1
	Cambridge Journals		+													1
	Highwire Press American Heart Association													+		1
	Highwire Press British Medical Journal Publishing Group														+	1
	Highwire Press Massachusetts Medical Society								+							1
	Nature									+						1
Total		1	1	1	3	1	1	1	2	1	4	2	3	2	3	

Users of the medical ejournals under consideration here, on the other hand, will always have an alternative choice in case of dissatisfaction and frustration. As Vaughan and Dillon (2006) have revealed, ejournal interface design can influence usage. For the authors (Ibid, p. 502), a complex interface, for example, can increase "the amount of work required of users to navigate and manipulate digital resources" and this workload can go against the very reason why a user would choose the HTML format instead of the PDF format: navigation resources.

Sometimes, however, rather than being too complex, an ejournal interface can be extremely shallow. This is how the aggregator *Wilson* presents its ejournals, represented in the corpus by the *Journal of Speech, Language and Hearing Research*. Via *Wilson*, there are no hyperlinks whatsoever in the HTML version of this journal, which may not be so serious an issue as the omission of all figures. Fortunately, besides *Wilson*, *Journal of Speech, Language and Hearing Research* is also available via another two aggregators – *Gale Group* and *ProQuest* – and, hence, users have more than one option when interested in the ejournal's HTML version. Nevertheless, as will be discussed in the following section (5.2.2), neither *Gale Group* nor *ProQuest* make much use of the medium's affordances and do not follow the hyperlinking patterns of most aggregators in the corpus, which makes the HTML research articles published by these three platforms less typical examples.

In the corpus, the aggregator database *Wilson* is the only one which does not use any hyperlinks in its HTML research articles (one could even ask why it offers an HTML format of the RAs at all, given that it does not explore any of the

medium's potentialities that most of the other aggregators in the corpus present), while the other platforms use several forms of hyperlinking. In the following section, I will explore the use and nature of these hyperlinking forms in HTML research articles in each one of the remaining 12 aggregators in the corpus in an attempt to find out whether there is a hyperlinking pattern across aggregators and ultimately, whether this pattern realizes the communicative purposes of the HTML research article as seen by their users.

5.2.2 Hyperlinking patterns in HTML research articles

In previous literature, hyperlinking patterns have been classified in several ways based on a number of models. One of these models is presented by Landow (1997, p. 12-14) and identifies six forms of linking: 1) unidirectional lexia to lexia, 2) bidirectional lexia to lexia, 3) string (word of phrase) to lexia, 4) string to string, 5) one to many, and 6) many to one (Figure 5.4).

As Figure 5.4 shows, all forms of linking apply to those connections made between two lexias (pages, nodes). This is coherent with the concept of hyperlink presented in Section 5.2 above, i.e., that a hyperlink is that element of a hypertext responsible for implementing the connection between *two nodes*. However, other studies (Kok, 2004) have shown that hyperlinks can also connect strings (to use Landow's terminology) within the *same* node. That is the case of the majority of links found in the HTML research articles in the corpus. For example, every time an in-text citation (i.e. a reference to a previous study) is made, it is presented in the form of a hyperlink (Figure 5.5) which connects the in-text citation to the full reference, presented in the list of references placed at the bottom (or *at the end*) of the document within the *same node*.

143

1) Unidirectional lexia to lexia
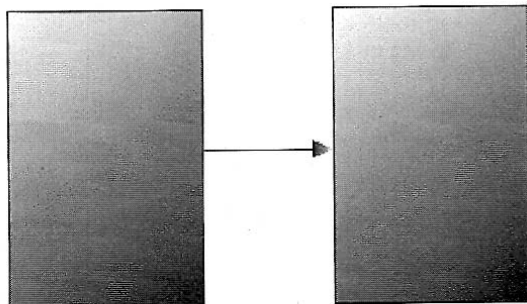
4) String to string
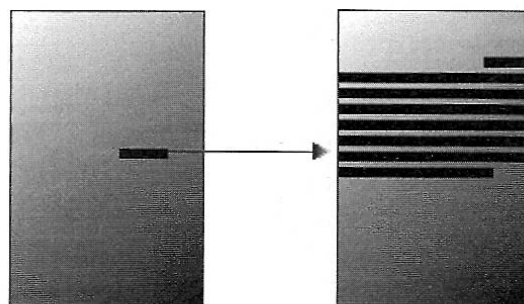
2) Bidirectional lexia to lexia
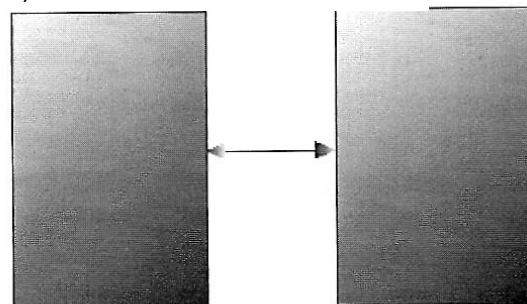
5) One to many
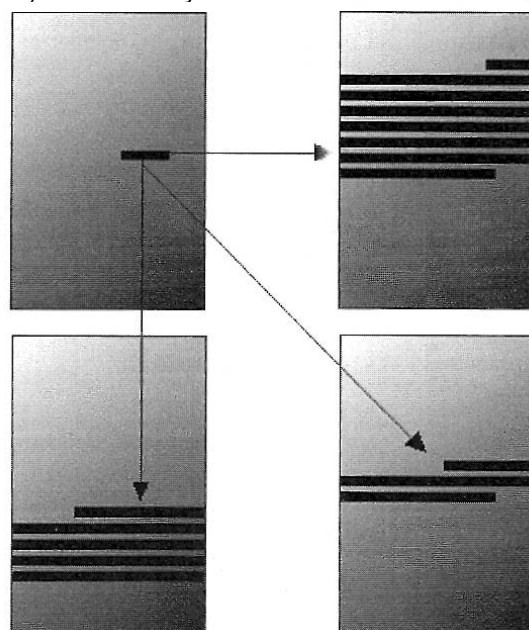
3) String (word of phrase) to lexia
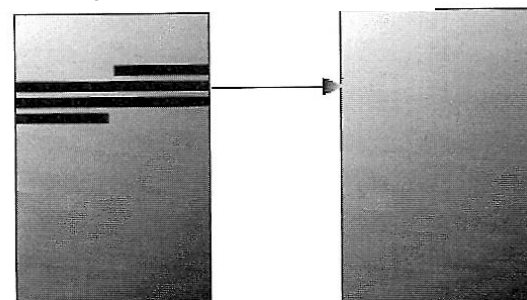
6) Many to one
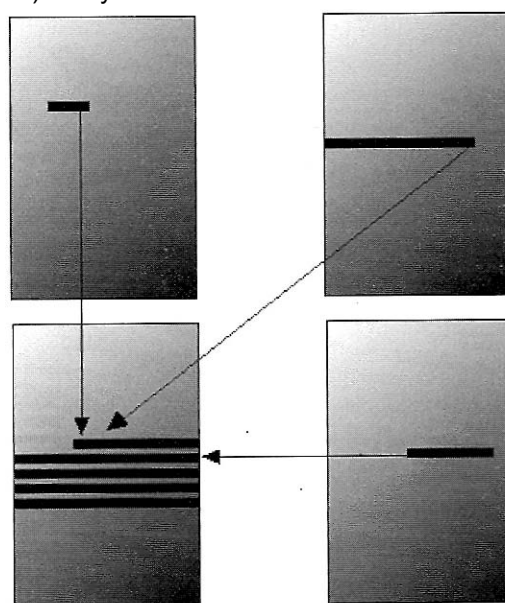
Figure 5.4. Forms of linking according to Landow (1997, p. 12-14)

Figure 5.5. Hyperlinks within the same node in HTML research articles: in-text citations which are connected with full reference at the end list

In the corpus, the intra-node connections can establish several kinds of relationships: bidirectional, unidirectional and many-to-one. As for the inter-node forms of linking proposed by Landow (Ibid.), only two out of his six (Figure 5.4) were identified in the corpus. Hence, in total, five forms of linking were found in the HTML research articles in the corpus, namely: 1) bidirectional string to string within the same lexia, 2) unidirectional string to string within same lexia, 3) unidirectional many to one within same lexia, 4) unidirectional many to one lexia to lexia, and 5) unidirectional string (word or phrase) to lexia. Following Landow's (Ibid.) design, a tentative representation of these five forms of linking is shown in Figure 5.6.

If activated all together, these five linkage patterns would result in something similar to the image shown in Figure 5.7, which, for reasons of space dimensions, shows a shortened version of a Linguistics HTML research article from the corpus.
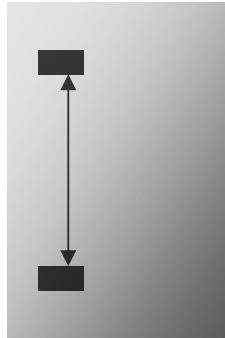
1) Bidirectional string to string within the same lexia
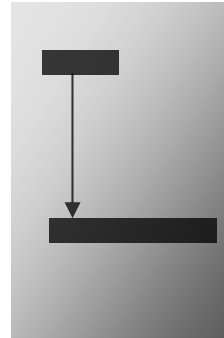
2) Unidirectional string to string within the same lexia
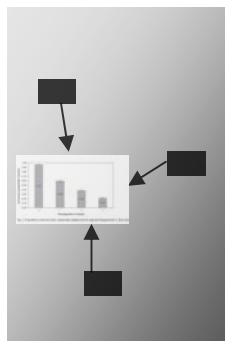
3) Unidirectional many to one within the same lexia
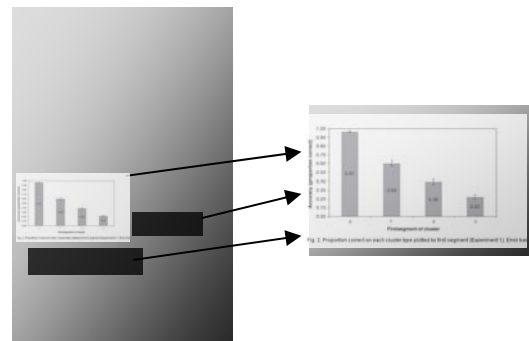
4) Unidirectional many to one lexia to lexia
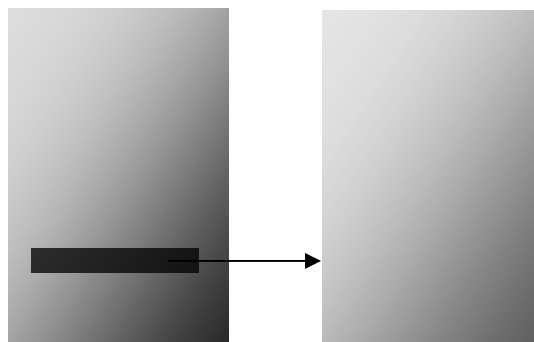
5) Unidirectional string (word or phrase) to lexia

Figure 5.6. Forms of linking in HTML research articles in the corpus

Figure 5.7. Types of hyperlinks in an HTML research article

Hyperlink types 1, 2 and 3 (Figure 5.6), which establish connections within the same lexia, can either connect section headings with the actual sections or connect reference expressions to their referents. The first kind of connection helps the HTML article users to move across the different sections of article more quickly. They work as an article outline and thus also help the user visualize which part of the HTML research article he is reading or browsing (Figure 5.8). Because these links highlight the structure of the HTML research article, i.e., the different sections that constitute the article, I will call them *structure hyperlinks*.
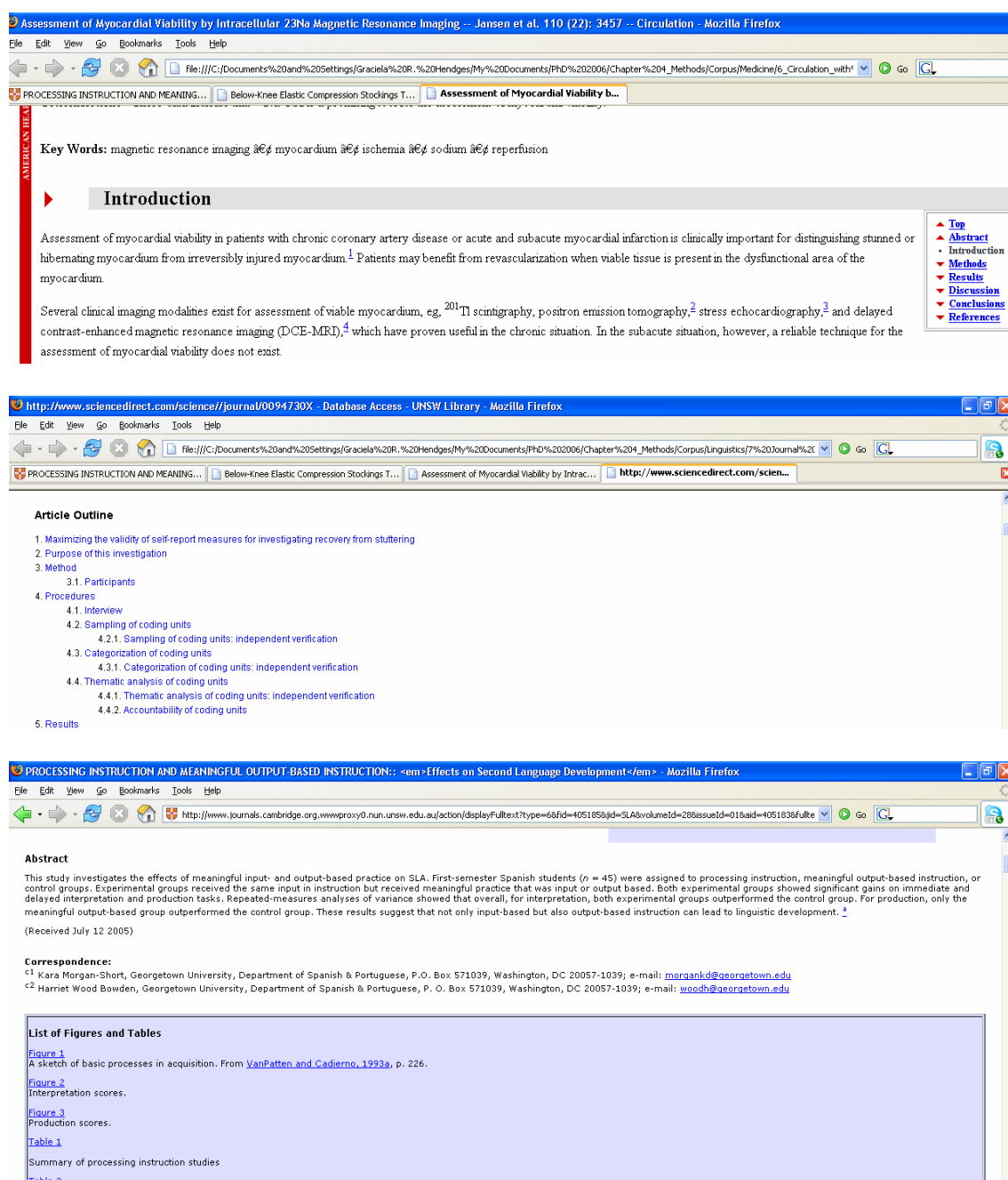
Figure 5.8. Examples of *structure hyperlinks*

The second kind of links within the same lexia establishes the connection between reference expressions to their referents, which can either be in-text references to images that, when activated, take the user to where the image within that lexia is, or in-text citations which are connected to the full reference

of the cited text at the end list presented at the bottom of the lexia. Because of their referential nature, I called this type of links *reference hyperlinks*. It has been argued that the kind of connection reference links establish already existed in print RAs, although these were abstract connections, while in the electronic medium they have become (virtually) material, visible and touchable and, therefore, the reading experience when it comes to HTML research articles can be very different from reading print RAs. In some aggregators, reference links are bidirectional, a tool which can speed up the online reading process significantly, and thus, help realize the purpose of the genre of *rapid viewing of article content*. This purpose is equally realized by structure hyperlinks.

The main function of structure and reference hyperlinks (which correspond to hyperlink types 1, 2, and 3 in Figure 5.6 and connect different strings within the same lexia), thence, seems to be to speed up the location of information in an HTML research article, and, consequently, the online reading process, by making already existing ties between different parts of an RA more visible. In this sense, structure and reference hyperlinks do not add new information or materials to the HTML research article when compared to its print counterpart, but add connections between strings of information.

On the other hand, hyperlink types 4 and 5 in Figure 5.6, that is to say, those hyperlinks that establish connections *between* two lexias, add new information and materials to HTML research articles. The nature of the information they add varies. One type of hyperlinks gives access to higher quality versions of the images (figures, tables, charts, graphs, pictures, etc.) presented in an HTML research article. Most of the times, more than one
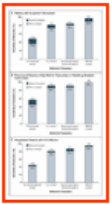
resolution level and format for the same image are offered, with quality standards that are better than those of the image available in the print version of the RA (Figure 5.9). For this reason, I have named them *quality hyperlinks*.

NEJM (HPMMS)

Added Value of Second Culture for Sensitivity of De

The incremental benefit of a second smear for acid-fast ba
tuberculosis but offered no added value among prescreen

Figure 2. Incremental Benefit of O

The percentage of cases detected
smear or culture. For patients with
smears stained with the Ziehlâ€"Ne
samples from the hospitalized patien
Peruana Cayetano Heredia, becaus

View larger
version (24K):
[in this window]
[in a new window]

Get PowerPoint
Slide ▶

mycobacterial cultures, and 1 Löwenstein-Jensen culture) from
defined as a positive culture according to at least one method f
methods yielded negative results or in which two were negative
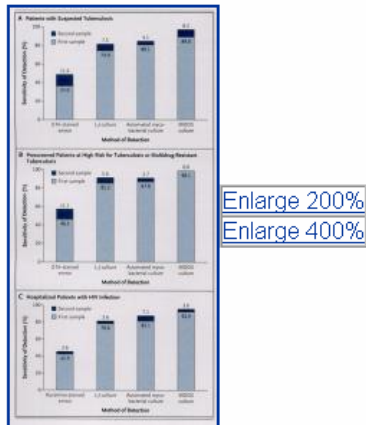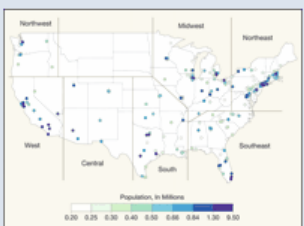only cause of the positive culture.36

Enlarge 200%
Enlarge 400%

Figure 2. Incremental Benefit of One and Two Sputum Cultures as
The percentage of cases detected during examination of the first sr
tuberculosis (Panel A) and prescreened patients at high risk for tub

JAMA (AMA)

infections [464-466, 480-487]), and hospitalizations ca
1999-2002 appear in Table 1.

View this table: | Table 1. Percentag
[in this window]
[in a new window]
[as a PowerPoint slide]

The study population includes 11.5 million Medicare enr
larger than 200 000. Of these 204 counties, 90 had dai
year. The locations of the 204 counties appear in Figur
for the counties.[10-11]

Figure 1. US Co

View larger version (102K):
[in this window]
[in a new window]
[as a PowerPoint slide]

NEJM (*ProQuest*)

Figure 5.9. Examples of *quality hyperlinks*

Another kind of hyperlink between two lexias are those which give *immediate* access to the author of the HTML research article, via her/his email address, to the full text of studies cited in an HTML research article and give access to additional material in the format of what journals have denominated *Supplements*. Because they provide access to such less reachable 'goods and services' in print, I called them *access hyperlinks*.

Particularly in the case of the immediate access to the full text of references made in an HTML research article, this feature is highly valued by ejournal users and is usually indicated as one of the biggest advantages of the HTML research article over print and PDF RAs. Back in 2000 when I took my masters research on free online journals (Hendges, 2001), the amount of references whose full texts were available via hyperlinks was very low and I speculated that one reason for that scarceness was that most of the texts being referred to did not exist in digital format yet. Today, in areas such as Medicine that rely mostly on journal publications as previous literature, everything is online and through such projects as CrossRef, the majority of texts (when not 100% of them) in the reference list is immediately accessible (Figure 5.10). Even in Linguistics the number of direct references is high.

20. Eckart RE, Uyehara CF, Shry EA, Furgerson JL, Krasuski RA. Matrix metalloproteinases in patients with myocardial infarction and percutaneous revascularization. J 2004; 17: 27â€"31.[CrossRef][Medline] [Order article via Infotrieve]

21. Zeng B, Prasan A, Fung KC, Solanki V, Bruce D, Freedman SB, Brieger D. Elevated circulating levels of matrix metalloproteinase-9 and -2 in patients with symptom: disease. Intern Med J. 2005; 35: 331â€"335.[CrossRef][Medline] [Order article via Infotrieve]

22. Blankenberg S, Rupprecht HJ, Poirier O, Bickel C, Smieja M, Hafner G, Meyer J, Cambien F, Tiret L. Plasma concentrations and genetic variation of matrix metallop prognosis of patients with cardiovascular disease. Circulation. 2003; 107: 1579â€"1585.[Abstract/Free Full Text]

23. Pollanen PJ, Karhunen PJ, Mikkelsson J, Laippala P, Perola M, Penttila A, Mattila KM, Koivula T, Lehtimaki T. Coronary artery complicated lesion area is related to polymorphism of matrix metalloproteinase 9 gene: an autopsy study. Arterioscler Thromb Vasc Biol. 2001; 21: 1446â€"1450.[Abstract/Free Full Text]

24. Jaffer FA, Weissleder R. Molecular imaging in the clinical arena. JAMA. 2005; 293: 855â€"862.[Abstract/Free Full Text]

25. Tung CH, Bredow S, Mahmood U, Weissleder R. Preparation of a cathepsin D sensitive near-infrared fluorescence probe for imaging. Bioconjug Chem. 1999; 10: 892â€"896.[CrossRef][Medline] [Order article via Infotrieve]

26. Tung CH, Mahmood U, Bredow S, Weissleder R. In vivo imaging of proteolytic enzyme activity using a novel molecular reporter. Cancer Res. 2000; 60: 4953â€"4958.[Abstract/Free Full Text]

27. Weissleder R, Tung CH, Mahmood U, Bogdanov A Jr. In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. Nat Biotechnol. 1999; 17 375â€"378.[CrossRef][Medline] [Order article via Infotrieve]

28. Bremer C, Tung CH, Weissleder R. In vivo molecular target assessment of matrix metalloproteinase inhibition. Nat Med. 2001; 7: 743â€"748.[CrossRef][Medline] [C Infotrieve]

29. Weissleder R, Ntziachristos V. Shedding light onto live molecular targets. Nat Med. 2003; 9: 123â€"128.[CrossRef][Medline] [Order article via Infotrieve]

30. Ntziachristos V, Ripoll J, Wang LV, Weissleder R. Looking and listening to light: the evolution of whole-body photonic imaging. Nat Biotechnol. 2005; 23: 313â€"320.[CrossRef][Medline] [Order article via Infotrieve]

31. Chen J, Tung CH, Allport JR, Chen S, Weissleder R, Huang PL. Near-infrared fluorescent imaging of matrix metalloproteinase activity after myocardial infarction. Cir 1800â€"1805.[Abstract/Free Full Text]

32. Kridel SJ, Chen E, Kotra LP, Howard EW, Mobashery S, Smith JW. Substrate hydrolysis by matrix metalloproteinase-9. J Biol Chem. 2001; 276: 20572â€"20578.[Abstract/Free Full Text]

33. Chen J, Tung CH, Mahmood U, Ntziachristos V, Gyurko R, Fishman MC, Huang PL, Weissleder R. In vivo imaging of proteolytic activity in atherosclerosis. Circula. 2766â€"2771.[Abstract/Free Full Text]

Figure 5.10. Hyperlinked reference list in HTML research articles: immediate access to previous literature

Finally, a third new kind of information added to HTML research articles through hyperlinks are the supplementary materials (Figure 5.11). Five out of the 12 aggregators offer this kind of hyperlinks, and most of them are the kind of platforms I am calling *home aggregators* here (Table 5.1): *Ovid* (which publishes only the medical ejournals in the corpus), *Highwire Press Massachusetts Medical Society* (which publishes the NEJM online), *Nature* (which publishes NM online)*, Highwire Press American Heart Association* (which publishes *Circulation* online) and *Highwire Press British Medical Journal Publishing Group* (which publishes BMJ online). Currently, as far as Linguistics and Medicine are concerned, supplements are an exclusive feature of medical journals. Because the maximum number of pages for a medical RA to occupy in a print journal is very limited and has been decreasing (in part as a consequence of impact factor policies, because the more articles published per journal, the higher the chances that the journal will be cited), many times

authors have to leave a lot of methodological details as well as results out of the

print RA, as reported by [MI]:

> [MI] If you look at the highest impact factor journals, most of them have the shortest articles. But sometimes you can't tell the whole story in 4 pages. […]people cramp material into whatever page limit that you have and then I mentioned to you also what has become increasingly common is that you provide additional information that does not fit into your space limit or page limit as supplementary material and sometimes that supplementary material is larger than your actual application.

With the Internet, extra data can be published online as supplements,

without affecting the print edition of journals. It is not clear at this point to what

extent ejournal users value and use these extra materials. In the online

questionnaire they were not mentioned as a reason why users access HTML

research articles, maybe because to see extra materials users do not

necessarily have to *visit* the HTML article, since aggregators (four out of the five

in the corpus) are also placing links to supplements in the table of contents of

ejournals (Figure 5.11).

Circulation

Nature Medicine

Figure 5.11. Examples of aggregators and ejournals that include supplementary materials in HTML research articles

Still among the categories of links that connect two lexias are those by which related articles and letters and citation records of the HTML research article under consideration can be reached. I see these links as contributing to show the relevance of the HTML research article in the field and therefore I named them *status hyperlinks*. Similarly to what happens with the hyperlinks to supplements, the links to related letters and related articles an exclusive feature of those aggregators which distribute the medical ejournals, specifically of those I call *home aggregators* (Table 5.1): *Highwire Press Massachusetts Medical Society*, *Nature*, *American Medical Association*, *American College of Physicians*, *Highwire Press American Heart Association* and *Highwire Press British Medical Journal Publishing Group*. It seems that this uniformity among the home aggregators is influenced by the competitive race that journal publishing has become (boosted in part by the *publish or perish* kind of policy as well by impact factor standards). The links to related letters and related articles appear at the bottom of the page and/or in the right-hand columns that are usually present in HTML research articles at the top of the page (Figure 5.12 (a) and (b)).

(a) JAMA

42. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res.* 1993;2:221-226. FULL TEXT | ISI | PUBMED

43. McTigue KM, Garrett JM, Popkin BM. The natural history of the development of obesity in a cohort of young US adults between 1981 and 1998. A
FREE FULL TEXT

RELATED ARTICLES

This Week in *JAMA*
JAMA. 2004;291:1415.
FULL TEXT

Clinical Outcomes and Costs With the Levonorgestrel-Releasing Intrauterine System or Hysterectomy for Treatment of Menorrhagia: Ran
Ritva Hurskainen, Juha Teperi, Pekka Rissanen, Anna-Mari Aalto, Seija Grenman, Aarre KivelÃ×, Erkki Kujansuu, Sirkku Vuorma, Merja Yliskoski, and Jor
JAMA. 2004;291:1456-1463.
ABSTRACT | FULL TEXT

Evaluating Gynecological Surgical Procedures: Trials and Tribulations
Roy M. Pitkin and James R. Scott
JAMA. 2004;291:1503-1504.
EXTRACT | FULL TEXT

Hysterectomy
Janet M. Torpy, Cassio Lynm, and Richard M. Glass
JAMA. 2004;291:1526.
EXTRACT | FULL TEXT

THIS ARTICLE HAS BEEN CITED BY OTHER ARTICLES

Oophorectomy in Premenopausal Women: Health-Related Quality of Life and Sexual Functioning
Teplin et al.
Obstet Gynecol 2007;109:347-354.
ABSTRACT | FULL TEXT

AIM

32. Klee AL, Maldin B, Fosmit T, Mostashari F, Fine A, Layton M, et al. Long-term prognosis for clinical West Nile virus infection. Emerg Infect Dis. 2004;10:1405-11. [PMI

33. Catalog of preference scores. CEA Registry: standardizing the methods and practices of cost-effectiveness analysis. Harvard Center for Risk Analysis and Harvard School
http://www.hsph.harvard.edu/cearegistry on 2 October 2003.

34. **Vamvakas EC.** Uses and sources of data on long-term survival after blood transfusion. Transfus Med Rev. 2003;17:194-208. [PMID: 12881780]. [Medline]

35. **Geist A, Lai QJ.** Gen-Probe Incorporated (GRPO). Baird U.S. Equity Research. Milwaukee, WI. 28 May 2004. Accessed at http://www.rwbaird.com/research on 7 Novemb

Related articles in Annals:

Editorials
Meeting Transfusion Safety Expectations
James P. AuBuchon
Annals 2005 143: 537-538. (in) [Full Text]

Summaries for Patients
The Cost-Effectiveness of Screening the U.S. Blood Supply for West Nile Virus
Annals 2005 143: I-44. (in) [Full Text]

This article has been cited by other articles:

JAMA   ▸HOME
West Nile Virus Transmission Through Blood Transfusion--South Dakota, 2006
JAMA, March 21, 2007; 297(11): 1186 - 1188.
[Full Text] [PDF]

JOURNAL WATCH INFECTIOUS DISEASES   ▸HOME

NEJM

15. Ettinger B, San Martin J, Crans G, Pavo I. Differential effects of teriparatide on BMD after treatment with raloxifene or alendronate. J Bone Miner Res 2004;19:
16. Lane NE, Sanchez S, Modin GW, Genant HK, Pierini E, Arnaud CD. Parathyroid hormone treatment can reverse corticosteroid-induced osteoporosis: results of
Clin Invest 1998;102:1627-1633. [Free Full Text]

Appendix

In addition to the principal investigators, the following persons participated in the PaTH study: *Columbia University* â€" K. Lee and J. Sliney (study coordinators); *Min
Center* â€" V. Wyum and N. Michaels; *University of Pittsburgh Medical Center* â€" J.L. Ryan (study coordinator) and J.M. Wagner; *Maine Center for Osteoporo
â€" L. Fowler and D. Storm (study coordinators); *University of California, San Francisco* â€" T. Hue (project director), L. Palermo (statistician), D. Sellmeyer, and
Safety Monitoring Board â€" L. Raisz, S. Hui, R. Recker, D. Kiel, and D. Hanley.

Related Letters:

Alendronate and Parathyroid Hormone
Muldowney F. P., Black D. M., Sellmeyer D., Rosen C. J., Cosman F., Nieves J., Lindsay R.
Extract | Full Text | PDF
N Engl J Med 2005; 353:2618-2619, Dec 15, 2005. **Correspondence**

This article has been cited by other articles:

- Shoback, D. (2007). Update in Osteoporosis and Metabolic Bone Disorders. *J. Clin. Endocrinol. Metab.* 92: 747-753 [Abstract] [Full Text]
- Gafni, R. I., Baron, J. (2007). Childhood Bone Mass Acquisition and Peak Bone Mass May Not Be Important Determinants of Bone Mass in Late Adulthood.

BMJ

Drugs 1993;46: 253-6.

37. Szczeklik A, Sanak M. Molecular mechanisms in aspirin-induced asthma. *Allergy Clin Immunol Int* 2000;12: 171-6 [CrossRef]

38. Szczeklik A, Nizankowska E, Duplaga M. Natural history of aspirin-induced asthma. *Eur Respir J* 2000;16: 432-6 [Abstract]

(Accepted 14 November 2003)

Related Articles

Aspirin induced asthma: Clinical relevance of finding was not explained
Indu Sivanandan and Stephen M Robinson
BMJ 2004 328: 1076. [Extract] [Full Text]

Aspirin induced asthma: Effect size needs to be clarified
Robert D Southward
BMJ 2004 328: 1076. [Extract] [Full Text]

Aspirin induced asthma is more common than previously suggested
BMJ 2004 328: 0. [Full Text]

This article has been cited by other articles:

(Search Google Scholar for Other Citing Articles)

- Knowles, S. R, Drucker, A. M, Weber, E. A, Shear, N. H (2007). Management Options for Patients with Aspirin and Nonsteroidal Antiinflamma

(b)
JAMA

Circulation

Figure 5.12. Examples of eRAs which present links to related letters and related articles at the end of the page (a) and in right-hand column (b)

Right below the "Related articles" and "Related letters" links, in the body of the HTML research article (Figure 5.12 (a)) are the other kind of *status hyperlinks*, those links which indicate citation records of the HTML research article under consideration. It is interesting to notice that when status hyperlinks appear in the body of the HTML research articles, at the bottom of the page, the

title of the section that introduces them follows the same design characteristics (same color, size, and typeface) of the other section titles in the RA, indicating that they are of the same hierarchy and as much a section of the HTML research article as the Introduction, the Methods, the Results, the Discussion and References sections are.

Yet when they appear both in the body of the HTML research article as well as in right-hand column (Figure 5.12 (b)), it is less clear whether they are to be considered a part of the HTML research article or whether they are to be read as an extra service. That is because it is also in the right-hand column where the last kind of inter-lexia links are provided: those which offer access to a variety of services related to the HTML research article, such as the possibility of downloading the PDF format, searching for other articles by the same author(s) of the HTML research article (sometimes with *Google Scholar*), directly sending an HTML research article through email via "Send to a friend" (JAMA) or "Email this article to a friend" (Circulation) links, and "Request permissions" to use copyright materials. These links are never in the body of the HTML research article, but always clearly separated in this right-hand column, working as a sort of table of contents rather than adding new information and material to the HTML research article per se.

The presence and frequency with which each one of these hyperlink types is used across the 13 aggregator platforms is shown in Table 5.2. As can be noted, all types of hyperlinks were identified with relative frequency in the HTML research articles in the corpus, except for those exemplars published by *ProQuest* (18%), *Gale* (9%) and *Wilson* (0%).

Table 5.2
Types, presence and frequency of hyperlinks used in the eRAs across the 13 aggregator databases in the corpus

		Structure links		Reference links		Quality links	Access links			Status links		Service links	Total
		Article outline	Outline of images	Image reference to in-text image	In-text references to reference list	In-text image to higher resolution image	Author email	Full text of cited study	Data supplement	Related articles/letters	Citations	Other services	
General Aggregator	*ProQuest*	−	−	−	−	+	−	−	−	−	−	+	2 (18%)
	ElsevierSD	+	−	+	+	+	+	+	−	−	+	+	8 (73%)
	Ovid	+	+	+	+	+	−	+	+	−	+	+	9 (82%)
	Gale	−	−	−	−	−	−	−	−	−	−	+	1 (9%)
	Wilson	−	−	−	−	−	−	−	−	−	−	−	0
	Informaworld	+	+	+	+	+	−	+	−	−	−	+	7 (64%)
Home Aggregator	*Cambridge Journals*	+	+	+	+	+	+	+	−	−	+	+	9 (82%)
	Highwire Press Massachusetts Medical Society	−	+	+	+	+	+	+	+	+	+	+	10 (91%)
	Nature	+	+	+	+	+	+	+	+	+	−	+	10 (91%)
	American Medical Association	+	−	+	+	+	+	+	−	+	+	+	9 (82%)
	American College of Physicians	+	+	+	+	+	+	+	−	+	+	+	10 (91%)
	Highwire Press American Heart Association	+	−	+	+	+	+	+	+	+	+	+	10 (91%)
	Highwire Press British Medical Journal Publishing Group	−	−	+	+	+	+	+	+	+	+	+	10 (91%)
Total		8 (62%)	6 (46%)	10 (77%)	10 (77%)	11 (85%)	8 (62%)	10 (77%)	5 (38%)	6 (46%)	8 (62%)	12 (92%)	

What Table 5.2 clearly shows is that home aggregators make higher use of hyperlinks than general aggregators. Home aggregators are a recent development if compared to general aggregators. The latter appeared first and maybe home aggregator designers have learned from the mistakes of general aggregators and developed databases with better interfaces that correspond to user needs as well as the purpose of the genre.

If the hyperlink patterns are contrasted across the disciplines, the outstanding results is that all types of hyperlinks are more frequently found in Medical HTML research articles than in Linguistics HTML research articles (Table 5.3). In addition, it can be observed that there are more types of hyperlinks in Medical than in Linguistics HTML research articles (Table 5.3)

Table 5.3
Frequency of types of hyperlinks per discipline

	Structure links		Reference links		Quality links	Access links			Status links		Service links
	Article outline	Outline of images	Reference to image to in-text image	In-text references to reference list	In-text image to higher resolution image	Author email	Full text of cited study	Data supplement	Related articles/letters	Citations	Other services
L	62%	13%	62%	62%	77%	56%	62%	0%	0%	56%	87%
M	52%	38%	66%	66%	89%	49%	66%	45%	42%	58%	100%

If the percentages of the types of hyperlinks found in the HTML research articles is compared to the percentages of the types of hyperlinks ejournals users in linguistics and medicine expect to find, as revealed by the contextual analysis (Table 4.12 in Chapter 4), the figures are closely related in terms of hyperlinks to references and to supplementary data, but are quite contrasting in relation to hyperlink to figures and tables and to contact details (Table 5.4).

Table 5.4
Comparison between contextual an textual analyses in terms of hyperlink types

Types of hyperlinks	Contextual analysis	Textual analysis
References	84	77
Data supplement	43	38
Figures and tables	32	85
Contact details	5	62

These results may indicate the HTML research article is still in transition.

In general, however, the types of hyperlinks found in HTML research articles match the set of communicative purposes of the genre found in the contextual analysis perfectly, except the purpose of being a substitute for the PDF format, which does not apply in the present research because the question about software availability was not included in the questionnaire, nor did any of the respondents express that they read HTML research articles because they have no PDF readers to view the PDF version of articles. Table 5.5 shows the types of hyperlinks along with the percentage of hyperlinks incidence across the 13 aggregators and the purpose of the HTML research article they realize.

Table 5.5
Types of hyperlinks and the purposes they realize

Types of hyperlinks	Purpose of the HTML article	% across aggregators
Structure links and *Reference links*	allow rapid viewing of article content	77%
Access links to full text of cited study and *Status links* to related articles	offer immediate access to previous literature	77%
Quality links	offer access to quality graphics	85%
Access links to data supplements	offer access to supplementary information	38%

It has been suggested in previous literature that "it might be useful to define the form of a hyperdocument (and thus the overall genre) by the pattern of links it exhibits" (Cowston & Williams, 2000, p. 210).

In this sense, if hyperlinks realize structure and structure realizes purpose (Figure 5.1 in this chapter), then the purposes of the HTML research article as far as textual analysis is concerned match those obtained in the contextual analysis. The consistency of purposes across the two levels of the genre analysis eliminates the need for genre repurposing.

5.3 Summary of textual analysis

The main results obtained in this chapter are:

- The forms of linking found in the corpus are different from those proposed in previous literature on hypertext (Landow, 1997);

- There are considerable differences across aggregators in relation to the hyperlinking patterns shown in HTML research articles;

- The variety of hyperlinking patterns presented by home aggregators is considerably higher (in average, they present 88% of the total amount of hyperlink types found in the genre) than the one of general aggregators (in average, 41%). This suggests that the HTML research articles published by home aggregators are more prototypical examples of HTML articles than those published by general aggregators;

- Given that 6 (86%) of the 7 home aggregators in the corpus are from the area of Medicine, there are more chances for medical researchers to have access to prototypical HTML research articles than linguists;

- *Access hyperlinks* to data supplements were found exclusively in medical HTML research articles, indicating that this genre is sensitive to disciplinary idiosyncrasies; in Hasan's Generic Structure Potential model, this kind of links would be optional rather than obligatory elements in the structure of the HTML research article.

The most relevant outcome of the textual analysis, however, is that it corroborates the argument that in online hypertextual genres, hyperlinks can be responsible for giving shape to these genres and ultimately, realizing communicative purpose. Omitting the hyperlinks found in the HTML research articles would make the texts unsuitable for its intended and expected uses, which would change the genre.

In this view, it is the affordances of the medium that create and constrain the development of a new genre. Hence, it may be necessary to consider *medium* at the level of genre, as one of the generic features which contribute for the realization of communicative purpose.

Chapter 6 Conclusion

> "As a society we are over-stimulated: we have vitamins in our water now, we got caramel and chocolate in our coffee, we have cheese in our crust, cameras in our phones, TVs in our cars, we got blackberries, iPods, TiVo, PSP, GPS… thank God your kids have ADD! They are gonna need it to deal with all this stuff! Your kids aren't hyper, they are paying attention!"
>
> (Stand-up comedian Nick Griffin on July 20, 2007, in Late Show with David Letterman)

The central objective of this study was to verify whether the HTML research article is a new genre or not. To achieve this objective, I employed the methodological framework proposed by Askehave and Swales (2001) and by Swales (2004) for genre analysis and genre classification, in which the authors argue for the need of contextual analysis along with textual analysis in genre studies. According to them, only the combined result of these two levels of analysis would reveal the *communicative purpose* of genres, communicative purpose being the main genre classification criterion.

Following Askehave and Swales's (2001) and Swales's (2004) argument, the present work developed both types of analysis: contextual and textual. The contextual analysis consisted of examining answers provided by researchers from the disciplines of Linguistics and Medicine to an online questionnaire about their ejournal and HTML research article usage practices.

The textual analysis consisted of identifying, describing and classifying hyperlink patterns in HTML research articles published in ejournals from the same disciplinary domains: Linguistics and Medicine.

Both analyses evidenced that the medium of the Internet affects the reason consumers and producers use HTML research articles as well as how HTML research articles are structurally and functionally organized. Both the usage practices and the structural organization of HTML research articles are different from the usage practices and structural organization of print RAs as reported in previous literature about print RAs. This means that the primary *communicative purposes* of the HTML research article is different from the communicative purposes of the print RA: while the first aims at allowing rapid viewing of article content, offering immediate access to previous literature, to quality graphics/images and to supplementary information, besides serving as a substitute for the PDF format, the latter, according to previous literature, aims at reporting research experiments, disseminating scientific information, advancing the state of knowledge in a field, controlling quality of research, providing recognition for authors and career upgrade, and securing research funding.

It could be argued, however, that these latter are also purposes of the HTML research article as well as of other scientific genres, such as conference presentations, research proposals, conference abstracts. This suggests that the set of purposes of the print RA are on a different level of those identified for the HTML research article in the sense that the former are *epistemological* purposes while the purposes of the HTML research article are more practical. This issue adds new complexity to the communicative purpose identification

methods proposed by Askehave and Swales (2001), given that identification of communicative purpose alone does not suffice: it has to be further characterized and categorized (e.g. how abstract is the communicative purpose?) so that the genre classification process can be more precise. Such characterization and categorization of the levels of communicative purpose, however, fall outside the scope of this work, but should be considered in future genre classification accounts.

In light of these findings and within the perspective of the theoretical and methodological framework adopted here, my thesis is that *the HTML research article is a new genre.* This argument answers the main question investigated in this study: whether the HTML research article is a new genre or not.

In regards to specific objectives listed in Chapter 1, section 1.4 (p. 14), the results obtained for each one are the following:

7. *Does the communicative purpose of the HTML research article echo the communicative purpose of the traditionally print RA, based on contextual analyses, i.e., on interviews with HTML research article users (producers and consumers)?*

No. As reported above, HTML research article users expressed that they use these texts with purposes that are different from those why they use PDF RAs. Regarding PDF RAs, it can be argued on the basis of previous literature that they (PDF RAs) are used with the same purpose of print RAs.

8. *Does the communicative purpose of the HTML research article echo the communicative purpose of the traditionally print RA, based on textual analyses, i.e., on an investigation of the HTML research article's schematic structure?*

No. Although all the sections (IMRD) and textual information of print RAs are also present in HTML research articles, the schematic structure of HTML research articles is defined primarily by the types of hyperlinks they present. The functional units or moves of HTML research articles are, in this sense, different from the functional units of print RAs, namely, the IMRD pattern. Instead of the IMRD organization, HTML research articles are organized around two forms of linking: 1) within lexia and 2) between lexia. *Within lexia* links allow rapid viewing of article content. *Between lexia* links offer access to materials and resources which are not available in print RAs, namely, higher quality images; data supplements, author's email, full texts of cited literature and other online services. These hyperlinks realize the structure of HTML research articles and, as a consequence, considering that *structure realizes purpose*, it can be argued that the communicative purpose of HTML research articles does not echo the communicative purpose of print RAs.

9. *What features are consistent and what features vary (Hasan, 2004) in HTML research articles at the level of genre across the disciplines of Linguistics and Medicine?*

The results obtained in this study show that the new scientific genre of the HTML research article is sensitive to disciplinary variations just as print scientific genres have been shown to be.

Three main disciplinary differences were observed: 1) number of types of hyperlinks per article varies across them; 2) some types of hyperlinks are discipline-specific; and 3) medical HTML research articles present *PowerPoint* slides for images.

In relation to the first difference, it was found that there are less types of hyperlinks per HTML research article in Linguistics articles (an average of 5.3 types per article) than in Medicine (an average of 6.8 types per article). Maybe the lower variety of hyperlinks in Linguistics HTML research articles explains why linguists *click* on hyperlinks less frequently (18%) when compared to medical researchers (36%) (see Figure 4.11 in Chapter 4, p. 122).

The second difference is that two types of hyperlinks are used exclusively in medical HTML research articles: links to *Data supplements* and links to *Related articles*, suggesting that these two types of links are discipline-oriented features of HTML research articles. As with other fields (Berkenkotter, 2007; Ayers 2007), data supplements in the Medical ejournals generally contain further details about the methods of a given research. The emergence of this resource in ejournals, particularly in the journal *Nature*, is explained by Ayers (2007, p. 7):

> methods and materials have become too complex and specialized to be relevant to the average Nature reader (albeit informed). In other words, the audience to whom information concerning methods and materials is meaningful is being continually reduced. Further, to describe and include all these more complex methods would fill up a great amount of space on the main paper which would explain Nature's indications ("Guide to Authors") to authors as how to present this information: unless methods are brief (less than 200 words in total of a text extending to up to 1300 words), they should
> be described at the end of the text in a Section 3. If this last section exceeds 800 words, authors are encouraged to submit the extra-material as on-line 'Supplementary Information', which accompanies the published paper on Nature's web-page.

The absence of this type of links in Linguistics could suggest that methodological detailing is less relevant in this area. However, as reported by

the participants from Linguistic, they would welcome links to extra information about methods:

LB# 11 Espero encontrar hiperlinks de referências ou maiores discussões sobre determinado conceito, dados ou explicações complementares sobre conceitos ou metodologia.

LB#13 O tipo de hiperlinks que espero encontrar são: email do autor, página pessoal do autor, figuras e tabelas, referências/citações, dados e instrumentos da pesquisa relatada no artigo, links para banco de dados e programas, links que dão acesso aos textos citados.

With this new possibility presented by ejournals aggregators, publishers and editors in the area of Linguistics become aware of users' expectancies, they may start to encourage authors to submit supplementary information.

Another difference observed in the cross-disciplinary analysis of HTML research articles is that Medicine articles present PowerPoint *slides as one alternative with which to visualize the images in the articles.* PowerPoint *slides started to appear in the Medical ejournals in the corpus in 2002. Interestingly, in that year, LaPorte et al. wrote a comment article in the* British Medical Journal *entitled* Papyrus to PowerPoint (P 2 P): metamorphosis of scientific communication. *In this paper, they suggest that ejournals should give way to PowerPoint presentations with the argument that it is the lingua franca of science given that as much as 95% of presentations use PowerPoint and*

Each day 30 million PowerPoint presentations are produced. PowerPoint is on 250 million computers worldwide. There are four million PowerPoint lectures on the web, and the number is increasing logarithmically (Google search). Reasons for the rapid spread are obvious. PowerPoint is easy,

relatively inexpensive, and fast, and scientists control production. (p. 1479)

In the journals of the corpus, so far power PowerPoint slides are used exclusively for the presentation of images. Given the importance of images for the area, having clearly defined high resolution graphics is crucial for research communications in the field. According to the [MI], these PowerPoint slides are also incorporated into lessons and courses by the medical community. The main tool used for the delivery of these lessons is PowerPoint slides, so if the images are already in the right format, medical professionals save time.

Nevertheless, as a suggestion for further research, it would be interesting to accompany the evolution of medical HTML research articles and see to what extent LaPorte et al.'s wishes come true. It was point out earlier in this thesis that one new ejournal, the *Journal of Visualized Experiments*, has already proposed a new form or article, the video article. Both types of articles, PowerPoint articles or video articles will require or already demand new abilities from researchers – producers and consumers because manipulation of the media along with manipulate of research results will be necessary.

10. *To what extent and how differently do the contexts of Brazil and Australia affect the HTML research article at the level of genre?*

It was discussed in Chapter 4 that there are significant differences between Brazil and Australia in terms of Internet access and ejournal access. Nevertheless, despite of these differences, if cross-country ejournals reading practices are contrasted, there seem to be no relevant disparities. This seems

to suggest that the pressures of science, of the *publish or perish* policies transcend technological shortcomings.

11. *What is the role of the medium, i.e. of the technology of the Internet, in shaping the HTML research article?*

The role of the context of the electronic medium of the Internet in shaping HTML research articles is central. If it were not for the medium, HTML research articles would probably not exist. The communicative purpose of the genre HTML research article is realized by hyperlinks, which are essentially a medium-based affordance. For this reason, the classification of the HTML research article as a new genre is essentially due to the medium where it is produced, distributed and consumed.

12. *To what extent does the theoretical approach for analyzing print genres introduced by Swales (1990, 2004) and by Askehave and Swales (2001) apply to the analysis of digital genres?*

Since Askehave and Swales (2001) decided to leave the concept of *context* as 'black box', it is hard to argue that their framework does not consider medium as a context and, therefore, that the framework is flawed. In this sense, depending on how one interprets the model, the theoretical approach for analyzing print genres introduced by Swales (1990, 2004) and by Askehave and Swales (2001) applies to the analysis of digital genres.

Nevertheless, it would be useful if a systematic procedure which dealt with the nature and function of hyperlinks and their impact on genre classification were proposed. In this study, I suggested, following Crowston and Williams

(1999, 2000) and Askehave and Nielsen (20005a, 2005b), that the communicative purpose of digital genres – as far as textual analysis is concerned – can be determined based on the hyperlinking patterns these genres exhibit.

Specifically in the case of the HTML research article, saying that it does not fulfill the same communicative purpose of the print RA does not mean that it does not share scientific knowledge. It does share scientific knowledge just as thesis, conference presentations and books do. These, however, are all different genres. Therefore, my argument is that the HTML research article is not simply a replica of the print RA, but a genre that stands on its own, as an independent text with its own communicative purposes.

So what? I hope that the discussion presented in this thesis will not only contribute for genre analysis, but also for the areas of English for Academic Purposes and digital literacy. It was pointed out earlier in this thesis that throughout my research process I have encountered people in Brazil and in Australia who did not know what and HTML research article was. And there are people who do not know what a hyperlink is, as reported by one of the participants.

MA#11 I don't know what a hyperlink is.

This participant's age is 55 or older. Younger generations may not have the choice of not knowing what hyperlinks are or what HTML research articles are. At this point, we only need to be able to *read* HTML research articles, because it is the publishers and aggregators who create them. How long will it take until

it may be necessary that we *create/write* them ourselves? Will we be prepared? In the *Journals of Visualized Experiments* (JoVE), for example, researchers themselves create the video articles and submit them to the ejournal. Will those who do not know how to make video articles be left outside the scientific mainstream? The answer is *probably*.

When I began my research about online articles in 1999, the HTML research article was a total stranger to me. I know it better now, but it evolves so quickly that if I do not keep track of the changes, it may be a total stranger again in a few years. This fluidity (fast transformation and evolution) of digital genres is another issue that genre analysis has to be prepared to deal with.

References

Agosti, M.; Crestani, F.; Melucci, M. (1997). On the use of information retrieval techniques for the automatic construction of hypertext. *Information Processing & Management, 33*(2), 133-144.

Anthony, L. (1999). Writing research article introductions in software engineering: howaccurate is a standard model? *IEEE Transactions on Professional Communication, 42*(1), 38-46.

Araújo, A. D. (1996). *Lexical signaling: a study of unspecific nouns in book reviews*. Ph.D dissertation, Universidade Federal de Santa Catarina, Florianópolis.

Araújo, A. D. (2006). Práticas discursivas em conclusões de teses de doutorado. *Revista Linguagem em (Dis)curso, 6*(3). Retrieved June 4, 2006, from http://www3.unisul.br/paginas/ensino/pos/linguagem/0603/05.htm.

Araújo, J. C. R. de (2004). A conversa na web: O estudo da transmutação de um gênero textual. In L. A. Marcuschi & A. C. Xavier (Orgs.), *Hipertexto e gêneros digitais* (pp.91-109). Rio de Janeiro: Lucerna.

Araújo, J. C., & Biasi-Rodrigues, B. (2005). A natureza hipertextual do gênero chat aberto. In J. C. Araújo & B. Biasi-Rodrigues (Orgs.), *Interação na internet: Novas formas de usar a lingaugem* (pp. 48-62). Rio de Janeiro: Lucerna.

Araújo, J. C., & Biasi-Rodrigues, B. (2007). Questões de estilo no gênero chat aberto e implicações para o ensino de língua materna. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (pp.78-92). Rio de Janeiro: Lucerna.

Araújo, J. C., & Costa, N. (2007). Momentos Interativos de um chat aberto: A composição do gênero. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (pp.21-34). Rio de Janeiro: Lucerna.

Askehave, I.; Swales, J. M. (2001). Genre identification and communicative purpose: A problem and a possible solution. *Applied Linguistics, 22*(2), 195-212.

Askehave, I.; Nielsen, A. E. (2005a). Digital genres: A challenge to traditional genre theory. *Information Technology & People, 18*(2), 120-141.

Askehave, I.; Nielsen, A. E. (2005b). What are the characteristics of digital genres? Genre theory from a multi-modal perspective. *Proceedings of the 38th Hawaii International Conference on System Sciences.*

Atkinson, D. (1999). Language and science. *Annual Review of Applied Linguistics, 19*, 193-214.

Australian Bureau of Statistics (2006). *8146.0 - Household use of information technology, Australia, 2005-06.* Retrieved July 13, 2007, from http://www.abs.gov.au/Ausstats/abs@.nsf/0/acc2d18cc958bc7bca2568a9001393ae?OpenDocument.

Ayers, G. (2007). The evolutionary nature of genre: An investigation of the short texts accompanying research articles in the scientific journal Nature. *English for Specific Purposes.* Available online 24 August, doi:10.1016/j.esp.2007.06.002.

Bakhtin, M. M. (1986). *Speech genres and other late essays.* Austin: University of Texas Press.

Bar-Ilan, J.; Fink, N. (2005). Preference for electronic format of scientific journals: A case study of the Science Library users at the Hebrew University. *Library & Information Science Research*, *27*, 363-376.

Bazerman, C. (1984). Modern evolution of experimental report in Physics: spectroscopic articles in *Physical Review*, 1893-1980. *Social Studies of Science*, *14*, 163-196.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science.* Madison, WI: University of Wisconsin Press.

Berkenkotter, C. (2006). *North American genre theorists: The North American Genre School (NAGS) – a brief history.* Retrieved August 24, 2007, from http://blog.lib.umn.edu/wardx278/genre/2006/01/north_american_genre_theorists.html.

Berkenkotter, C. (2007). Multimodal genres and cognitive change. Conference Presentation. *4º Simpósio Internacional de Estudos de Gêneros Textuais.* Tubarão, SC.

Berkenkotter, C.; Huckin, T. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power.* Hillsdale: Lawrence Erlbaum Associates.

Bezerra, B. G. (2007). Gêneros introdutórios mediados pela web: O caso da homepage. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (p.113-125). Rio de Janeiro: Lucerna.

Bhatia, V. K. (2004). *Words of written discourse: A genre-based view*. London: Continuum.

Biasi-Rodrigues, B. (1998). *Estratégias de condução de informações em resumos de dissertações*. Ph.D dissertation, Universidade Federal de Santa Catarina, Florianópolis.

Biasi-Rodrigues, B. (2007). O papel do propósito comunicativo na análise de gêneros: Diferentes versões. *Anais do IV SIGET*. Tubarão, SC: UNISUL.

Biesenbach-Lucas, S.; Weasenforth, D. (2001). E-mail and word processing in the ESL classroom: How the medium affects the message. *Language Learning & Technology*, *5*(1), 135-165. Retrieved January 4, 2007, from http://llt.msu.edu/vol5num1/weasenforth/default.html.

Bittencourt, M. (1995). *Academic abstracts: A genre analysis*. MA thesis, Universidade Federal de Santa Catarina, Florianópolis.

Bittencourt, M. (1996). The textual organization of research paper abstracts in Applied Linguistics. *Text*, *16*(4), 481-499.

Bostock, W. W. (2006). The electronic dissemination of research findings. *Academic Exchange Quarterly*. *10*(1), 237-241.

Brett, P. (1994). A genre analysis of the results section of Sociology articles. *English for Specific Purposes*, *13*(1), 47-59.

Breure, L. (2001). *Development of the genre concept*. Retrieved March 2, 2005 from http://www.cs.uu.nl/people/leen/GenreDev/GenreDevelopment.htm.

Butler, D. (1999). The writing is on the Web for science journals in print. *Nature*, *397*(6716), 195-200.

Caballero, R. (2005). The influence of hypertext on genre: Exploring online book reviews. In C. R. Caldas-Coulthard & M. Toolan (Eds.), *The writer's craft, the culture's technology* (p. 57-70). Amsterdam: Rodopi.

Caiado, R. V. R. (2007). A ortografia no gênero weblog: Entre a escrita digital e a escrita escolar. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (p.35-47). Rio de Janeiro: Lucerna.

Caldas-Coulthard, C. R. (2005). Personal web pages and the semiotic construction of academic identities. In C. R. Caldas-Coulthard & M. Toolan (Eds.), *The writer's craft, the culture's technology* (p.23-46). Amsterdam: Rodopi.

CAPES. (2007). *O que é o Portal .periodicos. CAPES?* Retrieved July 13, 2007, from http://www.periodicos.capes.gov.br/portugues/index.jsp.

Chambers, M. B.; S. SooYoung. (2004). Full-Text aggregator database vendors and journal publishers: A study of a complex relationship. *Serials Review*, *30*(3), 183-193.

Chen, C-F. E. (2006). The development of e-mail literacy: From writing to peers to writing to authority figures. *Language Learning & Technology*, *10*(2), 35-55. Retrieved January 4, 2007, from http://llt.msu.edu/vol10num2/chen/.

Coburn, M.; Burrows, P.; Loi, D.; Wilkins, L. (2001). Ebook readers: directions in enabling technology. In B. Cope & D. Kalantzis (Eds), *Print and electronic text convergence* (p. 145-181). Altona, Vic: C-2C Project.

Comitê Gestor da Internet no Brasil (2006). *TIC domicílios: TIC domicílios e usuários 2006.* Retrieved July 13, 2007, from http://www.cetic.br/usuarios/tic/2006/index.htm.

Cross, C.; Oppenheim, C. (2006). A genre analysis of scientific abstracts. *Journal of documentation*, *62*(4), 428-446.

Crowston, K.; Williams, M. (1999). The effects of linking on genres of Web documents. *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*.

Crowston, K.; Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society*, *16*(3), 201-215.

Curran, C. (2002). The medical journal meets the Internet. *First Monday*, *7*(6). Retrieved September 2, 2004, from http://firstmonday.org/issues/issue7_6/curran/index.html.

Delamothe, T.; Smith, R. (1999). Moving beyond journals: The future arrives with a Crash - New ways to disseminate research from NIH and the BMJ. *British Medical Journal*, *318*, 1637-1639.

Delamothe, T.; Smith, R. (2001). PubMed Central: creating an Aladdin's cave of ideas. *British Medical Journal*, 322, 1-2.

Djonov, E. (2005). *The organisation of information in websites: From hypermedia design to systemic functional hypermedia discourse analysis.* PhD dissertation, University of New South Wales, Sydney.

Dubois, B. L. (1980). The use of slides in biomedical speeches. *The ESP Journal*, *1*(1), 45-50.

Dudley-Evans, T.; Henderson, W. (Eds.). (1990). *The language of economics: The analysis of economic discourse.* England: Modern English.

Eckkrammer, E. M. (2004). Drawing on theories of inter-semiotic layering to analyse multimodality in medical self-counselling texts and hypertexts. In E. Ventola, C. Charles & M. Kaltenbacher (Eds.), *Perspectives on multimodality* (p. 211-226). Amsterdam: John Benjamins.

Fontes, M. C. M. (2007). O uso de emoticons en chats: Afetividade em ensino a distância. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (p.64-77). Rio de Janeiro: Lucerna.

Fortanet, I.; Palmer, J. C.; Posteguillo, S. (1998). Netvertising: content-based subgeneric variations in a digital genre. *Proceedings of the 31st Hawaii International Conference on System Sciences.*

Gains, J. (1999). Electronic mail – a new style of communication or just a new medium?: an investigation into text features of e-mail. *English for Specific Purposes, 18*(1), 81-101.

Garfield, E. (2005). *The agony and the ecstasy: The history and meaning of the journal impact factor.* Retrieved October 16, 2006, from http://www.garfield.library.upenn.edu/papers/jifchicago2005.pdf.

Garfield, E. (2006). The history and meaning of the Journal Impact Factor. *Journal of the American Medical Association, 295* (1), 90-93.

Glänzel, W.; Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management, 35,* 31-44.

Gledhill, C. (2000). The discourse function of collocation in research article introductions. *English for Specific Purposes, 19,* 115-135.

Graddol, D. (2006). *English next.* London: British Council.

Groote, S. L. de; Doranski, M. (2204). The use of personal digital assistants in the health sciences: results of a survey. *Journal of the Medical Library Association, 90*(2), 341-348.

Gross, A. G; Harmon, J. E.; Reidy, M. (2000). Argument and 17th-Century Science: A Rhetorical Analysis with Sociological Implications. *Social Studies of Science, 30*(3), 371-396

Gross, A. G.; Harmon, J. E.; Reidy, M. (2002). *Communicating science*: *The scientific article from the 17th century to the present*. Oxford: Oxford University Press.

Halliday, M. A. K. (1978). *Language as social semiotic.* London: Edward Arnold.

Halliday, M. A. K. (1985/1989). Part A. In M. A. K. Halliday & R. Hasan, R., *Language, context, and text: Aspects of language in a social-semiotic perspective* (p. 3-49). Oxford: Oxford University Press.

Halliday, M. A. K. (1991). The notion of "context" in language education. In T. Lê & M. McCausland (Eds), *Language education: Interaction & development. Proceedings of the International Conference, Ho Chi Minh City, Vietnam* (p. 1-26). Launceston: University of Tasmania.

Halliday, M. A. K. (1994). *Introduction to functional grammar.* London: Edward Arnold.

Halliday, M. A. K.; Martin, J. M. (1993). *Writing science: Literacy and discursive power*. University of Pittsburgh Press.

Harnad, S. (1998). On-line journals and financial fire-walls. *Nature, 395,* 127-128.

Hartley, J.; Sydes. M. (1997). Are structured abstracts easier to read than traditional ones? *Journal of Research in Reading*, *20*(2), 122-136.

Hasan, R. (1985/1989). The structure of a text. In M. A. K. Halliday & R. Hasan, R., *Language, context, and text: Aspects of language in a social-semiotic perspective* (p. 52-69). Oxford: Oxford University Press.

Hasan, R. (2004). Analysing discursive variation. In L. Young & C. Harrison (Eds.), *Systemic Functional Linguistics and Critical Discourse Analysis: Studies in social change* (p. 15-52). London/New York: Continuum.

Hemais, B. (2001). The discourse of research and practice in marketing journals. *English for Specific Purposes*, *20*(1), 39-59.

Hemais, B.; Biasi-Rodrigues, B. (2005). A proposta sócio-retórica de John M. Swales para o estudo de gêneros textuais. In J. L Meurer, A. Bonini & D. Motta-Roth (Orgs.), *Gêneros: teorias, métodos, debates* (p. 108-129)*.* São Paulo: Parábola.

Henderson, W.; Dudley-Evans, T.; Backhouse, R. (Eds.). (1993). *Economics and language.* London/New York: Routledge.

Hendges, G. R. (2001). *Novos contextos, novos gêneros: A seção de Revisão da literatura em artigos acadêmicos eletrônicos*. MA thesis, Univeridade Federal de Santa Maria, Santa Maria, RS.

Hendges, G. R. (2002). Citando na Internet: Um estudo de gênero da Revisão da literatura em artigos acadêmicos eletrônicos. In D. Motta-Roth & J. Luiz Meurer. (Orgs.), *Gêneros textuais* (p.117-139). Bauru, SP: EDUSC.

Hendges, G. R. (2006). Gramática da imagem: modalidades semióticas não-verbais em artigos acadêmicos de Lingüística. In D. Motta-Roth, N. C. A. de Barros & M. G. Richter (Orgs.), *Liguagem, cultura e socieldade* (p. 23-38). Santa Maria, RS: Programa de Pós-Graduação em Letras.

Hendges, G. R.; Motta-Roth, D. (2000). Padrões de citação em artigos acadêmicos eletrônicos: When less means more. *Expressão*, *3*(2), 76-83.

Hill, S. S.; Soppelsa, B. F.; West, G. K. (1982). Teaching ESL students to read and write experimental research papers. *TESOL Quarterly*, *16*(3), 333-347.

Hitchcock, S.; Carr, L.; Hall, W. (1996). A survey of STM online journals 1990-95: the calm before the storm. Retrieved September 4, 2001, from http://journals.ecs.soton.ac.uk/survey/survey.html.

Holmes, R. (1997). Genre analysis and the Social Sciences: An investigation of the structure of research article Discussion sections in three disciplines. *English for Specific Purposes*, *16*(4), 321-337.

Hopkins, A.; Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, *7*(2), 113-121.

Houghton, B. (1975). *Scientific periodicals: Their historical development, characteristics and control*. London: Bingley.

Hydén, L-C.; Mishler, E. G. (1999). Language and Medicine. *Annual Review of Applied Linguistics*, *19*, 174-192.

Hyland, K. (2002). Genre: Language, context, and literacy. *Annual Review of Applied Linguistics*, *22*, 113-135.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, *30*, 693-722.

IBGE. (2007). Instituto Brasileiro de Geografia e Estatística. Retrieved online May 7, 2007, from http://www.ibge.gov.br/home/.

Ihlström, C.; Åkesson, M. (2004). Genre characteristics: A front page analysis of 85 Swedish online newspapers. *Proceedings of the 37th Hawaii International Conference on System Sciences.*

Internet World Stats: Usage and population statistics. (2007a). *Internet usage and population in South America*. Retrieved September 20, 2007, from http://www.internetworldstats.com/stats15.htm.

Internet World Stats: Usage and population statistics. (2007b). *Internet usage and population in Oceania.* Retrieved September 20, 2007, from http://www.internetworldstats.com/stats6.htm.

Internet World Stats: Usage and population statistics. (2007c). Top 36 *countries with the highest Internet penetration rate*. Retrieved September 20, 2007, from http://www.internetworldstats.com/top25.htm.

Jewitt, C. (2004). Multimodality and new communication technologies. In P. Levine & R. Scollon (Eds.), *Discourse and technology: Multimodal discourse analysis* (p. 184-195). Washington: Georgetown University Press.

Jeon–Slaughter, H.; Herkovic, A. C.; Keller, M. A. (2005). Economics of scientific and biomedical journals: Where do scholars stand in the debate of online journal pricing and site license ownership between libraries and publishers?
First Monday, 10(3). Retrieved April 15, 2006, from http://firstmonday.org/issues/issue10_3/jeon/index.html.

Johns, A. M; (2004). *Defining and teaching genre*. Retrieved March 4, 2006, from http://www-rohan.sdsu.edu/~annjohns/tesolgenre04.htm.

Jones, S. L.; Cook, C. B. (2000). Electronic journals: Are they a paradigm shift? *Online Journal of Issues in Nursing, 5*(1). Retrieved May 19, 2004, from http://www.nursingworld.org/ojin/topic11/tpc11_1.htm.

Kling, R.; McKim, G. (1997). *A typology for electronic-journals: Characterizing scholarly journals by their distribution forms*. Retrieved April 24, 2006, from http://rkcsi.indiana.edu/archive/CSI/WP/wp97-07B.html.

Knox, J. (2007). Visual–verbal communication on online newspaper home pages. *Visual Communication, 6*(1), 19-53.

Kok, K. C. A. (2004). Multisemiotic mediation in hypertext. In K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives* (p. 129-159). London: Continuum.

Komesu, F. C. (2004). Blogs e as práticas de escrita sobre si na Internet. In L. A. Marcuschi & A. C. Xavier (Orgs.), *Hipertexto e gêneros digitais* (p.110-119). Rio de Janeiro: Lucerna.

Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, *34*(2), 337-340.

Kress, G. (2003). *Literacy in the new media age*. London/New York: Routledge.

Kress, G. (2004a). Media discourse – extensions, mixes, and hybrids: Some comments on pressing issues. *Text*, *24*(3), 443-446.

Kress, G. (2004b). Reading images: Multimodality, representation and new media. *Information Design Journal + Document Design*, *12*(2), 110-119.

Kress, G. (2005). Gains and losses: New forms of texts, knowledge, and learning. *Computers and Composition*, *22*(1)*,* 5-22.

Kress, G.; van Leeuwen, T. (1990). *Reading images*. Geelong: Deakin University Press.

Kress, G.; van Leeuwen, T. (1996). *Reading images*. London: Routledge.

Kress, G.; van Leeuwen, T. (2001). *Multimodal Discourse: the modes and media of contemporary communication*. London: Arnold.

Kwan-Hoong, N. (2006). Electronic publishing. *5th regional workshop on medical writing for radiologists*. Singapore. Retrieved July 2, 2006, from http://www.biij.org/resources/spore/4-06-ElectronicPublishing.mov.

Landow, G. (1997). *Hypertext 2.0: the convergence of contemporary critical theory and technology*. Baltimore/London: John Hopkins.

LaPorte, R. E.; B. Hibbits. (1996). Rights, wrongs, and journals in the age of cyber-
space. *British Medical Journal*, *313*, 1609-1612.

LaPorte, R. E.; Linkov, F.; Villasenor, T.; Sauer, F.; Gamboa, C.; Lovalekar, M.; Shubnikov, E.; Sekikawa, A.; & Eun Ryoung Sa. (2002). Papyrus to PowerPoint (P 2 P): Metamorphosis of scientific communication. *British Medical Journal,* 325, 1478-1481. Retrieved August 9, 2004, from http://bmj.bmjjournals.com/cgi/content/full/325/7378/1478.

LaPorte, R. E.; Marler, E.; Akazawa, S.; Sauer, F.; Gamboa, C.; Shenton, C.; Glosser, C.; Villasenor, A.; & Maclure, M. (1995). The death of biomedical journals. *British Medical Journal, 310*, 1387-1390. Retrieved August 9, 2004, from http://www.bmj.com/archive/6991ed2.htm.

Larivière, V.; Archambault, E.; Gingras Y.; Vignola-Gagné, E. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering

with social sciences and humanities. *Journal of the American Society for Information Science and Technology, 57*(8), 997-1004.

Leal, V. P. L. V. (2007). O chat quando não é chato: O papel da mediação pedagógica em chats educacionais. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (p.48-63). Rio de Janeiro: Lucerna.

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science* (p. 87-113). London: Routledge.

Lemke, J. L. (1999). Discourse and organizational dynamics: Website communication and institutional change. *Discourse & Society, 10*(1), 21-47.

Lemke, J. L. (2002). Travels in hypermodality. *Visual Communication, 1*(3): 299-325.

Llewellyn, R. D.; Pellack, L. J.; Shonrock, D. D. (2002). The use of electronic-only journals in scientific research. *Issues in Science and Technology Librarianship, 35*. Retrieved June 23, 2005, from   http://www.istl.org/02-summer/refereed.html.

Luzon Marco, M. J. (2000). Collocational frameworks in medical research papers: a genre-based study. *English for Specific Purposes, 19*, 63-86.

MacIntosh-Murray, A. (2007). Poster presentations as a genre in knowledge communication: A case study of forms, norms, and values. *Science Communication, 28*(3), 347-376.

Martin, J. R. (1984). Language, register and genre. Reprinted in A. Burns & C. Coffin (Eds.) (2001), *Analysing English in a global context* (p. 150-166). London: Routledge.

Martin, J. R. (1992). *English text: System and structure.* Philadelphia/Amsterdam: John Benjamins.

McLuhan, M. (1964). *Understanding media: The extensions of man.* New York: McGraw-Hill.

Meurer, J. L. (1998). Aspects of language in self-help counselling. Pós-Graduação em Inglês: Florianópolis.

Meurer, J. L. (2003). Genre as diversity, and rhetorical mode as unity in language use. *Ilha do Desterro, 43*(2), 61-82.

Meurer, J. L. (2004) Role prescriptions, social practices, and social structures: A sociological basis for the contextualization of analysis in SFL and CDA. In L. Young & C. Harrison. (Eds.), *Systemic Functional Linguistics and Critical*

Discourse Analysis: Studies in social change (p. 85-99). London/New York: Continuum.

Meurer, J. L. (2006). Integrando estudos de gêneros textuais ao contexto de cultura. In A. M. Karwoski, B. Gaydecka, K. S. Brito. (Orgs.), *Gêneros textuais: Reflexões e ensino* (p. 165-185). 2 ed. Rio de Janeiro: Lucerna.

Miller, C. (1984). Genre as social action. *Quarterly Journal of Speech*, *70*, 151-167.

Miller, T. (1998). Persuasion: A comparison of visuals in academic texts and the popular press. *English for Specific Purposes*, *17*(1), 29-46.

Miller, C.; Shepherd, D. (2004). Blogging as social action: A genre analysis of the weblog. In L. J. Gurak, S. Antonijevic, L. Johnson, C. Ratliff & J. Reyman (Eds.), *Into the blogosphere: Rhetoric, community, and culture of weblogs.* Retrieved May 23, 2007, from http://blog.lib.umn.edu/blogosphere/blogging_as_social_action.html.

Misser, E. (n/d). *The research paper: A writing centre handout.* Retrieved June 25, 2006, from http://www.wlu.ca/forms/926/The_Research_Paper.pdf.

Mitchell, W. J. T. (1995). *Picture theory: Essays on verbal and visual representation*. Chicago: University of Chicago Press.

Montgomery, S. (2004). Of towers, walls, and fields: Perspectives on language in science. *Science*, *303*(5662), 1333-1335.

Motta-Roth, D. (1995). *Rhetorical features and disciplinary cultures: A genre-based study of academic book reviews in linguistics, chemistry, and economics.* Ph.D dissertation, Universidade Federal de Santa Catarina, Florianópolis.

Motta-Roth, D. (1998). Discourse analysis and academic book reviews: A study of text and disciplinary cultures. In I. Fortanet, S. Posteguillo, J. C. Palmer & J. F. Coll. (Eds.), *Genre studies in English for Academic Purposes* (p. 29-48). Vol. 1. Castelló de la Plana: Publicacions de la Universitat Jaume.

Motta-Roth, D. (2006a). Questões de metodologia em anaálise de gênero. In A. M. Karwoski, B. Gaydecka, K. S. Brito. (Orgs.), *Gêneros textuais: Reflexões e ensino* (p. 145-163). 2 ed. Rio de Janeiro: Lucerna.

Motta-Roth, D. (2006b). Ensino de produção textual com base em atividades sociais e gêneros textuais. *Revista Linguagem em (Dis)curso, 6*(3). Retrieved April 4, 2007, from http://www3.unisul.br/paginas/ensino/pos/linguagem/0603/07.htm.

Motta-Roth, D. (2007). Contribuições da análise crítica de gêneros discursivos para o estudo e o ensino da linguagem. Conference Presentation. *4º Simpósio Internacional de Estudos de Gêneros Textuais.* Tubarão, SC.

Motta-Roth, D.; Hendges, G. R. (1996). Uma análise de gênero de resumos acadêmicos (abstracts) em Econimia, Lingüística e Química. *Revista do Centro se Artes e Letras*, *18*(1-2), 53-90.

Motta-Roth, D.; Hendges, G. R. (1998). Uma análise transdisciplinar do gênero abstract. *Intercâmbio, 7*, 117-125.

Motta-Roth, D.; Hendges, G. R. (1999). A genre analysis study in computer science. *Anais do XIV Encontro Nacional de Professores Universitários de Língua Inglesa (ENPULI*), 103 – 109.

Motta-Roth, D.; Hendges, G. R.; Amoretti, E. C.; Kloss, M. (2000). O tradicional e o novo: Análise de artigos acadêmicos eletrônicos. *Intercâmbio, 9*, 29-38.

Motta-Roth, D.; Marshall, D.; Reis, S. C. dos. (2005). Aprender inglês para a comunicação: A construção da home page pessoal na www. *Calidoscópio*, 3(1), 39-46.

Motta-Roth, D.; Reis, S. C. Dos; Marshall, D. (2007). O gênero página pessoal e o ensino de produção textual em inglês. In J. C. Araújo (Org.), *Internet e ensino: novos gêneros, outros desafios* (p.126-143). Rio de Janeiro: Lucerna.

Motta-Roth, D.; Siqueira, F. O.; Kurtz, F. D.; Marcuzzo, P.; Reis, S. C. dos. (2004). Uma análise de periódicos acadêmicos eletrônicos brasileiros nas áreas de Ciências Humanas, Biológicas e Agrárias. *Revista de Letras, 1/2*, 103-111.

Myers, Greg. (1990). *Writing Biology: Texts in the Social Construction of Scientific Knowledge.* Madison, WI: University of Wisconsin Press.

Nascimento, R. G. do. (2002). *A interface texto verbal e texto não-verbal no artigo acadêmico de engenharia elétrica.* Unpublished MA Thesis, Universidade Federal de Santa Maria, Santa Maria.

Nascimento, R. G.; Motta-Roth, D.; Hendges, G. R. (2003). Contributions of genre analysis research to English for Academic Purposes teaching. *Anais do XVI ENPULI*, 383-390.

Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes, 16*(2), 119-138.

Odlyzko, A. M. (1994). Tragic loss or good riddance? The impending demise of traditional scholarly journals. *Journal of Universal Computer Science, 0*(0), 3-53.

Retrieved October 8, 2003, from http://www.jucs.org/jucs_0_0/tragic_loss_or_good/Odlyzko_A.html.

Ozturk, I. (2007). The textual organisation of research article introductions in Applied Linguistics: Variability within a single discipline. *English for Specific Purposes*, *26*(1), 25-38.

O'Toole, L. M. (1994). *The language of displayed art*. London: Leicester University Press.

Paiva, V. L. M. O. (2004). E-mail: Um novo gênero textual. In L. A. Marcuschi A. C. Xavier (Orgs.), *Hipertexto e gêneros digitais* (p. 68-90). Rio de Janeiro: Lucerna.

Paltridge, B. (2001). *Genre and the language learning classroom*. Ann Arbor: University of Michigan Press.

Paltridge, B. (2002). Thesis and dissertation writing: An examination of published advice and actual practice**.** *English for Specific Purposes*, *21*(2), 125-143.

Peek R. P.; Newby G. B. (Eds.) (1996). *Scholarly publishing: The electronic frontier*. Cambridge: MIT Press.

Peters, P.; Smith, A. (2003). Textual structure and segmentation in online documents. *Language and Computers*, *48*(1), 71-88.

Posteguillo, S. (1999). The schematic structure of computer science research articles. *English for Specific Purposes*, *18*(2), 139-160.

Pullinger, D.; Baldwin, C. (1997). *SuperJournal: What readers really want from electronic journals.* Retrieved October 8, 2003, from http://www.superjournal.ac.uk/sj/elvira.htm.

Reis, M. M. de O. (2005). *Acesso e uso do Portal de Periódicos CAPES pelos professores da Universidade Federal do Acre*. MA thesis, Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópilos.

Roth, W. M.; Bowen, G. M.; Roth, W. M.; McGinn. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching, 36*(9), 977-1019.

Roth, W. M.; Welzel, M. (2001). From activity to gestures and scientific language. *Journal of Research in Science Teaching*, *38*(1), 103-136.

Rowland, F. (1997). Print journals: fit for future?. *Ariadne*, *7*. Retrieved October 8, 2003, from http://www.ariadne.ac.uk/issue7/fytton/.

Rowley-Jolivet, E. (2002). Visual discourse in scientific conference papers: a genre-based study. *English for Specific Purposes*, *21*(1), 19-40.

Rowley-Jolivet, E. (2004). Different visions, different visuals: A social semiotic analysis of field-specific visual composition in scientific conference presentations. *Visual communication*, *3*(2), 145-175.

Rowley-Jolivet, E.; Carter-Thomas, S. (2005). The rhetoric of conference presentation introductions: context, argument and interaction. *International Journal of Applied Linguistics*, *15*(1), 45-70.

Samraj, B. (2002). Introductions in research articles: Variations across disciplines. *English for Specific Purposes*, *21*(1), 1-17.

Salager-Meyer, F. (1990). Metaphors in medical English prose: A comparative study with French and Spanish. *English for Specific Purposes*, *9*(2), 145-159.

Salager-Meyer, F. (1991). Medical English abstracts: How well are they structured? *Journal of the American Society for Information Science. 42*(7), 528-531.

Salager-Meyer, F. (1992). A text-type and move analysis study of verb tense and modality distribution in medical English abstracts. *English for Specific Purposes*, *11*(2), 93-113.

Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*, *13*(2), 149-170.

Salager-Meyer, F. (1999). Referential behavior in scientific writing: a diachronic study (1810-1995). *English for Specific Purposes*, *18*(3), 279-305.

Salager-Meyer, F. (2000). Debate-Creating vs. Accounting References in French Medical Journals. *Technical Communication Quarterly*, 9(3), 291-310.

Salager-Meyer, F.; Alcaraz Ariza, M. A. (2003). Academic criticism in spanish medical discourse: A cross-generic approach. *International Journal of Applied Linguistics*, *13*(1), 96-114.

Salager-Meyer, F.; Alcaraz Ariza, M. A.; Zambrano, N. (2003). The scimitar, the dagger and the glove: Intercultural differences in the rhetoric of criticism in Spanish, French and English medical discourse (1930–1995). *English for Specific Purposes*, *22*, 223-247.

Sathe, N. A.; Grady, J. L.; Giuse, N. B. (2002). Print versus electronic journals: A preliminary investigation into the effect of journal format on research processes. *Journal of the Medical Library Association*, *90*(2), 235-243.

Shepherd, M.; Watters, C. (1998). The evolution of cybergenres. *Proceedings of the 31st Annual Hawaii International Conference on System Sciences.*

Shepherd, M.; Watters, C. (1999). The functionality attribute of cybergenres. *Proceedings of the 32nd Hawaii International Conference on System Sciences.*

Silva, L. F. 1999. *Análise de gênero: uma investigação da seção de Resultados e Discussão em artigos científicos em Química.* MA Thesis, Universidade Federal de Santa Maria, Santa Maria, RS.

Sony. (2007). *Sony® Reader.* Retrieved July 2, 2007, from http://products.sel.sony.com/pa/prs/reader_features.html.

Spooner, M.; Yancey, K. (1996). Postings on a genre of email. *College Composition and Communication*, *47*(2), 252-278.

Starfield, S.; Ravelli, L. J. (2006). "The writing of this thesis was a process that I could not explore with the positivistic detachment of the classical sociologist": Self and structure in New Humanities research theses. *Journal of English for Academic Purposes*, *5*(3), 222–243

Swales, J. M. (1986). Citation analysis and discourse analysis. *Applied Linguistics, 17*(1), 39-56.

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

Swales, J. M. (1996). Teaching the conference abstract. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues* (p. 45–59). John Benjamins.

Swales, J. M. (2004). *Research genres: Explorations and applications.* New York: Cambridge University Press.

Swales, J. M.; Feak, C. B. (1994/2004). *Academic writing for graduate students: Essential tasks and skills.* Ann Arbor: The University of Michigan Press.

Swales, J. M.; Feak, C. B. (2000). Research on display – the conference poster. In *English in today's research world: A writing guide* (p. 80-113). Ann Arbor, MI: The University of Michigan Press.

Swales, J. M.; Ahmad, U. K.; Chang, Y-Y.; Chavez, D.; Dressen, D. F.; Seymour, R. (1998). Consider this: The role of imperatives in scholarly writing. *Applied Linguistics*, *19*(1), 97-121.

Tardy, C. M. (2005). Expressions of disciplinarity and individuality in a multimodal genre. *Computers and Composition, 22* (3), 319-336.

Tenopir, C. (2003). Use and users of electronic library resources: An overview and analysis of recent research studies. Washington, D.C.: Council on Library and Information Resources.

Tenopir, C.; King, D. W. (1998). Designing electronic journals with 30 years of lessons from print. *The Journal of Electronic Publishing*, *4*(2). Retrieved February 23, 2006, from http://www.press.umich.edu/jep/04-02/king.html.

Tenopir, C.; King, D. W. (2000). *Towards electronic journals: Realities for scientists, librarians, and publishers.* Washington, DC: Special Libraries Association.

Tenopir, C.; King, D. W. (2001). Lessons from the future of journals. *Nature*, *413*, 672-674.

Tenopir, C.; King, D. W. (2002). Reading behaviour and electronic journals. *Learned Publishing*, *15*, 259-265

Tenopir, C.; King, D. W.; Boyce, P.; Grayson, M.; Zhang, Y.; Ebuen, M. (2003). Patterns of journal use by scientists through three evolutionary phases. *D-Lib Magazine*, *9*(5). Retrieved May 23, 2005, from http://www.dlib.org/dlib/may03/king/05king.html.

Todorov, T. (1976). The origin of genres. *New Literacy History*, *6*, 159-170.

The Free On-line Dictionary of Computing. (n.d.). *Portable Document Format*. Retrieved August 07, 2007, from http://dictionary.reference.com/browse/Portable Document Format.

The Royal Society. (2006). Retrieved November 8, 2005, from http://www.pubs.royalsoc.ac.uk/index.cfm?page=1244.

Thetela, P. (1997). Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes*, *16*(2), 101-118.

Thomas, S.; Hawes, T. (1994). Reporting verbs in medical journal articles. *English for Specific Purposes*, *13*(2), 129-148.

Thompson, G.; Yiyun, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, *12*(4), 365-382.

Thornes, J. E. (2004). The visual turn and Geography (Response to Rose 2003 Intervention). *Antipode*, *36*(5), 787-794.

Valauskas, E. J. (1997). Waiting for Thomas Kuhn: First Monday and the evolution of electronic journals. *First Monday, 2*(12). Retrieved February 22, 2006, from http://firstmonday.org/issues/issue2_12/valauskas/index.html.

Varian, H. R. (1998). The future of electronic journals. *The Journal of Electronic Publishing, 4*(1). Retrieved October 8, 2003, from http://www.press.umich.edu/jep/04-01/varian.html.

Varttala, T. (1999). Remarks on the communicative functions of hedging in popular scientific and specialist research articles on Medicine. *English for Specific Purposes, 18* (2), 177-200.

Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes, 20*(1), 83-102.

Vaughan, M. W.; Dillon, A. (2006). Why structure and genre matter for users of digital information: A longitudinal experiment with readers of a web-based newspaper. *International Journal of Human-Computer Studies, 64*, 502–526.

Ventola, E. (1995). Generic and register qualities of texts and their realization. In P. H. Fries & R. Hasan (Eds.), *Discourse in society: Systemic functional perspectives* (p. 3-29). Norwood, NJ: Ablex.

Xavier, A. C.,; Santos, C. F. (2005). E-forum na Internet: Um gênero digital. In J. C. Araújo & B. Biasi-Rodrigues (Orgs.), *Interação na internet: Novas formas de usar a lingaugem* (p. 30-38). Rio de Janeiro: Lucerna.

Wang, S-C. (2006). Preparing effective medical illustrations for publication. *5th regional workshop on medical writing for radiologists*. Singapore. Retrieved July 2, 2006, from http://www.biij.org/resources/spore/2-04-EffectiveImaging.mov.

Williams, I. A. (1996). A contextual study of lexical verbs in two types of medical research report: Clinical and experimental. *English for Specific Purpose, 15*(3), 175-197.

Williams, I. A. (1999). Results sections of medical research articles: Analysis of rhetorical categories for pedagogical purpose. *English for Specific Purposes, 18*(4), 347-366.

Yang, R.; Allison, D. (2003). Research articles in Applied Linguistics: Moving from results to conclusions. *English for Specific Purposes, 22*(4), 365-385.

Yang, R.; Allison, D. (2004). Research articles in Applied Linguistics: Structures from a functional perspective. *English for Specific Purposes, 23*(3), 264-279.

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas