universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Exploring patient information sheets – Analysing a genre for ESP learning and teaching

Verfasserin

Ulrike Gruböck

Angestrebter akademischer Grad

Magistra der Philosophie (Mag. phil.)

Wien, im August 2008

Studienkennzahl lt. Studienblatt:	A 190 344 299
Studienrichtung lt. Studienblatt:	Lehramtsstudium UF Englisch
Betreuerin:	Ao. Univ.-Prof. Dr. Christiane Dalton-Puffer

Acknowledgements

First and foremost I would like to thank my advisor Univ.-Prof. Dr. Christiane Dalton-Puffer for her continuous support and encouragement. Her advice was of invaluable help to me. Without her guidance it would have been impossible to complete this thesis. I am also greatly indebted to my teachers Dr. Ute Smit and Mag. Amy Krois-Lindner for encouraging me to choose a topic from the field of ESP. Thanks also to Dr. Julia Isobel Hüttner for her support as regards computer-based genre analysis.

Furthermore, I would like to thank Martha Kostera and Verena Tischberger, my specialist informants, for providing impetuses and feedback. Their specialist opinion and advice made an enormous contribution to the present thesis. Special thanks, of course, also go to my family, friends and colleagues.

Last, but not least, I am grateful to the following for the permission to use copyright material for educational or scholarly purposes: the National Health Service (UK), Health Canada, the American Association of Family Physicians, the Government of Victoria (The Better Health Channel) and the Cancer Society of New Zealand.

Declaration of Authenticity

I confirm to have conceived and written this paper in English all by myself. Quotations from other authors and any ideas borrowed and/or passages paraphrased from the works of other authors are all clearly marked within the text and acknowledged in the bibliographical references.

Vienna, August 2008 _____

Table of Contents

Abstract (in English)

Genre analysts, course book designers and English language teachers have ignored for a long time the enormous potential of written medical genres and the genre of patient information sheets (PIS) in particular as sources for course material for EMP (English for Medical Purposes) courses. To remedy this problem this paper explores the genre of patient information sheets (PIS) in its contextual setting. For this purpose, a corpus-based genre analysis is undertaken. It is shown that both genre analysis and corpus studies (despite ongoing disputes in the field of applied linguistics) are popular and highly useful in the fields of ELT and especially ESP. Among the various approaches to genre analysis is the ESP approach. ESP served as a basis for the analysis performed in this study. Using patient information sheets as an example, a computer-assisted genre analysis is proposed to analyse vocabulary, whereas a manual analysis is used to identify the basic move structure and the realisation of the communicative purpose. For the purpose of this study, 20 patient information sheets (PIS) on the topics of breast cancer, prostate cancer, skin cancer and bowel cancer have been compiled and analysed.

The results of the thesis show that various types of vocabulary can be found in patient information sheets (PIS). These include academic and technical as well as general vocabulary. Although general vocabulary is the commonest type found in patient information sheets, the quantity of academic and technical vocabulary is striking. In fact, the findings of the analysis suggest that patient information sheets (PIS) are more academic and technical than academic texts. Moreover, it is shown that multi-word units such as compounds, as well as lexical phrases, play a central role in patient information sheets (PIS).

As far as the basic build-up of the genre texts is concerned, the findings of this study suggest that there is a structure that is used by many authors of patient information sheets (PIS). There are various moves that are particular to the genre.

Authors of patient information sheets (PIS) tend to include information about the disease, including the number of people affected, warning signs or symptoms, diagnostic procedures, treatment options, prevention, and references to other sources of information.

Abstract (in German)

Sprachwissenschaftler/Innen und Englischlehrende haben das Potential von medizinischen Genres, vor allem das der Patientenaufklärungsbögen, als Quellen für Unterrichtsmaterial für den Englischunterricht für medizinisches Personal lange Zeit nicht erkannt. Ziel der vorliegenden Diplomarbeit ist es jene linguistische Eigenschaften von Patientenaufklärungsbögen hervorzuheben, die für einen Englischunterricht interessant und relevant wären. Als Methode wurde die (auf einen Korpus basierende) Genreanalyse gewählt. Zu den Kriterien, die während der Zusammenstellung des Korpus zum Tragen kamen, zählten neben der Aktualität der Texte und vertrauenswürdige Quellen auch die Themenwahl. So wurden ausschließlich Patientenaufklärungsbögen zu den Themen Brustkrebs, Prostatakrebs, Hautkrebs und Darmkrebs gewählt. Der Korpus welcher auf 20 verschiedenen Bögen basiert wurde mit Hilfe eines Computerprogramms hinsichtlich Wortschatzarten analysiert. Danach folgte eine Analyse der Makrostruktur der Bögen. Wichtig war es zu zeigen, welche typischen ‚Bausteine' Patientenaufklärungsbögen auszeichnen, welche kommunikativen Zwecke diese haben und wie diese in der Praxis realisiert werden. Dabei wurden den so genannte *language functions* wie Definieren und Beschreiben große Aufmerksamkeit geschenkt. Die Resultate der Analyse der Bögen zeigen, dass Vokabular der Allgemeinsprache, wie auch der Akademischen und Technischen Sprache in den Genretexten zu finden sind. Obwohl das Vokabular der Allgemeinsprache dominiert, sind das Vokabular der Akademischen und das der Technischen Sprachen nicht außer Acht zu lassen. Die Ergebnisse der Analyse zeigen, dass Patientenaufklärungsbögen akademischer und technischer sind als akademische Texte. Im Rahmen der Vokabelanalyse wurden auch Mehrwortlexemen Aufmerksamkeit geschenkt. Die Ergebnisse zeigen, dass diese überaus häufig und wichtig sind. Ein Kernbereich der Arbeit befasst sich mit der so genannten *move structure*, der prototypischen Struktur von Patientenaufklärungsbögen. Die Resultate der Analyse zeigen, dass viele Informationsbögen eine ähnliche Struktur aufweisen. So wurden einige Elemente identifiziert, die als genre-spezifisch beschrieben werden können. Im Großen und Ganzen tendieren Autoren der Bögen die jeweilige Krankheit, Symptome dieser, wie auch Arten der Diagnose und der Behandlung zu thematisieren. Informationen bezüglich Krankheitsprävention und weiterführender Literatur werden meistens ebenfalls angeboten.

Key to abbreviations

AWL	Academic Word List
BNC	British National Corpus
EAP	English for Academic Purposes
EMP	English for Medical Purposes
EOP	English for Occupational Purposes
ESP	English for Specific Purposes
EST	English for Science and Technology
GSL	General Service List of English
IMG	International medical graduates
LSP	Language for Special/ Specific Purposes
MWU	Multi-word unit
PIS	Patient information sheets
VET	Vocational Education and Training

List of figures

1 Introduction

Specialised language teaching and learning has awakened academic and professional interest over the last few years. Evidence for this development includes qualifications such as the BEC (Business English Certificate), the ILEC (International Legal English Certificate) and the ICFE (International Certificate in Financial English). These qualifications are symptomatic of the increasing demand for specialised language learning and teaching. Specialised language skills are increasingly gaining prominence as more businesses are encouraging their employees to attend courses and take accredited exams. This growing interest in specialised language can be observed in the medical profession as well, although there are several reasons for the rising demand. Firstly, both aspiring and practicing doctors and nurses have to communicate with patients from different cultures. Although interpreters are often consulted to assist medical staff in informing patients, they are not always immediately accessible. Foreign language skills are therefore essential. Secondly, most scientific research is conducted and published in English. As medical articles and journals are mainly written in English, the language plays a significant role in medical education at the tertiary and higher levels. When considering a career in a foreign country, it is extremely useful to be familiar with the English language, as well as the genres and conventions of the relevant discipline. The branch of teaching concerned with medical purposes is generally known as English for Medical Purposes (EMP). It belongs to the field of English for Specific Purposes (ESP), the main function of which is:

> to help language learners cope with the features of language or to develop the competencies needed to function in a discipline, profession, or workplace. (Basturkmen 2006: 6)

In the case of EMP, the function and/or objective is also the same. The main goal is to equip aspiring and practicing medical personnel with the skills that enable them to function within the discipline of medicine. These skills include language and communication proficiency, as well as knowledge regarding common practices and genres.

The interest in EMP is slowly but gradually growing, but further research is needed – especially in regard to medical genres and their roles in the discipline. The focus of this paper is therefore on the genre of patient information sheets (PIS). It is attempted to demonstrate that the genre is well worth considering when compiling material and designing courses for medical personnel. This study aims to determine the key features of patient information sheets (PIS) that might be of relevance for EMP teaching and learning (in terms of vocabulary and move structure). For this purpose a corpus-based genre study is undertaken, the aim of which is to investigate the relation and coverage of general, academic and technical vocabulary, but in addition will examine in closer detail the use of multi-word units in patient information sheets (PIS). Another goal of this study is to determine the so-called move structure of the genre, hereby the communicative purposes of the genre and their realisations will also be discussed.

Regarding the organisation of this paper, the first part prepares the groundwork for the study. Chapter Two provides general information on genre theory, the development of written discourse analysis, the different approaches to genre and present developments in this field; Chapter Three discusses the convergence of genre- and corpus-based methodologies in the field of language pedagogy and especially in the field of ESP. This chapter is of special interest since the method in the study is a synthesis of genre and corpus analysis. More information on the setting of the study, including the genre of patient information sheets (PIS), previous studies in this and related fields, the institutional setting, the corpus design of the corpus (as well as the methods applied) are offered in Chapter Four. In Chapters Five and Six the results of the analyses are presented. The patient information sheets (PIS) are analysed according to vocabulary and the generic structure. The results of each analysis are dealt with in a separate chapter. The final chapter explores how genres and genre-based data can be used and exploited in the language classroom.

2 Genre theory

In the last few years genre theory has shed light on the way that discourse is used in academic, professional or other institutional settings (Bhatia 2004: XIV). This genre-based view of discourse has become popular, triggering "a wave of studies" (Hyon 1996: 693). Interest in the genre-based view has since continued. As the name implies, the underlying concept of the theory is based on genre. Genres are commonly referred to as:

> recognizable communicative events, characterised by a set of communicative purposes identified and mutually understood by members of the professional or academic community in which they regularly occur. (Bhatia 2004: 23)

Another characteristic of genres is that they are typically structured in nature and subject to conventions. However, experts of a specific community can exploit and manipulate these conventions for their own 'private' purposes. Given that they can be manipulated according to intentions, one has to acknowledge that genres are constantly changing – and in this sense they are dynamic rather than static. Even though genres are dynamic, they tend to mirror the culture and the (changing) practices of a community (Bhatia 2004: 23-24). According to what has been described so far, genre analysis can be described as "the study of situated linguistic behaviour in institutionalised academic or professional settings" (Bhatia 2004: 22).

This paper will deal with genre theory and genre analysis, and its application in an ESP setting. In order to set the stage, it is necessary to discuss the history of written discourse analysis. The focus here is on the development of written discourse analysis and in particular that of genre analysis. In addition, this chapter will provide an overview of three approaches to genre and genre analysis: the systemic functional approach, the new rhetoric approach and the ESP approach. In particular, there will be a closer examination of the ESP approach to genre and genre analysis, since it serves as a framework for my own study. Finally, the last section of the chapter will explore a new development in genre theory, one towards a multi-perspective view of genre.

2.1 Historical development of written discourse analysis

This paper will explore a medical genre, and also use the framework of analysing texts as genres. The approach of genre analysis was chosen as it "has become an important approach to text analysis, especially in the field of English for Specific Purposes (ESP)" (Dudley-Evans 1994: 219). In order to write about genre analysis it is necessary to consider its history and origins. Genre theory should be regarded as a result of a gradual development in written discourse analysis (Bhatia 2004: XIV-XV), thus it is essential to take into consideration the historical development. It is possible to divide the historical development of written discourse analysis into three distinct, yet overlapping phases. These are "(a) textualisation of lexico-grammar, (b) organisation of discourse and (c) contextualisation of discourse" (Bhatia 2004: 4), each to be discussed in turn:

Textualisation of lexico-grammar

The first phase of written discourse analysis dates back to the 1960s and 1970s. During this period, written discourse was investigated from a formal linguistic perspective, hence language description primarily involved the surface level of a language (Bhatia 2004: 4). The main focus was on recurring features in actual language use. One of the first studies exploring frequently occurring language patterns was the work of Barber (1962) on research articles (Bhatia 2004: 4, Gavioli 2005: 55). Among the aims of the study were to identify recurring lexico-grammatical features and common modal verbs that were particular to scientific prose (Barber 1962, Gavioli 2005: 55-56) [1]. For this purpose, a genre-based corpus study was carried out (ibid.). In the literature there is common agreement on the importance of Barber's (1962) contribution to the description of language. It is still widely seen as influential, especially in regard to English for Specific Purposes (ESP) (Swales 1988:1-2, Gavioli 2005: 55). The interest in specialised fields and texts, as well as the focus on the surface level description (and thus on the description of recurring lexical items and grammatical structures) served the purpose of ESP studies well (Bhatia 2004: 5).

[1] For more information see section 3.3.1, Barber (1996) or Swales (1988).

This phase of written discourse analysis was also linked with other developments. Firstly, there was a growing interest in the functional or discoursal value of lexico-grammatical features, thus the focus was on the textualisation of those features. According to Bhatia (2004: 5), Swales' (1974) work on the function of –en particles in chemistry texts is doubtless one of the most representative and striking examples of this development. Swales (1974) investigation of –en particles in scientific discourse demonstrates that these particles occupy a central role in this type of discourse, as they fulfil specific rhetorical functions (ibid.).

Secondly, another development in this phase involved the investigation of discoursal organisation. Cohesion and coherence were of particular interest as well as the macro-structures of discourse. To conclude, it can be said that text-linguistics deviated from descriptive linguistics insofar as it focused on particular samples of language use rather than on the whole language (Bhatia 2004: 7-8).

Organisation of discourse

Slowly but gradually there was a shift in written discourse analysis, away from textualisation to discourse. Unlike the first phase of written discourse analysis, the second phase was characterised by a focus on discourse and the identification of discourse patterns. This phase itself can be subdivided into three different developments (Bhatia 2004: XIV-XV): the first of which was concerned primarily with the structure of mainly specialised discourse, while the second development investigated more general patterns of organisation; in contrast, the third development focused on the macro-structure and the correlating communicative purpose(s) of academic or professional genres (e.g. Swales 1990, Bhatia 1993)[2]. The notion of communicative purpose was and is especially important for genre theory and genre analysis, since it "governs choice at the grammatical and lexical level" and also highlights the communicative purpose of an author and thus of a text, thus demonstrating that "in fact, the defining feature by which a genre […] can be distinguished from other genres […]" (both Dudley-Evans 1994: 219). In general terms, this approach of analysing texts as genres opened possibilities to investigate how and why members of an academic or professional community produce genres the way they do. In contrast to text-linguistics and the focus on lexico-grammatical features, genre theory allowed "a thicker description of language" (Bhatia 1993: 39-

[2] For more information see Bhatia (2004: 9).

40), taking into account socio-cultural as well as psychological factors (Bhatia 1993: 16). Genre theory as a means of written discourse analysis emerged in the 1990s and soon became a fashionable field of research (Bhatia 2004: 9-10). Studies performed in the 1990s mainly concerned academic (Swales 1990) and professional genres, in particular legal and business genres (Bhatia 1993). Even though the interest in genre was constantly growing, different frameworks for genre analysis were established. Among the various approaches to genre analysis are the systemic functional approach to genre, which centres on the works of Martin (1993); the new rhetoric approach as represented by Miller (1984), Bazermann (1994) and Berkenkotter and Huckin (1995); and the ESP approach to genre analysis, described in detail by Swales (1990) and Bhatia (1993) (Bhatia 2004, Flowerdew L. 2005, Hyon 1996, Kay & Dudley-Evans 1998). The growing interest in genre knowledge went hand in hand with an awakened interest in the social contexts and hybridisation of genres (Bhatia 2004: 10-11).

Contextualisation of discourse

So far, it has been demonstrated that the focus of written discourse analysis shifted from lexico-grammatical features of texts to the patterns of discourse organisation. In recent years, the focus has been on the context of discourse (Bhatia 2004: 11). Whereas some researchers continued with the tradition of genre analysis and expanded the approach, in the sense that they raised issues such as the hybridisation, dynamism and constellations of genres (Bazermann 1994, Berkenkotter and Hucking 1995), others moved towards a multi-perspective view towards genre (Bhatia 2004)[3]. Other developments focused on the language as critical discourse or on language as social interaction[4] (Bhatia 2004: 11).

[3] For more details see section 2.5.

[4] For more information on Critical discourse analysis see Fairclough (1995); for more information on discourse as social interaction see Scollon (1998).

In this section of this paper I have already mentioned the developments of written discourse analysis and thus also of genre theory. I will now turn to approaches to genre and genre analysis, since these are of relevance for my thesis. In the following sections, three different approaches to genre are introduced: the systemic functional approach, the new rhetoric approach and the ESP approach. The basic principles of these approaches (such as the genre views adopted in each case, as well as overall goals in regard to language pedagogy) will also be outlined in greater detail. I will begin by introducing the systemic functional approach.

2.2 The systemic functional approach to genre

As mentioned in the previous section, the systemic functional approach is just one approach to genre. It is also referred to as the Australian systemic functional approach or simply the Australian School or Sidney School (Bhatia 2004: 10, Hyon 1996: 696). It is, as the name implies, based on systemic functional linguistics, which can be traced back to M.A.K. Halliday (1978) and his book *Language as Social Semiotic* (Halliday 1978). In systemic functional linguistics it is of interest to see how language functions, or is used when people communicate (Halliday & Hasan 1989: 15). Language is seen as the realisation of social context. Social context determines and influences language, and in turn is determined and influenced by language (Halliday & Martin 1993: 24-25). This is illustrated by the figure below:

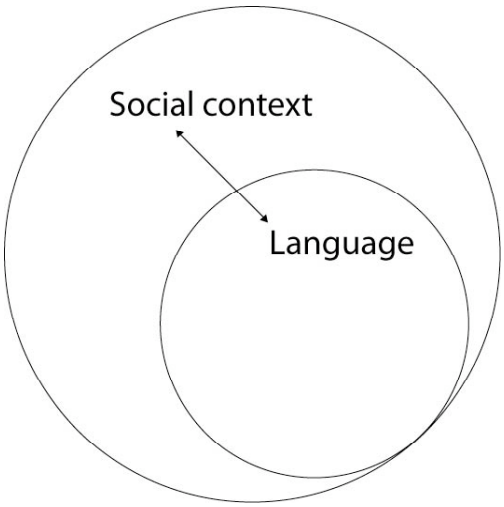


Figure 1: The relationship between language and social context (Halliday & Martin 1993: 25)

The genre view adopted by systemic functional linguists can be traced back to the definitions outlined by Martin (et al. 1987), a student of Halliday. In his model of genre (1992), he "takes language as departing point" as "linguistic choices are the first to be contextualised" (Halliday & Martin 1993: 32) and then moves towards social context, which according to him consists of the level of genre and the level of register (Martin 1992: 495). It has been mentioned that the starting point of Martin's (1992) model of genre is language. As regards language in systemic functional linguistics one distinguishes three 'levels of context' (Halliday 1985): field, tenor and mode. Field refers to what is happening in a specific situation, whereas tenor is linked to the participants, and their roles and relationships in the situation. Mode describes the language that is used in more detail, and the channel of communication and function of language in a specific context (Halliday & Martin 1993: 32-33). In Martin's (1992) model, field, tenor and mode together form register (Martin 1992: 495, Halliday & Martin 1993: 33). Register is "a semantic concept" that describes "a configuration of meanings that are typically associated with a particular situational configuration of field, mode and tenor" (Halliday & Hasan 1989: 38-39). In Martin's (1992) model of genre, register functions as "the expression form of genre" (Martin 1992: 495), and so genre and register are seen as complementing each other. Genres are viewed by Martin (1987) as "staged, goal-oriented social processes"; or in more colloquial terms, "how things get done" (Martin 1985: 250). According to the definition(s), many items fall into the category of genre, including recipes, lectures, and even making appointments (Martin 1989 referred to in Kay & Dudley-Evans 1998: 309). Broadly speaking, the concept of genre "embraces each of the linguistically realised activity types which comprise so much of our culture" (Martin 1985: 250).

Martin (et al 1987) is the most prominent figure in terms of genre theory within systemic functional linguistics. Together with other scholars (Macken et al 1989) Martin (et al 1987) has also focused on the teaching of genre. Generally, genre-based teaching applications of the systemic functional approach concern primary, secondary and adult education (Hyon 1996: 699). In the 1980s and 1990s the focus of researchers (Martin 1989) was on primary school students and the genres they had to produce at this level of education (Hyon 1996: 699, Kay & Dudley-Evans 1998: 310); but more recently the focus has shifted towards adult education, in particular

that of adults with a migrant background (Hyon 1996: 699). Broadly speaking, the main aim and concern of genre-based teaching application is to empower language students. Genre-based knowledge can be used to empower language learners to master genres and also to empower them to be successful in society. Underlying this approach is the ideological idea of social (in)justice (Hyon 1996: 701). Not all language learners have access to socially important genres such as reports or letters (ibid.). This lack of exposure can have negative effects on school performance (in the case of students) or on the life and work performance, in the case of (migrant) adult learners (ibid.). This explains why genre knowledge is regarded as beneficial and thus should be considered when designing syllabi or tasks.

Systemic functional scholars have developed a range of instructional frameworks[5]. One framework that was used in the course of the Disadvantaged Schools Program (DSP) in Sidney is the so-called 'Martin model of genre', also known as the 'DSP 'Wheel' Model of Genre Literacy Pedagogy' (Macken et al. 1989 referred to in Cope & Kalantzis 1993: 10 and Hyon 1996: 704). This takes the form of a wheel, since it presents a cycle of teaching and learning (ibid.) The teaching-learning cycle can be divided into three phases: (a) the modelling phase, which is teacher-centred as the teacher begins by presenting a type of text, its structure and its lexico-grammatical features; (b) the joint negotiation phase, in which teacher and language students work with and construct the relevant text type; and (c) the independent construction phases, which provides students with the possibility to construct the relevant text type independently (Macken et al. 1989 referred to in Cope & Kalantzis: 10-11 and Hyon 1996: 704-705). To illustrate how the different phases can be realised in the classroom the model is presented below in its entirety:

[5] For more information see Hyon (1996).

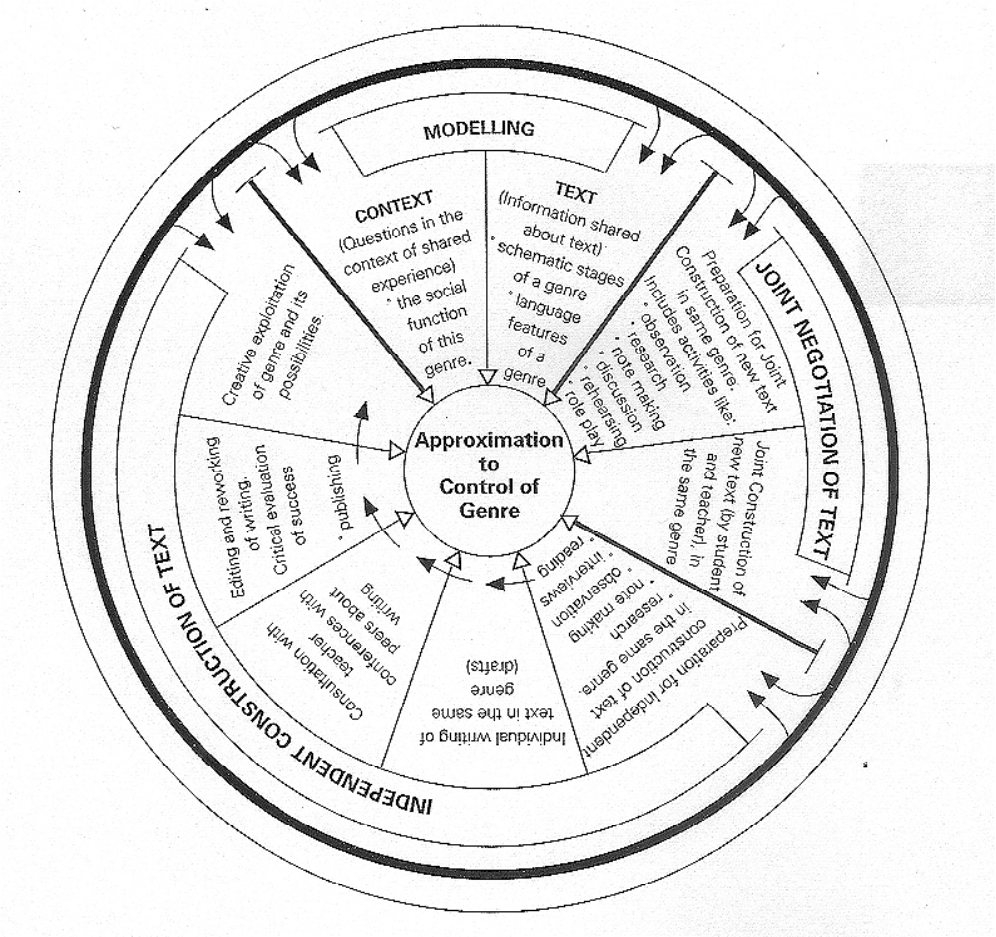


Figure 2: The Martin/ DSP 'Wheel' Model of Genre Literacy Pedagogy (Macken et al. 1989 referred to in Cope & Kalantzis: 11)

In conclusion, it can be stated that due to numerous publications of instructional frameworks, the systemic functional approach to genre attracted academic as well as professional audiences (Hyon 1996: 706). In the next section, I will introduce another approach of genre – the new rhetoric approach. Unlike the systemic functional approach, the new rhetoric approach was published for and also attracted mainly an academic audience (ibid.).

2.3 The new rhetoric approach to genre

The new rhetoric approach, also called (North American) New Rhetoric Studies, the American School or simply New Rhetoric, is another approach concerned with genre (Bhatia 2004: 10, Hyon 1996: 695). In general terms, New Rhetoric is traditionally linked to L1 teaching at the tertiary level of education, in particular to the teaching of rhetoric and professional writing (Hyon 1996: 696). Among the scholars representing this approach are Bazerman (1994), Freedman & Medway (1994) and Miller (1984, 1994). Miller is of especial importance since her article *Genre as Social Action* (1984) has had a lasting effect on the view of genre adopted by the approach (Bhatia 2004, Flowerdew. L. 2005, Hyon 1996). For Miller (1994: 25), genre is "more than a formal entity; it [is] pragmatic, fully rhetorical, a point of connection between intention and effect, an aspect of social action". Genres are specifically seen as "typified rhetorical actions based in recurrent situations" (Miller 1994: 31). This typification of rhetorical actions makes it possible to recognise a specific genre as such (ibid.). In addition, genres "as recurrent patterns of language use, […] help constitute the substance of our cultural life" (Miller 1994: 37). Recurrent patterns result from the needs of individual writers to accomplish their communicative goals (Bazerman 1988: 316). This explains why regularities occur. Regularities in turn, can lead to institutionalisation and codification because they are regarded "to be generally and explicitly advisable" (Bazerman 1988: 316). However, it is important to appreciate regularities as 'fluid' rather than timeless (Bazerman 1988: 315) since they are a means to accomplish goals. In general terms, for New Rhetoric scholars genres function as connectors or mediators between "private intentions and social exigence" (Miller 1994: 37). New Rhetoric scholars therefore see genres as fulfilling social purposes in their situational contexts, and so the focus is on functional and socio-contextual aspects of genres (Hyon 1996: 696).

One important concept linked to the New Rhetoric approach is that of genre systems (Bazerman 1994). System of genres are "interrelated genres that interact with each other in specific settings" (Bazerman 1994: 97). Depending on the situation and on the discipline, only specific genres may precede other genres. This sequencing may vary in fixedness according to domain (Bazerman 1994: 98). In New Rhetoric genres and genre systems are sometimes analysed from an ethnographic point of view. Some New Rhetoric scholars (Bazerman 1994) have made use of ethnographic methods such as interviews and observations in order to investigate genres. Since

genres are seen as social action, ethnographic methods enable researchers to retrieve information regarding socio-contextual aspects such as the attitude of community members towards specific genres (Hyon 1996: 698). In general terms, New Rhetoric scholars mainly explore genres in terms of social functions and their contexts rather than text structure or form, thus describing genres, their roles and context (Hyon 1996: 703). In regard to teaching, the overall aim of this approach is therefore to point out these aspects of genres to students (Hyon 1996: 698) in order to empower them to produce genres themselves. Genre knowledge is seen as beneficial, since only:

> [t]hrough an understanding of the genres available to us at any time we can understand the roles and relationships open to us. An understanding of generic decorum will let us know whether it is ours to ask or answer, to argue or clarify, to declare or request. (Bazerman 1994: 99)

Since New Rhetoric scholars have been preoccupied with enabling L1 university students to understand the socio-contextual aspects of genres, they were less concerned with the development of instructional frameworks. Accordingly, the approach failed in terms of introducing explicit teaching applications (Hyon 1996: 703) [6]. This may also explain why New Rhetoric literature addressed and attracted a mainly academic audience (Hyon 1996: 706).

To conclude, it can be said that despite the shortcomings of this approach to bring forth explicit guidelines for the use of genres in the language classroom, it has shaped the way that genres are perceived today. The viewpoint that genres are determined by intentions and thus are 'fluid' (Bazerman 1988) rather than static is a commonly-held one, and has influenced the concept of genre in other approaches, especially the one adopted by the ESP approach, which will be introduced in the following section. It, like the New Rhetoric school, attracted mainly an academic audience, although it is oriented more towards practicality (Hyon 1996: 706)

[6] For further reading see Hyon (1996).

2.4 The ESP approach to genre

Another approach to genre is the ESP approach, sometimes referred to as the British ESP School (Bhatia 2004: 10). It is represented in the works of Swales (1990) and Bhatia (1993). The underlying idea of this approach is to explore spoken and written genres from academic and professional settings according to their communicative purposes and realisations (Bhatia 2004: 9). It is the writer's communicative purpose that determines the formal characteristics, such as grammatical and lexical choices, of a genre text (Dudley-Evans 1994: 219). It is therefore the communicative purpose of a genre that allows it to be differentiated from other genres (ibid.). According to Dudley-Evans (1994: 219), the view of genre in the ESP approach has been influenced by definitions created by Miller (1984) and also Martin (1989), insofar as genre in the ESP approach is seen as:

> [a] means of achieving a communicative goal that has evolved in response to particular rhetorical needs and that a genre will change and evolve in response to changes in those needs. The emphasis is thus on the means by which a text realizes its communicative purpose rather than on establishing a system for the classification of genres. (Dudley-Evans 1994: 219)

According to this extract, genres are dependent on psychological factors such as strategic choices that can shape how a genre text is constructed. Genre can be therefore considered as a "dynamic social process" (Bhatia 1993: 16).

With regard to teaching, this approach of genre analysis is interesting since it is based on the assumptions that (a) writers in a certain field follow conventions when producing genre texts (Bhatia 1993: 29) and thus contribute to the development of 'prototypes' of genres, and that (b) language teachers and learners can explore these prototypical forms of genres (Upton & Connor 2001: 317) in order to create genre texts independently. Not only do teachers and researchers turn to the ESP approach to investigate conventions, content and language of a genre, they do this in order to explore how genres operate within discourse communities (Dudley-Evans 1994: 219-220). The concept of genre is closely connected to the one of discourse community (Swales 1990: 24), since members of a discourse community partly communicate through the genres that are particular to their community (Dudley-Evans 1994: 220, Swales 1990: 24).

So far, the focus has been on the basic principles of the ESP approach, the genre view adopted by it and the concept of discourse community that is linked to it.
In the following paragraphs I will explore how genre can be exploited for the ESP classroom. Researchers dealing with teaching applications include Swales (1990) and Bhatia (1993), the latter proposing a seven step process for analysing a genre. Language students can follow these steps in order to investigate and finally construct specific genres (Hyon 1996: 703). Bhatia's (1993) seven steps of analysing a genre are introduced in the following section since they also serve as a framework for the study presented in this paper. As mentioned previously, Bhatia (1993) offers a detailed model for analysing a genre. The seven step process can facilitate the investigation of any unfamiliar genre. Not all seven steps have to be utilised. Whether all the steps have to be taken into consideration, depends on the background knowledge one has, the emphasis one would like to put and, of course, the purpose of the analysis (Bhatia 1993: 22). The steps are:

1. Placing the given genre-text in a situational context
2. Surveying existing literature
3. Refining the situational/contextual analysis
4. Selecting corpus
5. Studying the institutional context
6. Levels of linguistic analysis
   Level 1: Analysis of lexico-grammatical features
   Level 2: Analysis of text-patterning or textualisation
   Level 3: Structural interpretation of the text-genre
7. Specialist information in genre analysis
   (Bhatia 1993: 22-36)

The first step involves the placing of a genre-text in a situational context. By drawing from individual experience and knowledge of the discipline, from real-life, and from what Bhatia (1993: 22) calls 'internal clues' in the text, the user is able to contextualise the genre text. This way, the user gains an insight of why a genre is written in a particular style. As not all users possess the same levels of background knowledge, those with a lack of background knowledge are capable of adjusting by surveying existing literature (Bhatia 1993: 22).

In order to analyse an unfamiliar genre, it is therefore necessary to follow step two of Bhatia's (1993) model - surveying available literature on linguistic analyses of the relevant genre or related ones; theories, methods and instruments of genre analysis; guide books of the speech community concerned; and other existing literature that might be of relevance in terms of gaining more information about the relevant speech community (22-23).

The third step involves refining the situational or contextual analysis. What this implies is that it is essential to (a) have a speaker or writer as well as an audience profile in mind, to (b) consider their relationship and (c) their aims, to (d) place the discourse community in socio-cultural, philosophical and/or historical terms, and to (e) give thoughts to linguistic traditions as well as related and surrounding texts (Bhatia 1993: 23ff). To achieve this, it is also important to consider genre constellations (ibid.). Individual genres stand in relation to each other, both within a single discipline and also in inter-disciplinary situations. The genre constellation of 'genre set' can be traced back to Devitt (1991). It describes genres that a " particular individual – or more usefully sometimes a class of individuals – engages in, either or both receptively and productively, as part of his or her normal occupational or institutional practice" (Swales 2004: 20). Bazerman (1994), in contrast, advocates 'genre system', which encompasses "a complete set of discursive forms" (Bhatia 2004: 55) that is not restricted to one particular individual. Another type of genre constellation is genre network, which describes the sum total of all genres of a particular sector (Swales 2004: 22). Another term that is used to refer to the same concept is 'disciplinary genres' (Bhatia 2004: 55).

Step four involves 'selecting the corpus'. It is important to determine one's own criteria for what should be and should not be incorporated into the corpus. Therefore, it is helpful to define and identify the relevant genre or sub-genre beforehand (23-24).

Following Bhatia's (1993) model, one should study the institutional context as a fifth step. This involves studying the rules and conventions in the relevant institutional setting. Bhatia (1993) refers back to step number three as they are strongly related (24).

In the sixth step, Bhatia (1993) advocates looking at those levels of linguistic realisation at which the most distinctive features occur. These three levels are the analyses of lexico-grammatical features, text-patterning and structural interpretation. The linguistic analysis of lexico-grammatical features can be useful, as the high or low frequency occurrence of certain syntactic properties confirms or rejects initial intuition. A focus on surface features such as word frequency is not useful, since no information is gained regarding how communicative purpose is achieved (24-25).
Another level of linguistic analysis is the analysis of text-patterning. This aspect of analysis is concerned with tactical choices employed by members of a particular speech community. It represents a way forward, from language description towards an explanation on why certain genres are written the way they are (Bhatia 1993: 26-29). The third level of linguistic analysis is concerned with the structural interpretation of the text-genre. This is sometimes referred to as the move structure or macro-structure analysis (Swales 1990). These analysis reveals information on cognitive aspects of how a genre and its communicative purpose are organised. Hereby, the focus is on moves that together are realisations of the underlying communicative purpose (Swales 1990: 141). Swales' (1990) publication on the move structure of research article introductions[7] (and the so-called CARS model – Create A Research Space model) is widely seen as pioneering work (Dudley-Evans 1994, Flowerdew L. 2005). Since his work in 1990 many studies have been published on the move structure of other genres (e.g. Henry & Roseberry 2001, Upton & Connor 2001).

The final step in the process is to find and consult a specialist, ideally is a practising member of the discipline, and to 'confront' him/her in order to confirm and add validity to the findings (34).

[7] For a more detailed account see section 4.5.

In this paper, Bhatia's (1993) seven steps of analysing unfamiliar genres serve as a basis for the analysis of the medical genre patient information sheets (PIS). However, the analysis performed in this study deviates from Bhatia's steps in the following respects:

1. Bhatia (1993) suggests consulting practising members of the relevant discipline in order to confirm the findings (step seven – the final step of the analysis). During the course of my research and the writing process, I was in constant contact with my 'specialist informants'[8]. I was therefore able to gain valuable insights into the practices of the medical discourse community. It was only possible once I had this information to 'place' and 'study the genre-text in a situational context' (Bhatia's steps one, three and five). Steps two and three ('surveying available literature' and 'selecting a corpus') are followed in this paper.[9]

2. The analysis present in this paper also deviates in another respect from Bhatia's (1993) seven steps of analysing a genre. Bhatia (1993) suggests analyses on three levels: the surface level (step 6.1), the level of textualisation (step 6.2) and the level of discourse patterns (step 6.3). However, he does not offer explicit frameworks for the analyses, and it is difficult to carry out these three analyses independently of each other. Therefore in this study the analyses of textualisation and the structure are combined. As far as the surface level of language is concerned, a statistical approach has been chosen for the purposes of this study. The vocabulary of patient information sheets (PIS) is analysed according to range and frequency. As a next step, the co-occurrences of words are considered. This only constitutes a small part of the thesis, since the scope of this paper is limited. Significant attention is paid to the move structure analysis of PIS (step 6.3), since it reveals valuable information regarding the communicative purposes of PIS.

In the previous sections the focus has been on distinct approaches to genres. In order to stress the most important points of each approach, I have produced a summary, which appears below. The main characteristics and differences between the systemic

[8] My specialist informants are two practicing nurses.
[9] For a more detailed account see chapter 4.

functional approach, the New Rhetoric approach and the ESP approach to genre are displayed:

	Systemic functional approach	***New Rhetoric approach***	***ESP approach***
Most important representatives	Martin (1992), Martin et al. (1987)	Miller (1984, 1994), Bazerman (1994)	Swales (1990), Bhatia (1993)
Genre definition	Genre as a staged, goal-oriented, purposeful activity (Martin 1987)	Genre as social action (Miller 1984/94)	Genre as a set of communicative purposes (Swales 1990)
Basic assumptions	Genre knowledge enables to successfully function in society (Kay & Dudley-Evans 1994)	Genre is rhetorical and connects intentions and effects (Miller 1994)	Communicative purpose determines grammatical and lexical choices (Dudley-Evans 1998)
Overall teaching aim	To help students to master academic and professional texts (Hyon 1996)		
Goals of genre-based pedagogy	Genre knowledge empowers (underprivileged) language learners to master texts and thus leads to social success	Students should be enabled to understand the social functions and the contexts of genres	Students explore the structure and conventions of a genre in order to construct it
Ideological framework	Social justice (see above)	Lacking	Lacking
Contexts of genre-based pedagogy	L1/ L2/ EFL Primary, secondary and adult education	L1 Tertiary education	EFL/ ESP Secondary and tertiary education
Instructional frameworks	e.g. Macken et. al (1989): 'The Martin model'	Lack of instructional frameworks (Hyon 1996)	e.g. Swales (1990), Bhatia (1993): 'seven-step analysis'
Audiences	Academic/professional	Academic	Academic

Table 1. Distinct approaches to genre: an overview (based on Hyon 1996)

2.5 Developments in genre theory: towards a multi-perspective model of discourse

In the past few years, genre theory in general was primarily concerned with language teaching and learning (Bhatia 2004: XIV). According to Bhatia (2004), this narrow focus on language pedagogy resulted in the investigation of "simplified and idealised genres" (Bhatia 2004: XIV). Genres in reality, however, are more complex and often subject to change. They should be regarded as dynamic and sometimes hybrid in nature. There is therefore a considerable gap between the idealised genres (mainly covered by the literature on genre) and the genres that are used in the real world (ibid.). In order to bridge this gap, Bhatia (2004) advocates another view of genre, which moves towards a multi-perspective and multidimensional genre concept. Bhatia (2004: 18) proposes what he calls "a multi-perspective four-space model of discourse analysis". This is illustrated in the figure below:

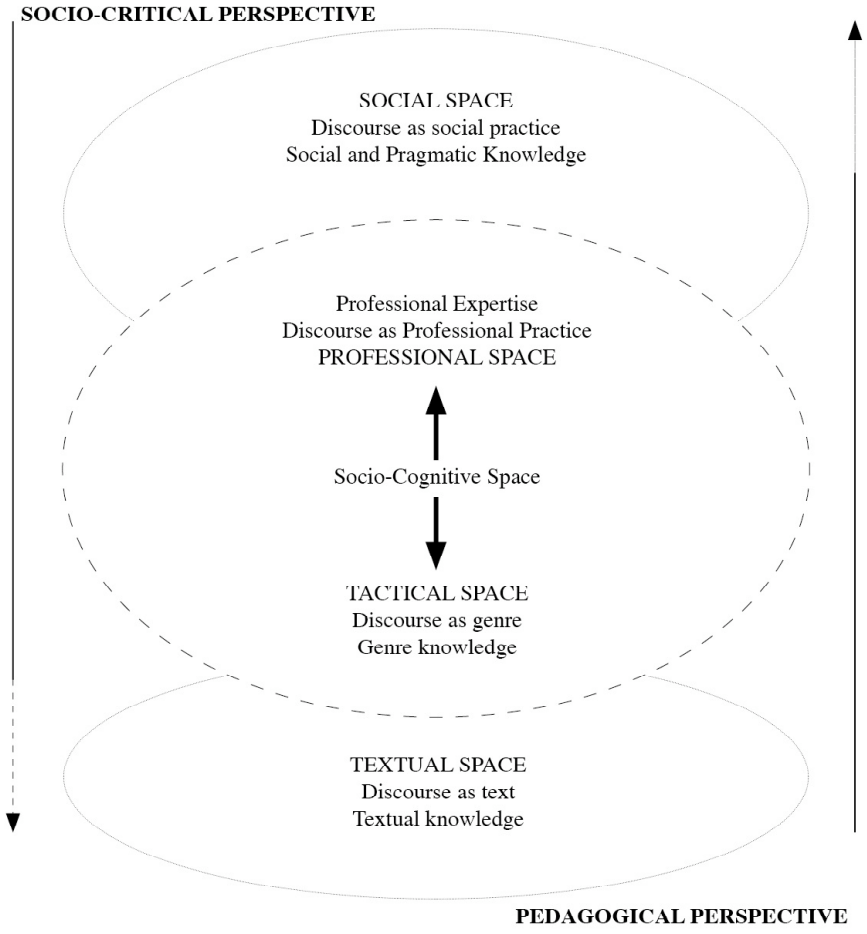


Figure 3. A multi-perspective four-space model of discourse analysis (Bhatia 2004: 19)

For Bhatia (2004) it is essential for genre theory to take into account the socio-cognitive factors that influence the construction, interpretation and analysis of genres. These socio-cognitive factors have traditionally been neglected (Bhatia 2004: 112), since genre theory was mainly focused on the analysis of linguistic factors, text-internal factors such as contextual features (e.g. audience, attitude, social distance, background), textual features (e.g. text patterns) or intertextual features (e.g. surrounding texts) (125-127). These text-internal features were used (mainly by the genre analysts themselves) to access genres. It has been suggested that members of a specific discipline tend to use text-external factors such as discursive practices (e.g. common modes of communication) and procedures (who does what at what stage and by which means) as well as knowledge regarding the disciplinary culture (professional goals and conventions) to approach specific genres (124, 127-133). Therefore, discursive practices, discursive procedures and disciplinary culture should be ascribed greater value.

In the multi-perspective model of genre as advocated by Bhatia (2004), professional expertise is a key concept (Bhatia 2004: 146ff). The concept itself is multidimensional in nature and therefore requires a multidimensional approach. In regard to teaching, it is important to understand how professional experts become experts. Professional expertise is developed slowly but gradually "in different ways and at different stages" (Bhatia 2004: 147). Broadly speaking, Bhatia (2004: 146) differentiates three elements that contribute to professional expertise: professional practice, disciplinary knowledge and discursive competence. The latter also includes generic competence (the competence of mastering genres in order to function within a discourse community) (Bhatia 2004: 145). These three key elements contribute to form professional expertise.

As far as genre analysis in the multi-perspective model is concerned, it can be said that genre is investigated in a four-part framework for analysis, involving: textual space (lexico-grammatical features, their textualisation, discourse patterns); socio-cognitive space, also referred to as tactical space (professional practices, discoursive practices); socio-critical concerns (ideology and power); and ethnographic concerns

(beliefs, attitudes and social structure) (Bhatia 2004: 160-163).[10] This model of genre and genre analysis is insightful as it draws from different perspectives (ibid.). By moving towards a multi-perspective view of genre and genre analysis:

> we are more likely to find increasing flexibility, fluidity and tentativeness in our understanding of generic integrity. This may be a consequence of blurred boundaries between genres increasing tendencies toward mixing and embedding of genres, and a natural inclination on the part of expert users of language for innovation and creativity in their construction as well as consumption of genres in real-life contexts. (Bhatia 2004. 181)

In this chapter I have explored the development of written discourse analysis, the shift of interest from text linguistic to discourse organisation, and to contextualisation of discourse and towards a multi-perspective approach to genre. I have also outlined the different schools of genre analysis, thereby placing special emphasis on the ESP approach and Bhatia's (1993) steps of analysing genre in particular, as they serve as frameworks for this paper. Genre analysis as introduced by Bhatia (1993) involves working with corpora. The convergence of genre and corpora has proven to be insightful, especially in the field of ESP. Therefore, the next chapter is dedicated to (genre-based) corpus studies in language pedagogy and in particular in ESP.

[10] For more information see Bhatia (2004).

3 (Genre-based) Corpus studies and ESP

In recent years corpus studies have attracted more and more attention from linguists and language teachers since corpora are useful tools for studying language use (Bowker & Pearson 2002: 9). Corpora as "store[s] of used language" (Hunston 2002: 3) do not provide information about language, but in combination with analysis software they do enable researchers and teachers to make observations, gain new insights and "perspective[s] on the familiar" (Hunston 2002: 2-3). This chapter will describe the origin and the issues of corpus studies in language teaching as well as the use of corpora in the language classroom. Finally, the chapter will take a closer look at the types of corpora that are of relevance to ESP studies and their use.

3.1 Corpus studies and pedagogy: origin and issues

It has been mentioned that research on corpora has become extremely popular over the last decades. This development has had an enormous impact on linguistics and language teaching. For Seidlhofer (2003), "corpus linguistics clearly represents a challenge to traditional ideas" (Seidlhofer, 2003: 78). Hunston (2001:1), in comparison, implies something even stronger. She sees corpus linguistics as revolutionising (ibid.). The scope of change or challenge that corpus study has evoked can be debated; however the fact remains that change was brought about. How corpus linguistics changed from a so-called "fringe activity" (Leech 1997: 4) to a mainstream one challenging traditional ideas is investigated now.

Technology always had and continues to have an immense influence on language teaching[11]. Innovations in technology are followed by or even mirrored in innovations in teaching (Stubbs 2000: 24-25). In the table below technological innovations of the past and the correlating teaching techniques, theories and syllabi are presented:

[11] See e.g. Howatt (1984) for more information.

Decade	***1880s***	***1950s***	***1990s***
Technology	Telephone, phonograph	Tape recorder, language labs	Computers plus corpora
Techniques	Broad and narrow transcriptions	Immediate constituent analysis	Collocational analysis
Theory	Phoneme theory	Structuralist linguistics	Lexico-grammar
Syllabus	Reform movement	Audio-lingual method	Lexical-syllabus

Table 2. Correlation of inventions and language teaching (Stubbs 2000: 25)

As much as 50 to 100 years ago, technological advances such as the telephone or the tape recorder were leading to the birth of techniques, theories and teaching syllabi (Howatt 1984: 177). The invention of the telephone, phonograph, tape recorder and language laboratory provided the basis for syllabi such as the reform movement and the audio-lingual method (Stubbs 2000: 25). A similar phenomenon could be observed in the 1990s. Corpus linguistics experienced a boom with the rising availability and broad accessibility of technological aids. Even though, this method of text analysis first came up in the 1960s with the compilation of the *LOB* and the *Brown* corpora, the initial breakthrough was still to come. The groundbreaking work by John Sinclair (1987) in the 1990s marked a milestone in corpus linguistics. The pioneering Sinclair (1987) and his Cobuild project group triggered a new popular, one might say even fashionable movement (Gavioli 2005: 17). The main aim of the Cobuild project was to develop a more realistic description of the English language for the language classroom. Teaching materials that were based on the outcome of the project were developed. In general, the project led to new insights in various respects. First of all, it highlighted the fact that more data could be compiled and analysed with the aid of new technology. Furthermore, the project demonstrated that language performance could be taken as a basis for language analysis (Gavioli 2005: 17-18). Notably, the Cobuild project evoked and revived discussions in the 1990s since it connoted "an equivalence between (a) a corpus and real language and (b) corpus-based descriptions and 'more realistic' students' language production" (Gavioli 2005: 18). These implications provoked considerable criticism. Various linguists voiced criticism (Cook 1998, Widdowson 1991), which mainly addressed the issues of language reality and the inclusion of real language in language lessons (Gavioli 2005: 18, Hunston 2002: 193).

The first argument to challenge the use of corpora in teaching was the view that corpora do not comprise 'real language' since "they are only a sample of language production" (Gavioli 2005: 18). This can be explained by the very nature and design of corpora. Large general corpora (such as the BNC) consist mainly of written texts. Spoken material only accounts for about 10 per cent in the BNC. Spoken language is thus presented only partially (Gavioli 2005: 18), therefore, the BNC as other corpora reveals (as do other corpora) "only a partial account of real language" (Widdowson 2000: 7).

Another issue that was addressed in the 1990s debate was the question of inclusion of 'real language' in the language classroom. Being exposed to 'real language' does not necessarily imply being able to understand or learn it (Gavioli 2005: 18).
Gavioli (2005: 18-19), with reference to Carter (1998), stresses that instances of real language (as found in corpora) very often do not exemplify features of spoken language as clearly as invented examples. In addition, it is still unclear to what extent corpus-based data should be manipulated for the purpose of language teaching (ibid.).

Another argument (linked to the one previously introduced) focuses on language reality. Whether students and learners should be exposed exclusively to 'genuine language' (in Widdowson's terms 1978: 79-80), meaning actually produced language is debated. Learners of a language also need to be able to produce language in a communicative context (authentic language) (Gavioli 2005: 19)[12].

So far, we have discussed criticism that has been made by applied linguists since the 1990s. Their objections mainly questioned the role of corpora and its over-enthusiastic use in pedagogy. Notably, Gavioli (2005: 20) insists that the controversies in the 1990s only highlighted the limitations of corpus linguistics without exploring its potentials for language teaching. The potentials of corpus work in language pedagogy will be discussed in the following section.

[12] For more information on the debates concerning corpus linguistics and language teaching see Seidlhofer (2003).

3.2 Corpus studies and the language classroom

In the previous sections, I have dealt with the discussions surrounding the rise of corpus studies. Interestingly, it was not just applied linguists who had reservations regarding corpus work, but also language teachers. Initially teachers were reluctant to use corpora directly at the secondary or tertiary level of education. One reason for this was the poor (or even non-existent) state of communication and negotiation between researchers and language teachers. Another reason for the resistance to corpora was the discrepancy between what is useful in terms of pedagogy and what is insightful and interesting in terms of linguistics. Finally, the reluctance to use corpora directly in schools and universities might also be demonstrated by the growing interest in learning processes rather than learning content and the limited accessibility to computer and corpora (Aston 2001: 1). The reluctance to use corpora directly in the language classroom began to slowly (but gradually) weaken. Corpus studies that highlighted the use of corpus work in language teaching and the revived interest in form-focused teaching contributed to a change in perception of corpus work in the language classroom (Gavioli 2005: 22).

In spite of teachers' resistance to explore corpora directly in schools and university, some researchers (Fligelstone 1993, Leech 1997, Renouf 1997, Flowerdew L. 2001, Gavioli 2005) have focused on the practical aspects of corpus linguistics and the integration of corpora in the language classroom. In the following section the direct and indirect use of corpora in language teaching will be outlined.

3.2.1 The direct use of corpora in language teaching

Corpora can be used in different ways in teaching. Fligelstone (1993) identified three ways in which corpora can be exploited in the language classroom. According to him, it is possible to (a) teach about corpora, (b) teach to exploit corpora or (c) exploit corpora to teach (Fligelstone 1993 in Leech 1997: 6).

'Teaching about corpora' usually takes place at the tertiary level. Students can or should enrol classes that familiarise them with corpus linguistics and important topics connected to it. Among the topics are existing corpora; ways and means of analysing corpora, the applications of and the language view inherent in corpus

linguistics (Leech 1997: 7). In other words teaching about corpora means "teaching about the principles and theories of corpus linguistics" (Renouf 1997: 256). An integral part of it is, of course, the actual "use and exploration of corpora" (Leech 1997: 7). This is clearly related to Fligelstone's (1993) second category 'teaching to exploit'.

Language learners are as capable of exploring corpora as researchers and teachers are. The teacher then leaves his/her role as the source of information in order to act as and advisor and guide. It is evident that in the case of Fligelstone's (1993) second category 'teaching to exploit', the learner (rather than the teacher) is at the fore. Learners should be enabled to investigate corpora with the aid of computer software, therefore the emphasis in on the practical aspect of corpus linguistics (Renouf 1997: 256). Sample tasks uphold a central role (Leech 1997: 8). These tasks can be 'divergent' or 'convergent'. If a task is 'divergent', then different learners are given the same task, instructions and data. The learners are likely to obtain different, but valid results. Hence, the assessment of a learner's performance relies on the student's ability to observe, analyse, infer, interpret and so on (Leech 1997: 8ff). If a learning task is 'convergent', then all students will be provided with the same instructions and data. However, they will be likely to deduce the same results as the nature of the corpora is usually limited or tailored to the needs of the learners. Computer learner packages on corpora usually comprise convergent tasks. These tasks typically fulfil three out of four qualities that are particular to computer-based tasks. They are automatic, learner-centred (in the sense that it is the learner who explores the corpus) and tailored (adaptive). The quality that is not realised is 'open-endedness'.
This is only realised when learners can freely select from a range of responses. Multiple-choice and gap-filling exercises are ideal for convergent tasks (Leech 1997: 11-12). It can be said that 'teaching to exploit' is becoming more popular at the tertiary level of education. There is interest and growing demand to equip future teachers with knowledge and skills regarding corpora and their exploitation. In this context, Renouf (1997) adds another category to those advocated by Fligelstone (1993). For Renouf (1997), another category of teaching corpus linguistics is 'teaching to establish resources' (Renouf 1997: 256). This somehow converges with what Fligelstone (1993) describes as 'exploiting to teach'.

Fligelstone's (1993) third category of how corpora can be used in teaching is 'exploiting to teach'. This category is aimed at teachers rather than students. The underlying idea is to familiarise teachers with corpus linguistics in order to enable them to derive teaching materials from corpora (Renouf 1997: 256). Corpora encapsulate real-life language and are therefore of interest to language teachers.

Instances of real language use can be used to explain vocabulary and/or grammar, as they provide information regarding frequency, phraseology and collocations (Hunston 2002: 13). Sinclair (1997) suggests five precepts that centre on corpus data, which are addressed to language teachers who wish to engage in corpus studies. Firstly, teachers should attempt to present only real-life examples in the classroom as invented examples are often far-fetched, inappropriate and do not mirror actual usage. Secondly, language teachers and learners should make use of their intuition about language and about how words and sentences are used and combined. Moreover, language teachers should inspect the contexts or rather co-texts. In this context, Sinclair (1997) advocates a closer investigation of verbal surroundings of a word or a phrase since multi-word choices will then become apparent (Sinclair 1997: 34). According to Sinclair (1997: 34), "[c]o-selection tends to undermine the notion of word-meaning […]"; hence, it should be placed at the forefront of attention. Rather than learning and teaching endless lists of vocabulary, space should be allocated for the teaching of multi-word units in the language classroom (Sinclair 1997: 35). Another advice to teachers who wish to use corpora to extract teaching material is to 'teach by meaning'. Dictionaries contain information on meanings of words, but fail to offer structures. Grammar, on the other hand, tends to regard words as possessing only one meaning (some words can function as noun and verb, for instance, and their meaning may vary accordingly). It is important to recognise that this may be a burden for language learners. Relating meaning and form is one possible way to reduce the learning load (Sinclair 1997: 35). Last, but not least, a corpus allows the researcher or teacher to observe a language, its changes and innovations. Moreover, it shows what is possible and permissible. The strengths of corpora is that they demonstrate the productivity of a language (Sinclair 1997: 37). Another advantage is that they "clarify, give priorities, reduce exceptions and liberate the creative spirit" (Sinclair 1997: 38). So far, I have dealt with the ways in which corpora can be explored directly in the language classroom. Corpora can, however, be exploited indirectly in teaching.

3.2.2 The indirect use of corpora in language teaching

The influence of corpus studies on language teaching is manifold. Corpora cannot only be explored directly in the classroom. The results of corpus studies also have an indirect impact on the language classroom, as reference sources and teaching materials heavily rely on corpora analyses. As mentioned previously, the *Collins Cobuild English Language Dictionary* (Sinclair 1987) was the first English language dictionary that was based on a corpus. Other publishers followed and thus contributed to a new movement. In the following years, numerous dictionaries such as the *Oxford Advanced Learner's Dictionary of Current English* (Hornby 1989) and grammar books such as *Collins Cobuild English Grammar* (Sinclair 1990) (both based on corpus studies) were published. Nowadays corpus studies provide the basis for the majority of reference publications, such as dictionaries and grammar books. This can be explained by the fact that corpus-based lexicography has various advantages. Among the advantages are that computer corpora provide frequency data and authentic language. Furthermore, they facilitate processes such as search, manipulation and updating (Leech 1997: 14).

Another important function of corpus studies is that they also contribute to the development of language teaching materials. They provide the basis for teaching materials in the sense that they (a) offer information on word frequency, (b) present instances of authentic language use and (c) go hand-in-hand with the development of computer-based learner packages on corpora exploitation (Leech 1997: 16).
Ongoing debates on corpora in the language classroom address the topic of frequency data and the value ascribed to it. Large corpora studies made it possible to obtain frequency data such as word frequency lists for the first time. The interpretation of this kind of data continues to constitute a challenge. Whereas some researchers or syllabi designers often praise frequency data, be it lexical or grammatical frequency data (Flowerdew J. 2001), others express reservations on the basis of that which occurs most frequently does not necessarily have to be the most important thing to teach and learn (Leech 1997: 16).

As corpora function as sources of information for language teaching, they can also function as sources of information for language testing. I have previously discussed the strengths of computer corpora. Qualities such as being automatic, learner-

centred, open-ended and tailored speak for the use of corpus-based material in language teaching and its testing. It is evident that a strict control of the task and the assessment in terms of 'correct' and 'incorrect' is necessary. Since the corpus is comprised of authentic language use, it is possible to select from different instances of real-life language in order to use them as test samples. In addition, computer corpora can be used to grade real-life samples (Hunston 2002: 193, Leech 1997: 16-17).

3.3 Corpus studies and ESP

3.3.1 Corpora in ESP

Defining corpus

> A corpus is defined in terms of both its form and its purpose. Linguists have always use the word *corpus* to describe a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study. (Hunston 2002: 2)

Bowker and Pearson (2002: 9), describe a corpus as a „large collection of authentic texts that have been gathered in electronic form according to a specific set of criteria". Several characteristics of corpora have been identified so far. Corpora should be 'authentic', 'electronic', and 'large' and should be compiled according to 'specific criteria'. These four characteristics make it possible to distinguish corpora from other types of text collections. Corpora ideally comprise texts that are authentic, i.e. those featuring language use. The texts are not artificial since they encapsulate naturally occurring language, and not language that was produced to be analysed. Nowadays corpora in corpus linguistics are computer corpora. The convergence of computer and corpora has facilitated the corpus compilation and the analysing process. Computer corpora and software packages (also known as corpus analysis tools) allow researchers to compile, access, manipulate and analyse large collections of data (Bowker & Pearson 2002: 9-10).

Regarding corpora as large collections of data or texts, there is no common agreement on the ideal size of a corpus. Although in the literature there is evidence that a corpus should be 'large', it is very often not stated how large a corpus should

be. 'The bigger the better' is a maxim that it is not necessarily applicable to corpora. Small-scale corpora especially in specific fields such as Technical English or Medical English sometimes offer more useful information than general language corpora of a larger size. Small corpora do have disadvantages though. If a corpus is too small, then it is difficult to make valid generalisations. According to Bowker and Pearson:

> well-designed corpora that are anywhere from about ten thousand to several hundreds of thousands of words in size have proved to be exceptionally useful in LSP[13] studies. (Bowker and Pearson 2002: 48)

Corpora also require explicit criteria. It is possible to include texts from a specific field such as business or medicine. Moreover, the criteria can be refined and narrowed down even further. A corpus may only comprise a specific spoken or written genre, such as telephone conversations and mission statements in Business English, or doctor-patient consultations and patient information sheets (PIS) in Medical English. Furthermore, it is possible to focus on a specific time period or frame, and on speaker groups such as women, men and children, or Britons and Canadians. Since the purpose of each corpus study differs, a different type and design of corpus is required – hence there is a multitude of corpus types (Bowker & Pearson 2002: 10-11). Relevant for this study is the broad distinction of corpora into large general (reference) corpora and smaller special purpose corpora.

Large general corpora vs. small specialised corpora

Since research with corpora is a mainstream activity, new developments are constantly being made. Among the developments of teaching-oriented corpus studies is the use of specialised corpora in ESP settings. These usually small corpora are often compared to large general ones. A large general corpus is regarded as representative of a whole language. It usually amounts to several million words and typically consists of written and/or spoken texts (and types of texts), and mainly covers general language (GL). On the other hand, special purpose corpora, also called specialised corpora, explore a specific subject field and are smaller in size. They usually comprise between ten thousand and several hundred thousand words

[13] Abbreviation for language for specific purposes.

(Bowker & Pearson 2002: 48). Specialised corpora vary in terms of degrees of specialisation, and can cover the language of specific purposes (LSP), a specific genre of a particular field, or even a specific demographic group. Specialised corpora are used to make observations about the aforementioned (Bowker & Pearson 2002: 12, Gavioli 2005: 7).

The emergence of small but specialised corpora naturally attracted attention in ESP circles. Although small-scale corpus studies often represent a challenge to general English description, as the corpora are too small to make generalisations about the whole language, small-scale studies are appropriate for investigating ESP (Gavioli 2005: 55). One of the first small-scale ESP (-like) studies was conducted as early as 1962. Barber's (1962) corpus was comprised of around 23,000 words, based on three research articles from the fields of astronomy, chemistry and engineering. Certain aims of the study including deriving the average sentence length or deriving frequently occurring modal verbs (Gavioli 2005: 55-56). Although Barber's (1962) corpus analysis dates back to 1962, it shares common features with contemporary corpus studies (Gavioli 2005: 55-56). The use of small but specialised corpora seems to be "a sort of natural development in this field" (Gavioli 2005: 56).

In regard to teaching, specialised corpora have various benefits. First of all, they are easier to handle as they are smaller in size. As mentioned previously, the ideal size of specialised corpora is between ten thousand and several hundred thousand words (Bowker & Pearson 2002: 48). General corpora, in contrast, comprise one to several million words – a quantity that is clearly more difficult to manage. In qualitative terms, specialised corpora also have an advantage, insofar as they display linguistic characteristics of a genre and therefore also highlight aspects of a LSP that are of relevance to the learners (Gavioli 2005: 62). It is common practice to use general reference corpora and special purpose corpora together. Hereby the corpora are usually compared in order to determine how the specialised corpus differs from the general one (Bowker & Pearson 2002: 12).

Among the developments in corpus studies in ESP is to focus on individual genres of a specific field, such as letters of applications in Business English or abstracts in

Academic English. In the following section I will discuss the convergence of corpus and genre-based methodologies in ESP corpus work.

3.3.2 Corpora and genres

Corpus work today is seen as an established and acknowledged method in ESP circles (Gavioli 2005: 55). Individual studies may focus on different teaching and/or learning needs, or on different genres. Interestingly, many of the corpus studies conducted in ESP are based on genres (Flowerdew L. 2005: 321). Among those who conducted genre-based corpus studies in ESP are Henry & Roseberry (2001) and Upton & Connor (2001). Henry & Roseberry (2001) as well as Upton & Connor (2001) focused on letters of applications. Upton & Connor (2001), however, also explored cultural differences and politeness strategies. Therefore, genre analysis is not primarily concerned about the language of a specific field such as the Business English or Medical English but is rather "about the conventions of thought and communication which define these areas of professional activity, and how, incidentally, these are given expression, or textualised, in English" (Widdowson 1998: 9). Genres involve typical features that are particular to specific fields. Included in these features is also the generic structure of a genre (Gavioli 2005: 68). According to Hunston (2002: 14), the main aim of a genre-based corpus study is to arrive at a representation or prototype of the relevant genre and to investigate its language. The genre-based corpus studies in ESP that have been performed so far can be divided into two types of studies; the first type investigating recurring combinations of words that reflect rhetorical choices or moves, and the second type concerned with language features and the ways they contribute to the relationship between writer and reader (Gavioli 2005: 56-58).

Regardless of the type of genre-based study, the synthesis of corpus-based methodology and genre investigation is fruitful. "Corpus-based methodologies have been informed by genre principles of text analysis", [… and] genre theories can profit from corpus-based methodologies" (Flowerdew L. 2005: 8). The convergence of genre studies and corpus work has proved very insightful in ESP settings, especially in terms of typical features of genres, such as the generic structure, but also in regard to lexis. In the next section I will deal with the role of specialised corpora in regard to specialised lexis.

3.3.3 Corpora and lexis

A discourse community is typically seen as "having acquired a specific lexis" (Swales 1990: 26). Therefore, novices of a specific discourse community such as students, who are outsiders with a lack of inside knowledge and experience, have to acquire such a lexis. Acquiring lexis means that novices should acquire the semantic meaning of words, but additionally, should acquire "a whole new way of conceptualizing things and talking about them" (Gavioli 2005: 68).

As far as lexis is concerned, corpora are extremely useful tools in identifying and describing core issues for teaching. Secondly, they facilitate and simplify the handling of specialised vocabulary (Gavioli 2005: 58-60). Traditionally, one distinguishes between the language expert, the language teacher and the subject expert (the teacher being responsible for the teaching of a specific subject, such as healthcare). In regard to language teaching, it is possible to identify a significant gap between (a) the text and its linguistic characteristics (teaching is usually regarded as the language teacher's job) and (b) the meaning of the text also covering technical vocabulary (traditionally linked to the subject teacher) (Gavioli 2005: 60). The teaching of technical vocabulary is hotly contested. In the literature, two points of view have manifested. Some researchers (Higgins 1966, Barber 1962) believe that the teaching of technical vocabulary is the subject teacher's rather than the language teacher's job, while other researchers strictly reject that view – on the basis that what is often considered as technical vocabulary is in fact vocabulary from the general language that has specialised meaning within a specific field (Flowerdew, J. 2001: 78). A strong point is made by Strevens (1973), who argues that it is often the teacher who struggles with technical vocabulary rather than the language learner since the latter usually brings knowledge about a specific field (Strevens 1973: 228). An interesting issue is raised here. Little research has been done on technical vocabulary and the identification of it. One reason might be that the teaching of technical vocabulary was not traditionally regarded as part of a language teacher's job (Gavioli 2005: 58). Another reason could be that language teachers simply tried to avoid this type of vocabulary out of convenience and/or insecurity. Since language teachers often lack knowledge of the subject area, the work with corpora that consists of specialised texts is insightful and yields information on the area, the texts, key issues, and more importantly on specialised vocabulary (Gavioli 2005: 58-60).

According to Gavioli (2005: 60), one main contribution of corpus work to ESP is that it enables researchers and especially teachers to investigate technical lexis, and therefore "bridges the gap".

Also in regard to lexis, computer programs enable researchers to access and analyse data more easily. What had to be done manually in the past is now done by machines. How small and specialised corpora can be explored will be discussed in the following section.

3.3.4 Corpus analysis - Exploring and exploiting corpora

In ESP quantitative data always seems to have played an important role. Traditionally the main concern of ESP description was to investigate and analyse actual language occurrence. Therefore, the focus had been on recurring lexical items and grammatical structures (Gavioli 2005: 55). Nowadays, corpora are exploited to extract information regarding, for instance, frequency and phraseology (Hunston 2002: 20, Leech 1997: 18). Small and specialised corpora enable researchers, teachers and learners to gain insights into a language they sometimes are not familiar with e.g. the language of engineering or the language of medicine. Furthermore, specialised corpora are used as tools to design language-teaching materials for a very specific learner group such as business people, architects, nurses or medical students. With the aid of computer software, computer corpora can be used to determine the linguistic characteristics, including word frequency data, collocations and typical grammatical structures of a LSP (Hunston 2002: 20, Leech 1997: 18). To what extent information regarding frequency data and phraseology is insightful will be discussed in the following section.

Frequency

Words occurring in a corpus can be arranged according to frequency, with the aim of developing a frequency list. It is even more insightful to compare different frequency lists, e.g. that of a small specialised corpus to a large general corpus. Word frequency comparisons of this type can show characteristics that are particular to the corpora. To illustrate this, an example is provided. In an attempt to highlight the potential of concordance programs as instruments for course design, John Flowerdew (2001) compared a specialist corpus to a general one. The specialist corpus focused on

biology, while the *Collins Cobuild English Language Dictionary (1987)* functioned as the general reference corpus. In order to compare the corpora, it is useful to consider the most frequent items of the corpora. In the case of Flowerdew's (2001) study, the results are as follows:

Rank	***Specialist corpus***	***Cobuild***
1	THE	THE
2	AND	OF
3	OF	AND
4	IS	TO
5	A	A
6	IN	IN
7	ARE	THAT
8	TO	I
9	IT	IS
10	THIS	WAS

Table 3. The most frequent items in the biology corpus and the *Cobuild* (Flowerdew J. 2001: 77)

According to Flowerdew (2001: 77), there is a similarity between the corpora, but the differences are also worth exploring. For example, in the table *was* is among the ten most frequent items of the *Cobuild* corpus. In the specialist corpus, however, it is only ranked fiftieth. One possible explanation may be that the past tense is more frequently used in the general language than in sciences.

Another possibility of comparing corpora is to contrast the most frequent lexical items. A comparison of lexical items such as nouns often shows great variation between large general corpora and smaller specialised ones. I want to return to the example I previously introduced. Flowerdew (2001) also compared the top twenty nouns of his specialist corpus to the ones of the *Cobuild* corpus (78). The outcome of his comparison is presented in the table below:

Top twenty nouns in the Cobuild general corpus
time, people, way, man, years, work, world, thing, day, children, life, men, fact, house, kind, year, place, home, sort, end
Top twenty nouns in the specialist corpus
cell, cells, water, membrane, food, plant, root, molecules, plants, wall, energy, concentration, organisms, cytoplasm, animal, stem, structure, body, part, animals

Table 4. The top twenty nouns in the *Cobuild* and the specialist corpus (Flowerdew J. 2001: 78)

The results of the comparison show that none of the nouns is featured in both lists. There is a great discrepancy between the most common nouns of the general and the specialised corpus. If one considers the nouns in the biology corpus, one realises that some words are particular to the field of biology such as *cytoplasm* and *membrane*. Yet, there are also words from the general language that have specialised usage such as *wall* and *concentration*. One argument for the use of specialised corpora therefore is that one is able to observe how particular items are used and whether these items have particular meanings (Flowerdew J. 2001: 78).

Phraseology

Corpora also hold information regarding phraseology. With the aid of computer programs and more specifically with the aid of a concordance program, the use of words and their co(n)-texts (Sinclair 1997) can be observed (Hunston 2002: 9). Here again, I want to refer to the study by John Flowerdew (2001), as it perceptively demonstrates how corpora can be exploited. In order to extract the uses of particular words Flowerdew (2001) used a concordance program. Such a program allows researchers to analyse the co(n)-texts of certain words. In the example presented below some of the uses of the item *well* are shown:

uoles/ you know what these are / **well** actually to be precise / leave out
w these words can be adjectives as **well** / akarytotic / prokaryotic / and euk
own food / how do they do it / yes **well** alright / we've heart lots of useful

Table 5. Some concordances for *well* (Flowerdew J. 2001: 91-92)

The analysis of the concordance lines for *well* suggests that *well* is used in three different ways in the biology corpus. It is used either as adverb (*can grow well*), conjunct (*as well as*) or discourse marker (*yes well alright*). In general, concordance programs can be used to explore phraseology. In particular, it can be employed to identify the most important items, their co-texts and syntactic patterns (Flowerdew J. 2001: 91-92).

To summarise briefly: although there is a consensus on the role of corpus linguistics in the study of language – it is widely seen as enrichment in the description of language (Seidlhofer 2003: 78-79) – there is no consensus over its role in language pedagogy. Using corpora in language teaching is not an undisputed enterprise, but in ESP circles this is not the case. (Genre-based) corpus studies are an acknowledged and insightful method in ESP, since information regarding the frequency of lexical or grammatical items, phraseology and co(n)-texts of words can be retrieved; in other words, information that is useful for syllabus design (Flowerdew J. 2001: 79). More generally speaking, "[a] corpus based analysis can provide valuable insights which are of use for language pedagogy and course design" (Nation & Kennedy, 1994: 49). The use of small and specialised corpora is commoner in ESP circles. These seem to be made for the purposes of ESP studies (Gavioli 2005: 55). Therefore, it comes as no surprise that Gavioli (2005) describes the relationship of ESP and corpora as "a happy marriage" (Gavioli 2005: 55).

In the following chapter my focus will be on the genre-based corpus study of this paper. I will discuss the relevant genre of this study – patient information sheets (PIS), before turning to studies that were previously conducted on PIS and other related topics. Furthermore, I will outline the design of the specialised corpus as well as the methods of the genre-based corpus study. In accordance with Bhatia's (1993) seven steps of analysing genre, the next chapter will cover step one 'Placing the given genre-text in a situational context' (in my paper section 4.1), step two 'Surveying existing literature' (section 4.2), steps three and five 'Refining contextual analysis' and 'Studying the institutional context' (sections 4.3 and 4.4) and step four 'Selecting the corpus' (section 4.5).

4 The study – genre, setting, data and methods

4.1 The genre of patient information sheets

Creating a context

As a discipline, medicine has undergone dramatic changes over the last few centuries. Great technological and scientific advances in medicine, such as the decoding of DNA, the advent of prenatal diagnosis, and the changing access to information have challenged both humans and the discipline itself. Contemporary debates over cloning and prenatal diagnosis have raised a number of issues concerning the worth of a human life, and thus have evoked ethical discussions in medicine[14] (Hoejke: 2007: 3, Wolff 1999: 184-188). In addition, the access to information has changed. Initially, doctors were considered as the only source of information regarding a patient's health. Nowadays, patients are able to familiarise themselves with medical topics by extracting information from, for instance, the Internet, in itself a technological milestone. Reliable and less reliable sources disseminate information and so challenge doctors as sources of information. The changing access to and distribution of information has had an impact on the traditional relationship between the roles of the doctor and of the patient. Whereas the doctor-patient relationship was once seen as paternalistic, in the sense that the doctor knew what was best for the patient, like a parent knows what is best for a child, this is not the case anymore. Nowadays, doctors and patients are seen as equal partners that cooperate and strive for the same goals – either healing or easing illness – and who collaborate in order to make decisions (shared decision-making) (Wolff 1999: 199). The patient is perceived as an expert regarding his/her own wellbeing. In the so-called cooperative model, patient empowerment is a key concept (Wolff 1999: 206-210). Educating or informing patients achieve this empowerment.

> Empowering patients through education can be divided into the following areas: biophysiological (identification of the symptoms and signs), functional (activities of daily living, illness and care), cognitive (receiving enough information and the ability to utilize it), social (experience of belonging to and support from the social network), experiential (feeling of appreciation with regard to one's experiences,

[14] Ethics in medicine first evolved in the U.S. where it became an acknowledged discipline in the 1970s. Courses on ethics were soon part of the curriculum for medical students. Only slowly did this interest expand over the borders and eventually reach Europe (Viefhues 1999).

> expectations and feelings), ethical (feeling of appreciation as a unique, autonomous individual) and financial. (Leino-Kilpi et al. 1998, 1999 in Johansson 2005: 13)

In this sense, patient empowerment can be understood as equipping patients with the right 'tools' to identify changes in their condition, thus enabling them to take appropriate independent action and therefore becoming experts of their wellbeing, fully equipped to take responsibility for their own health (Johansson 2005: 13).
Patients will find themselves in situations where decision-making will be necessary. Doctors therefore, should aim to empower the patient to make decisions that are based on sound information (informed consent). Handing out information sheets to patients prior to any medical intervention or procedure is a complementary means of disseminating information and sharing knowledge (Johansson 2005: 220).

Medical discourse

When speaking of medical discourse one implies professional discourse, 'discourse' being used in its broadest sense, as a self-contained system of communication covering jargon, common ideological position (ideology), socialisation, forms of discourse and interpersonal relationships (face systems) (Scollon & Scollon 1995: 95-98). Available evidence on professional discourse suggests that its major purpose is to convey information "as clearly, briefly, directly, and sincerely as possible" (Hoejke 2007: 6, Scollon & Scollon 1995: 94). This form of communication is also referred to as C-B-S style (Lanham 1983). Coined by Lanham (1983), the term stands for "clarity", "brevity" and "sincerity" (Lanham 1983 referred to in Hoejke 2007: 6). This style is regarded as being the most effective form of its kind in written professional communication (Scollon & Scollon 1995: 95-98). Freedom from bias is another characteristic of professional discourse. In addition, communication should be civil and respectful, so as not to appear offensive or disrespectful to the other person/people (Brown & Levinson, 1987 in Hoejke 2007: 6).

Medical discourse is comprised of professional medical discourse as well as institutional medical discourse (Hoejke 2007: 6). Students who enter the discourse community are slowly familiarised with the discourse system, which encompasses ideology, socialisation, forms of discourse and face systems (ibid.). When entering the world of work through internships aspiring medical personnel join the medical

discourse community. Broadly speaking, there are several characteristics that define a discourse community (1990: 24). According to Swales (1990: 24ff), a discourse community has (an) "agreed sets of goals" (24), (b) "mechanisms for intercommunication" and (c) "exchanging information and feedback" (both 25-26), (d) specific genres and thus (e) use specific lexis (26). In addition, a discourse community is (e) dependent on a balance of novice and expert members, the maintenance of balance being necessary for the survival of the discourse (ibid.). Therefore, when aspiring doctors and nurses join the medical discourse community they have to be familiar(ised) with these principles. I will now explore the more detailed implications of this. Most novice members enter the community when they start their career at medical institutions such as hospitals. Once aspiring doctors or nurses have joined the discourse community of medical personnel, e.g. when commencing internships in hospitals, they have to become familiar with the mechanisms of communication, be they mechanisms of intercommunication or mechanisms for providing information and feedback. In the case of medical personnel, for example, this involves the use of mechanisms such as conferences, staff meetings, reports, e-mails, and other forms of correspondence in order to communicate and thus to exchange information (Krois-Lindner 2007). Novice members must not only become familiar with the mechanism of communication, but also have to be aware of the genres that are particular to the discourse community. In medicine such genres can include medical reports, medical histories, patient information sheets (PIS), doctor-patient consultations etc. These genres also display specific lexis such as medical termini, abbreviations and acronyms, which can be confusing or misleading for outsiders and novice members (Swales 1990: 26). To conclude, it can be stated that novice members are important for discourse communities. As members also leave (e.g. through retirement), new members are needed. This constant movement within a discourse community is necessary for its survival, as it helps to preserve the balance between novices and expert members.

In medical discourse there are different types of communication, depending on who is communicating with whom (Krois-Lindner 2007). One type is inter-physician communication; doctors communicate with each other at conferences or via written journal articles (ibid.). Doctors communicate with nurses and other personnel and vice versa. This communication plays only a minor role in the eyes of some

researchers as it seems to be restricted to situations such as changing shifts or talking about patient facts and figures (Hoejke 2007: 8). Communication with other hospital personnel (e.g. secretaries) tends to be ignored. Pathologists, for example, record autopsy reports; secretaries then type these up. Communication with (medical) personnel can be is especially problematic for non-native medical residents. This will be explored later in more detail[15].

Another type of communication occurs between medical personnel and patients. Whereas several studies have focused on doctor-patient communication, comparatively little research has been done on the topic of nurse-patient communication. Patients are often intimidated and do not ask questions in the course of a doctor-patient consultation. They sometimes turn to nurses to ask for clarification, or for emotional support[16]. In the literature, however, doctor-patient communication has been the predominant topic of research, as opposed to nurse-patient communication. Genres within this field are doctor-patient consultations, patient education sheets, letters etc. (Krois-Lindner 2007). Lastly, medical personnel also often communicate with the wider public via health care websites or in other forms. Patient information sheets fall in this category, as they are made available online (ibid.).

Types of communication	***Genres***
Doctor – doctor *(inter-physician communication)*	Journal articles, abstracts, reports, referral and reply letters, standardised documentations
Doctor – patient	Doctor-patient consultations (interview, physical examination), patient information sheets
Doctor – public	Websites, patient information material

Table 6. Types of communication (Krois-Lindner 2007)

[15] For more information see section 4.2.
[16] According to my specialist informants.

The genre of patient information sheets

This medical genre is a relatively new one that has only evolved during the last few years (Krois-Lindner 2007). The rise of the genre emerged through developments in different but correlating disciplines such as sciences, ethics and law. Patient information sheets (PIS) are also called patient information leaflets, patient education materials or health fact sheets. All of these terms describe the same concept, which is the idea of offering patients information. This information can concern diseases, diagnostic procedures, types of treatment, medication or studies (Krois-Lindner 2007). The term 'patient information sheet' is sometimes wrongly used to refer to 'patient data sheets'. The latter are given to patients when admitted to hospital, in order to obtain personal data, i.e. first name, surname, address and so on[17].

Patient education varies from situation to situation, since the settings, goals and points of view are situation-specific (Anderson et al. 191, Anderson 1995, Leino-Kilpi 1999 in Johansson 2005:213). One main aim is to "empower patients by means of education, to help them make decisions about their care and manage with their own health and care situation" (Anderson et al. 1991, Anderson 1995, Leino-Kilpi 1999 in Johansson 2005:213). Patient education and patient empowerment are essential in contemporary medicine. Patient information materials are a means of accomplishing this empowerment (Johansson 2005: 220).

As far as the medical genre of patient information sheets is concerned, it is important to view it in combination with other genres from the discipline, since they do not exist in isolation. They are only one of numerous genres of medical discourse (Pettinari 1988: 12ff). Written genres include patient information material and medical history, but also reports, letters, articles and abstracts. Speech genres in medicine, for example, include doctor-patient consultations or pathology reports (Krois-Lindner 2007). In the following example I would like to illustrate the genres that are brought together in the overall genre of medical record. It is obvious that there is a chronological order – i.e. one genre after another is added to the record. A medical record is comprised of the following information:

[17] According to my research.

(1) Identification data (such as the patient's name, address, date of birth etc.)
(2) Medical history
(3) Report of the physical examination (can be combined with no. 2)
(4) Diagnostic and therapeutic orders
(5) Evidence of appropriate informed consent
(6) Clinical observation (progress notes, consultations, nursing notes, post-anaesthesia recovery report)
(7) Reports of procedures, tests and results (operative report, pathology reports, clinical laboratory reports, radiology and nuclear medicine examinations or treatment, anaesthesia reports, organ donor reports)
(8) Conclusion at the end of hospitalisation (discharge summary)
(Pettinari 1988:12-15)

This list was compiled by Pettinari in 1988, and it mirrors the common practices in the U.S. at that time. Although much time has passed since then, it is still valid. One thing that becomes obvious is that patient information and consent forms already played important parts in 1988. In Austria medical records are comprised of the same components.[18] Nowadays, special emphasis is given to informed consent and consent forms. Whenever an operation is advisable, doctors talk to their patients in a preoperative consultation. Here, procedures and risks etc. are discussed. Patient information sheets are often given to the patient in order to convey a clearer picture. Consultations are usually terminated with the patient agreeing to the operation or treatment by signing the consent form. Since Austria is a multi-cultural society, patient information materials in foreign languages are essential. As interpreters are not always available, the doctor equipped with patient information sheets in the relevant language can guide the patient through the consultation[19].

So far I have dealt with the challenges that medicine as a discipline and also doctors have had to face. In addition, an introduction to medical discourse and the genre of PIS was offered. I would now like to turn to past studies that addressed these topics. In the following section a brief overview of previous studies on patient information sheets and on medical communication (for language learners) will be given.

[18] According to my specialist informants.
[19] According to my specialist informants.

4.2 Previous studies

Patient information sheets

Growing numbers of studies on patient education mirror the interest, need and necessity of research in this field. Previous studies on patient education (material) investigated content, readability and design and identified deficiencies of education material, as the language employed was either too difficult or the design lacked readability (Chumbley et al 2002: 460). Other studies have focused on the effects of patient education, the implications being that patient education is instrumental in assisting patients to cope with their situations, for instance, by reducing anxieties, strengthening autonomy and positively influencing patients' dedication and serious-mindedness regarding post-operative proceedings (Johannson et al. 2005: 213). Among those conducting studies are Chumley et al. (2002) and Johansson et al. (2005). Although their studies do not primarily concern the field of oncology, they are of relevance since they address more general issues such as the involvement of patients in the designing process of patient information sheets (PIS) (Chumley et al. 2002). Chumbley et al. (2002) explored the readability of an existing patient information leaflet on patient-controlled analgesia (PCA), a method of postoperative pain control where patients administer and control their own dosage of medicine. In the introductory section of the study the authors offer a brief summary of what has been published on the topics so far. According to Chumbley et al. (2002: 460), "(p)atient education is usually delivered through bedside teaching or patient information leaflets". The problem with this is that medical professionals or product manufacturers design education material for laypeople and these professionals are not necessarily fully aware of the actual value ascribed to it by laypeople (ibid.). The aim of the quantitative study was to evaluate an existing leaflet in order to formulate a new one, and finally to compare it to the 'old' PCA information leaflet. Chumley et al. (2002) employed the method of eliciting patients' views in order to assess their perceptions, and thus derive what patients think is essential to understand (461). The findings suggest there was confusion about side-effects and use of PCA among the patients that read the 'old' leaflet (463). Shortcomings of the 'old' leaflet include the omission of information about morphine, which is used in PCA; another fault was that alternative methods to PCA were not mentioned at all. The new leaflet, on the other hand, included the information that patients wanted, as well as numerous illustrations that helped to facilitate understanding (Chumley et al. 2002: 467).

Another interesting study was published by Johansson et al. (2005). In *Preoperative Education for Orthopaedic Patients: Systematic Review* the authors undertake literature reviews as well as meta-analyses in order to explore the effects of patient education. The results indicate that written material such as booklets are powerful tools, as patients are always able to return to them (Johansson 2005: 220). However, according to the authors, readability is the key to success, in the sense that only patients who understand the content are empowered to manage their own healthcare. More importantly, it is explicitly stressed that written patient education material should be regarded as a complementary means of empowering and educating patients. By no means do written booklets or leaflets replace the need for face-to-face consultations between medical professionals and their patients (Johansson et al. 2005: 220).

Medical students, just like any other novices, have to be made familiar with the conventions, values and genres of the discipline (Hoejke 2007: 6). Genres they are likely to come across in their working life include medical reports, forms for taking a medical history, face-to-face consultations and PIS (Krois-Lindner 2007). How students ought to become familiar with medical communication in general has been the subject of numerous studies. These will be explored in the following section.

(Intercultural) medical communication

(Intercultural) medical communication has attracted attention from various disciplines, including communication studies, nursing, medicine, health education, and (more recently) the field of English of Specific Purposes (ESP) (Frank 2000: 31). Studies conducted in the field of ESP have focused on (a) medical professionals who seek to work in an English-speaking country; (b) medical professionals who would like to make publications in the mainly English-dominated research and publishing world or who wish to participate in international conferences; (c) international students who would like to study medicine in an English-speaking country (Frank 2000: 31). Among these studies, there are those conducted by Eggly et al. (1999) and Hoekje (2007). A study focusing on international medical graduates (in short, IMGs) (Hoejke 2007: 4), was undertaken by Eggly et al. in 1999 at the Wayne State University. The relationship between English language skills and success as medical

residents was investigated. The findings indicate that language proficiency correlates with patient satisfaction (Eggly et al. 1999: 2ff).

Another study dealing with IMGs was published by Hoekje (2007). In *Medical Discourse and ESP Courses for International Medical Graduates (IMGs)* she investigates the situation in the United States. According to her (2007: 2), international medical professionals constitute 23% of the physicians in the United States. Despite subject-specific core competencies such as medical knowledge, patient care, communication skills, professionalism, sensitivity to population diversity and the role of the health system, IMGs have to "acclimatize to the U.S.; learning cultural norms and language of the patient communit(ies) where they are practising; improving intelligibility where needed; and learning the components of professionalism in the U.S. context" (Hoejke 2007: 2). Hoejke (2007) developed various courses for different departments, such as psychiatry and pathology. Among the courses were also 'acculturation courses' for IMGs. In order to get started, a needs analysis was undertaken. In the courses of the needs analysis IMGs, their supervisors and training doctors were interviewed or asked to write reports (9-10). Additionally, observations and video- and tape-recordings were undertaken.

A situational syllabus was developed, which took in consideration spoken and written genres from medical discourse and "the wider U.S. cultural context" (Hoejke 2007: 2). The various course goals of the acculturation classes included familiarising aspiring doctors with the language of patients. This language of the patient community includes slang expressions for body parts, illnesses and symptoms, drugs, behaviours etc., as well as child language ("pee pee") and euphemisms ("I have a problem with my thingy") (Hoejke 2007: 9-11). Other goals included familiarising the doctors with the cultural values of the patients and the language of the institution. Hoejke's (2007) other courses emphasised other aspects, including intelligibility. The genre of autopsy reports, for example, is essential in pathology. Autopsy reports are usually recorded by doctors and then typed up by secretaries. Vocabulary and pronunciation (including stress and intonation) play significant roles (Hoejke 2007: 11). Hoejke's (2007) study provides many valuable insights and should be taken into account when designing courses for medical students.

Another study was carried out by Nestel & Kidd (2004). In *Teaching and Learning about Written Communications in a United Kingdom Medical School*, the authors

describe a two-hour session on written communication that they developed (and which was evaluated later by tutors and students). What is worth highlighting is the fact that abbreviations and acronyms were given a central role (Nestel & Kidd 2004: 29). As mentioned in section 4.1 acronyms and abbreviations can be especially confusing and misleading to outsiders of a discourse community (Swales 1990). Although Nestel & Kidd (2004) did not explicitly focus on international students, they made a contribution towards a better understanding of what medical students need to learn and become familiar with. In spite of the fact that not all medical professionals in Austria necessarily seek a career in an English-speaking country, one has to acknowledge the contributions that these studies and course descriptions have made. The results are invaluable in terms of designing syllabi for medical professionals in general since they (a) highlight the importance of language competence in regard to patient satisfaction and (b) stress the importance of acronyms, abbreviations and 'patient talk' in medical discourse.

4.3 The educational setting

In Austria pupils and students interested in medicine, health care and nursing can choose from a range of different school types and universities. Instead of pursuing an academic-oriented career at upper-secondary schools, pupils from Year 9 onwards can decide to attend a VET school (Vocational Education and Training), in this specific case, a school of social occupations (Schule für Sozialberufe) or an educational institution for nursing personnel (Schule für Gesundheits- und Krankenpflege). The main aim of VET schools is to offer initial vocational education and training together with a more in-depth general education. The educational objectives of VET schools are manifold depending on the individual school. In the case of schools for social occupations, the aims are to equip learners with the right tools that can enable them to specialise in the social field at a later point in their lives, and/or to train them in health care and nursing. Schools for social occupations also run part-time courses for people in employment. Mandatory internships during the school years uphold the most significant roles of this school type. Generally speaking, the vocational training in schools of social occupations can focus on the care for the elderly, or the work with families and disabled people (including health care and nursing) (Austrian Federal Ministry for Education, the Arts and Culture 2007a). This emphasis is also mirrored in the educational objectives of English as a

school subject. First and foremost, learners at schools of social occupations should be able to practise reading, writing, listening and speaking skills. Secondly, they should be able to have conversations related to health care and nursing. Other educational aims include autonomy and international communication. In terms of topics, the English language curriculum just offers key words such as family, youth, old age, disability, social problems or health, without necessarily narrowing the topics down (Austrian Federal Ministry for Education, the Arts and Culture 2007b). Therefore, a lot depends on the individual language teacher to interpret and fulfil the educational objectives.

Institutions for nursing personnel also fall into the category of VET schools, although there is an age limit for those wanting to attend the school, so it is only possible to apply from Year 11 onwards. The focus of individual schools differs, as some specialise in general care whereas others specialise in paediatrics or psychiatry (Austrian Federal Ministry of Health, Family and Youth, 2008). Theoretical education is provided alongside practical education in institutions for nursing personnel, thus Medical English is regarded as an integral part of the curriculum. The English education course covers three modules: a 40 lesson program in Year 1, and a 20 lesson programs in Years 2 and 3. Aspiring nursing personnel become familiar with subject-specific vocabulary and important language functions such as greeting, introducing oneself, explaining, narrating, seeking information, describing etc. In the English curriculum, prominence is given to general language as well as specialised language. In Year 1 the focus is on everyday situations at the future workplace. Emphasis is put on the enablement of conversational skills, since language functions such as greeting, introducing oneself, narrating and seeking information about the patient are important language functions. Slowly but gradually learners become familiarised with the specific field of health care and nursing, e.g. with the genre of admission interview or vocabulary concerning the human body, bodily functions, hospital wards and departments, medical personnel etc. During Years 1-3, the emphasis on learning and teaching subject-specific and related vocabulary becomes stronger. Language functions such as describing and informing are embedded in a very specific context, i.e. learners should be able to inform others about instruments, equipment and their functions (ÖBIG 2003). An overview of the English curriculum in educational institutions for nursing personnel is presented on the next page.

The curriculum has been summarised and translated into English.

Year	*Aims*	*Topics*
Year 1 **Module of 40 lessons** **Focus:** everyday situations, communication (health care)	▪ Greeting & Introducing Opening, having & terminating a conversation ▪ Changing the topic	Ways of greeting, introducing Having a conversation (opening and terminating)
	▪ Using expressions connected to time and intervals	Date, months, days of the week, time, intervals, etc.
	▪ Learning subject-specific terminology	The human body, functions, daily activities (describing the way, etc.), hospital wards, medical personnel and hierarchies
	▪ Hospital admission (Admission interview):	Admission (greeting, asking for, supplementing confirming information)
	▪ Explaining and finding out about everyday occurrences	Everyday occurrences
	▪ Informing about intervention, actions	Interventions, actions
Year 2 **Module of 20 lessons** **Focus:** Vocabulary (nursing)	▪ Translating expressions from everyday (nursing) life ▪ Applying vocabulary	Diseases: Aetiology, symptoms, diagnosis, therapy, prognosis, complications
	▪ Naming instruments and equipment ▪ Translating manuals ▪ Informing about the function and application of instruments and equipment	Instruments and tools/ devices ▪ In Nursing and Medicine ▪ Function ▪ Application ▪ Maintenance
	▪ Translating expressions from nursing process ▪ Using vocabulary ▪ Writing nursing reports	Nursing process Anamnesis, diagnosis, planning, intervention Documentation and evaluation
	▪ Translating expressions from First Aid ▪ Applying vocabulary ▪ Writing reports (accident, emergency)	First Aid Chain of survival (emergency call, wound management, etc.)
Year 3 **Module of 20 lessons** **Focus:** Vocabulary (health care system)	▪ Translating expressions concerning health care system	Health care system(s) (National and international)
	▪ Understanding the differences of a concept in context of different health care systems	
	▪ Discussing texts, articles etc.	Current topics: old age, physically and mentally handicapped people, euthanasia, alternative medicine, nursing research etc.
	▪ Working with (summarising) subject-specific texts	Sources (journals, books, Internet, etc.)

Table 7. Educational institution for nursing personnel: English curriculum (ÖBIG 2003)

At the tertiary level, people interested in medicine, health care and nursing can choose from a range of institutions, including medical universities or universities of applied sciences (Fachhochschulen). I have chosen to focus on medical universities, since it would be impossible to explore the English curriculum of every single type of tertiary institution. Several years ago, English language skills only played a minor role in the curricula of the various medical studies. Two courses on Medical English, introduced in 2006, can now be taken at the Medical University of Vienna. The first course focuses on equipping students with the necessary skills and knowledge to undertake research and studies, with prominence given to reading skills and the genres of abstract and report etc. The second course is of more interest to my study, and is aimed at students who intend to do their internship abroad (Medical University of Vienna 2008). Presented below is an excerpt of the course description:

> [...] Through intensive oral and written work related to medical treatment, this course prepares students for medical internships abroad. Based on listening and reading comprehension practices and on role-play and discussion activities, students will become familiar with the commonest expressions in medical diagnosis and treatment. The course focuses on doctor-patient communication as well as on oral and written professional interaction amongst medical staff. (Medical University of Vienna 2008)

As this course is focused on students who intend to do their medical internship abroad, emphasis is placed on different medical genres, with special focus on doctor-patient communication and consultation. Role-plays and other settings are used to simulate real-life situations.

In conclusion, it can be said that English plays an integral part in the education of health professionals. English language skills are considered essential for the studies of medicine and nursing as the majority of research articles are published in English, and also as English language skills are necessary professional assets in a multi-cultural society.

4.4 Research concept

Corpus-based genre analysis is a widely-used method to obtain data for teaching genre, especially so in EOP (English for occupational purposes). Those who have worked on the subject (Gavioli 2005, Flowerdew L. 2005) conclude that genre-based corpus studies in EOP provide invaluable findings for teachers and learners. Furthermore, specialised corpora are not only acknowledged as teaching tools, but also as learning tools. So far, genre-based studies in EOP have focused on academic genres such as research articles (Swales 1990) and abstracts, or business English, including, for instance, letters of application (Bhatia 1993, Henry & Roseberry 2001, Upton & Connor 2001). Legal English has also attracted attention, in particular legal propositions (Bhatia 1993) and consent statements (Mungra 2007). However, the range of (written) medical genres has largely been ignored. While the spoken genre of patient-doctor consultation/talk has become increasingly attractive for researchers from different disciplines, written genres such as medical history, pathology reports and patient information materials have been neglected. There are only a few exceptions: Pettinari (1988), for instance, investigated operative reports. However, there is little research available on the genre (analysis) of patient information sheets (PIS). For this purpose I am aiming in this study to explore PIS, highlighting interesting key features of the genre that are worth identifying for learners and teachers of EMP. The aim is to investigate the range of vocabulary used in PIS, since PIS cover vocabulary from the general, technical and academic language. In addition, I will investigate what types of multi-word units are used in the genre texts (Chapter Five). In the main body of the paper, I am more concerned with examining the communicative purpose of PIS – an objective that is achieved by analysing the so-called move structure of the genre. Special emphasis is also put on common language functions that are integrated in the individual moves (Chapter Six). As has been mentioned in previous sections, the English curriculum in institutions for medical personnel and medical universities often centres on language functions (such as defining and describing). I will also attempt to investigate what language functions are found in PIS and how these are realised in PIS. A detailed account of the procedures and methods is provided in the following section.

4.5 Methodology

Data collection

The data for this quantitative study was drawn from the medical genre patient information sheets (PIS). The materials were obtained from governmental or acknowledged healthcare websites from Great Britain, Australia, Canada, New Zealand and the United States. It has to be pointed out that the institutions (listed below) have produced their own patient information sheets (PIS). In Austria, by contrast, there is a tendency to 'buy in' standardised patient information sheets from the private sector. In other words, companies sell patient information sheets to individual doctors, hospitals or associations of hospitals. The patient information sheets in Austria are also available in various languages, including German, English, Polish and Turkish[20].

The patient information sheets used for this study were provided by the following institutions:

a) American Academy of Family Physicians
b) The Better Health Channel (Government of Victoria)
c) Cancer Society of New Zealand
d) Health Canada
e) National Health Service (NHS) CKS Clinical Knowledge Summaries (Great Britain)

In terms of inclusion or exclusion criteria, it can be said that first of all, reliable sources of patient information sheets (PIS) were sought, and secondly, the content or topic range itself was intentionally restricted. In order to simplify the analysing process and to facilitate the generation of specific knowledge I only selected patient information sheets (PIS) focused on the topic of cancer: more specifically, with the topics of breast, prostate, bowel (colorectal cancer) and skin cancer (melanoma). The relevant patient information sheets concentrate on information such as diseases, conditions, symptoms, diagnostic procedures and/or types of treatment. Patient information sheets (PIS) on cancer research studies and cancer medication were neglected. Table 8 summarises the criteria that were used to design the corpus.

[20] According to my specialist informants.

	Criteria
Corpus size	> 10.000
Number of texts	20
Full texts	Full texts
Medium	Written (patient information sheets/leaflets, health sheets, …)
Subject	Cancer (breast, bowel/ colorectal, prostate and skin cancer)
Authorship	PIS written by acknowledged experts
Language	Texts written in English by native speakers
Publication date	Recent years

Table 8. List of criteria (framework as suggested by Bowker & Pearson 2002: 72)

A sum total of 20 patient information sheets were compiled from five patient information sheets/leaflets providers. From each provider, one PIS on the topic of breast cancer, bowel (colorectal) cancer, prostate cancer and skin cancer (melanoma) was selected. In terms of topics, the most common cancer types are covered.
The overall number of tokens, meaning running words, in the corpus is 21537.
The average text length is thus 1076 words, although this varies according to topic. Interestingly, PIS on the topic of breast cancer exceed others in text size.

	Breast	***Bowel***	***Prostate***	***Skin***
Average text size	1549	925	902	929

Table 9. Average text size per topic

Methods and procedures

For this study I chose a statistical approach. The data processing was comprised of different stages, ranging from computer-based analyses of the vocabulary to the manual analysis of the structure of patient information sheets. At the initial stage, the patient information sheets were saved in .txt format in order to make them accessible for the computer programs. Afterwards, differences in spelling between American English and British English were resolved in order to guarantee that word count data was not falsified. As a final step, bullet points and freestanding hyphens were deleted, and the formatting was simplified as not all of the programs are able to read this type of data.

The relevant computer programs were RANGE and Oxford WordSmith Tools 4.0. Both program suites are acknowledged in linguistics. The data was accessed with the aid of these programs.

Software RANGE

First, the program RANGE was employed, which consists of two programs, RANGE and FREQUENCY. RANGE offers ways and means of generating and comparing vocabulary lists and or even comparing a specific text against vocabulary lists. RANGE is comprised of three basic word lists. The first consists of the first thousand commonest words in English, while the second lists the second thousand commonest words. The third base word list offers words that are used in an upper secondary school or tertiary education setting, and are drawn from various subjects. Essentially, two major sources were used to generate these three base word lists. These sources were *A General Service List of English Words* (*GSL*) by Michael West (1953) and *The Academic Word List* (*AWL*) by Coxhead (2000). The *GSL* served as the basis for base word lists one and two. Developed in 1953, the *GSL* can seem outdated, but it is still considered reliable and is therefore still used. It consists of 2,000 headwords that are taken from the general language (West 1953). The *AWL* was developed in 2000 by Averil Coxhead. The list comprises 570 word families and does not contain any words featured in the *GSL*.

The other program, FREQUENCY, enables the user to generate vocabulary lists ordered according to number of occurrences. It offers word type and family rankings, numbers of occurrences and cumulative percentages (Heatley, Nation & Coxhead 2005).

Software WordSmith Tools Version 4

WordSmith was used as a complementary tool. This computer software allows users to generate wordlists according to frequency and distribution, to identify key words and to make concordances on the basis of compiled texts. The WordList program of WordSmith was used to obtain a list of the words in the corpus, listed in descending order according to their frequency. Another program in the WordSmith suite is the concordance program, which was also used in this study. It is useful for identifying common clusters – in other words, showing how words are used (Scott 2004).

The computer-based analysis tools WordSmith and RANGE were used to access the data and to filter relevant information regarding high- and low-frequency vocabulary, technical and academic vocabulary, and information regarding co-occurrences of words.

The reference corpus

In order to compare data a reference corpus is needed. The comparison corpus in this study was the BNC corpus. It consists of 100 million words and comprises spoken and written English. Written texts – newspaper extracts, journals, periodicals, academic books, fiction, letters, essays and so on – make up about 90 % of the corpus. One might criticize the fact that the texts date back to the 1980s (or even 1960s). However, it nevertheless served as a reference corpus in this study since it is one of the biggest and most renowned corpora and also because written texts are featured so prominently (*British National Corpus*). Furthermore, it is most commonly employed in genre-based analyses in ESP.

Move structure analysis

A manual analysis of the move structure based on Swales' (1990) genre moves was also undertaken. Moves are identifiable units that contribute to the overall communicative purpose of a text (Bhatia 1993: 30). Identification of moves and the sub-types, known as steps, depends on the researcher and what he/she regards as realisations of underlying communicative functions. Swales' 'CARS model - Create A Research Space model' (1990) was introduced previously, and is now illustrated below. According to Swales (1990), a research article introduction is comprised of three moves, which can consist of various steps:

Move 1: Establishing a territory
- Step 1: Claiming centrality and/or
- Step 2: Making topic generalisation(s) and/or
- Step 3: Reviewing items of previous research

Move 2: Establishing a niche
- Step 1: A Counter-claiming or
- Step 1: B Indicating a gap or
- Step 1: C Question-raising or

Step 1: D Continuing a tradition

Move 3: Occupying the niche

Step 1: A Outlining a purpose or

Step 1: B Announcing present research

Step 2: Announcing principal findings

Step 3: Indicating research article structure

(Swales 1990: 141)

In this paper, Swales'(1990) analysis served as a framework, since his work on research articles has been seen as pioneering (Dudley-Evans 1994, Flowerdew L. 2005). This approach was chosen because analysing the move structure of a genre allows researchers to identify the basic structure and communicative purposes. I attempted to follow Swales' (1990) model of analysis and even his way of labelling moves and steps, using verb rather than noun phrases, as it has certain advantages. Verb phrases denote an active process and can enable a researcher to link the relevant move to the function of the move.

In addition, I attempted a categorisation of moves, in terms of optional, core and obligatory moves. Exemplars of genre texts are comprised of moves that are more or less frequent in the genre. In order to arrive at a representative structural interpretation, it was necessary to distinguish moves according to their frequency. Moves that are featured in 50% of the compiled genre texts can be regarded as optional (Henry & Roseberry 2001). Core moves are those occurring in 50-90% of all exemplars. Moves that, on the other hand, appear in 90%-100% of the compiled genre texts should be seen as obligatory (Hüttner 2007).

So far, I have dealt with Bhatia's (1993) first five steps of analysing genre. To recapitulate, these were (1) 'Placing the given genre-text in a situational context' (section 4.1), (2) 'Surveying existing literature' (section 4.2), (3) 'Refining the situational/contextual analysis' (section 4.3), (4) 'Selecting the corpus' (section 4.5) and (5) 'Studying the institutional context' (section 4.3) (Bhatia 1993: 22-36). Now, I would like to turn to Bhatia's (1993) next step of analysing genre – the linguistic analysis. In the following two chapters, the focus will be on vocabulary and the move structure of PIS.

5 Vocabulary in patient information sheets

5.1 Theoretical background

Vocabulary is a matter of major concern to language learners and teachers, and this is also true in the field of ESP. When English is taught for medical purposes, course materials often do not meet the needs of language learners. Computer-assisted corpus analysis can facilitate findings such as vocabulary and grammar of fields such as ESP and also of particular genres. This can provide valuable results for course book and task design. The genre of patient information sheets (PIS) is both interesting and important, since highly technical matters are translated into a language that is accessible for laypeople. Doctors, in fact, translate technical information for patients during doctor-patient consultations. It is this similarity that makes patient information sheets highly attractive for learners, teachers and researchers. PIS simply offer and highlight a range of useful vocabulary for doctor-patient communications (Krois-Lindner 2007).

In this study an analysis of vocabulary based on Nation's (2001a) four types of vocabulary was conducted. Nation's (2001a) developed categorisation of vocabulary into four types (high-frequency words, academic words, technical words and low-frequency words) has played and continues to play a significant role in vocabulary research (Nation 2001a: 11). The first type is high-frequency words, which occur in a wide range of language uses, and thus constitute a large proportion of a running (spoken or written) text. In most cases high-frequency words are function words, but they can also be content words (Nation, 2001a: 11). The second type is academic words, which are most often found in academic texts, but only constitute 9% of the running words of an academic text, whereas high-frequency words account for 80% (11-12). Thirdly, there are technical words. These range from subject-specific words that appear to be of high frequency to words that are characterised by having specialised meanings. Technical words constitute 5% of an academic text (Nation 2001a: 18). Finally, the fourth type of vocabulary is low-frequency words. Low-frequency words typically cover about 5% of a running text. In other words, low-frequency words are words 'of moderate frequency'.

So far, the focus has been on the different vocabulary types proposed by Nation (2001a). As far as vocabulary analysis is concerned, it is possible to argue that computer programs can facilitate and improve the description of vocabulary. These software tools enable researchers to analyse vocabulary according to vocabulary range and frequency (Nation 2001b: 33). In this study the vocabulary analysis software RANGE[21] was used, which was developed by Heatley, Nation & Coxhead (2002). In the following sub-chapters an overview regarding the vocabulary size of the corpus will be provided, before I turn to the outcomes of the analyses of frequency and range.

5.2 The vocabulary size of the corpus

Information regarding text and vocabulary size, as well as the distribution of words, is very important when exploring corpora. Text and vocabulary size are derived by counting either tokens, word types and/or word families. Tokens simply denote all the running words of a spoken or written text. If a specific word form occurs more often, these occurrences are counted as well. A word type is a single word form. When counting types, multiple occurrences are not taken into consideration. Not only can words be counted in the forms of tokens and types, they can be counted in word families. "A word family consists of a headword, its inflected forms, and its closely related derived forms" (Nation 2001a: 8). Derived forms cover words including affixes such as *–ly*, *-ness, -ment* or prefixes like *un-*. A word family therefore contains a base word and formally and semantically related word types (Chung & Nation 2003: 107). The main difficulty in counting word families is to decide what should and should not belong to a word family (Nation 2001a: 8).

The PIS corpus amounts to 21537 tokens or 2434 word types. A total of 1028 word families were identified.

Corpus	***Tokens***	***Types***	***Families***
PIS Corpus	21537	2434	1028 identified

Table 10. Text size and vocabulary size

[21] For more details on the program see chapter 4.

5.3 Vocabulary analysis according to frequency

Word frequency data is a very powerful and useful tool in the hands of teachers and course book designers (Flowerdew J. 2001: 75). It allows those who are interested to derive the relative frequency of vocabulary, that is the "inverse relation between frequency of occurrence and number of items occurring with a given frequency" (Flowerdew J. 2001: 75). Therefore, there are 2272 items that appear less than 20 times in the corpus. A considerably small number (10), on the other hand, occur over 300 times.

Frequency range	***PIS corpus***
> 1001 times	1
501 to 1000 times	4
301 to 500 times	5
101 to 300 times	22
51 to 100 times	29
21 to 50 times	101
1 to 20 times	2272

Table 11. The relative frequency of items: the distribution of words (word types)

Frequent and less frequent items will be dealt with in the following sub-chapters.
In the present study the expressions 'frequently occurring items' and 'less frequent items' are preferred to 'high-frequency' and 'low-frequency words' as far the frequency of specific words in the corpus is concerned. This can be explained by the fact that Nation (2001a, 2001b) uses the label 'high-frequency words' for words that "occur in a wide range of language uses" (Nation 2001b: 32). These words are also called 'general service words' (West 1953), and are not necessarily co-extensive with the most frequent PIS words. A different term is used for the latter in this paper.

5.3.1 Frequently occurring items

As the name implies, frequently occurring items are characterised by their high frequency. In most cases grammar words (rather than lexical words) constitute a high percentage of the most frequently occurring items (Hunston 2002: 3, Flowerdew J. 2001: 77). That is also the case in this corpus study. As can be seen in table 12, 'the'

and ‘of’ top the word frequency list of the specialised corpus and the comparison corpus.

Rank	*PIS corpus*	*BNC corpus*
1	THE	THE
2	OF	OF
3	CANCER	AND
4	TO	TO
5	AND	A
6	A	IN
7	IS	#
8	IN	THAT
9	YOUR	IS
10	YOU	IT

Rank	*PIS corpus*	*BNC corpus*
11	OR	FOR
12	ARE	WAS
13	FOR	I
14	BREAST	ON
15	BE	WITH
16	HAVE	AS
17	IT	BE
18	WITH	HE
19	CAN	YOU
20	PROSTATE	AT

Table 12. Word frequency comparison across corpora

Table 12 shows the twenty most frequent lexical items in the PIS corpus and the BNC corpus. In the top five of lexical items there is a striking similarity between the corpora as four items are featured in both (the item *cancer* being the only exception). The latter is featured throughout PIS, and so it therefore comes as no surprise to observe that it is ranked third in the word frequency list of the specialised corpus, with 593 occurrences (a rate of 27.53 per 1000 words).

More information regarding discrepancies between the corpora can be retrieved from the items ranked sixth onwards. The word family *you*, with the types *your* and *you* (with rates of 15.13 and 14.39 per 1000 words as opposed to 1.34 and 5.88 per 1000 words in the BNC corpus) play an important part in the PIS corpus. The personal as well as the possessive pronoun seems to be used deliberately, for several reasons. It can be assumed that the primary aim behind the use of *you* and *your* is to personalise the genre text, by addressing the reader (the patient) directly and thus establishing a relationship between author and reader. In doing this, the patient is able to feel responsible for her/his wellbeing. Other personal pronouns such as *I*, *he* and *you* are frequently-occurring lexical items in the general corpus. However, they appear to be of no importance in the specialised corpus as they hardly occur.

Was, for example, has a prominent position in the BNC corpus – with a rate of 8.63 per 1000 words – but in the specialised corpus it is ranked 408th, with only eight occurrences (a rate of 0.37 per 1000 words). Possible implications for learning and teaching might be that past tenses seem to play only a minor role in the genre of PIS.

Another frequent item in the PIS corpus is *can* (ranked 19th), as it occurs 170 times (a rate of 7.89 per 1000 words, as opposed to a rate of 2.11 in the BNC corpus). A similar observation can be made regarding the modal verb *may*, which accounts for a rate of 6.87 per 1000 words in the specialised corpus but only has a rate of 1.27 per 1000 words in the general corpus. One possible implication might be that modal verbs play an important role in expressing possibility, tentativeness and indirectness. It has to be noted that *can* and *may* outperform other modal verbs.

As apparent from table 12, the most common words in PIS are grammatical or so-called function words. I will now turn my attention to content words and more specifically to nouns. The twenty most common nouns of the specialised corpus and general corpus are presented below:

Top twenty nouns in the BNC corpus
Time, people, way, work, years, pound, year, Mr., day, government, man, life, world, part, home, house, number, end, place, case
Top twenty nouns in the PIS corpus
Cancer, breast, prostate, treatment, risk, bowel, skin, men, doctor, melanoma, age, women, body, screening, symptoms, cells, surgery, factors, people, test

Table 13. Top 20 nouns (Flowerdew J. 2001: 78)

The table illustrates the considerable differences between the general and the specific corpus and thus outlines differences between the general and the subject-specific language. The only noun to appear in both lists is *people* (and also *man* in its plural and singular forms). This comparison of corpora underlines the importance of specialised corpora in the field of ESP. One major advantage of the use of a specialised corpus for learning and teaching is that it provides valuable information regarding frequently occurring word types and their particular uses (78).

5.3.2 Less frequent items and single-occurrence items

Computerized word counts cover a range of items, including frequent and less frequent items. The latter are words that are "conspicuous by their low frequency or absence" (Flowerdew J. 2001: 78). These words can offer valuable and powerful information regarding a genre, its style and lexis, information that is of great value to teachers and learners. As the initial comparison of the PIS corpus to the BNC corpus showed, personal pronouns seem to play a minor role in PIS. Whereas the pronoun *you* is used to address the reader, other pronouns occur very rarely. *They* and *I* occupy are ranked 50th and 151st, as they occur 62 and 22 times. *We*, on the other hand, is used only 7 times (ranked 466th in the specialised corpus), *she* once (ranked 2254th), while *he* is not featured in the corpus at all. Among the less frequent items is also *was* (ranked 408th with 8 occurrences), *were* (ranked 373rd with 9 occurrences), and *had* (ranked 178th with 19 occurrences). Only *been* occurs more than twenty times, as it is ranked 135th with 25 occurrences. This can be explained by the fact that authors of PIS describe diseases, body parts, procedures etc., and thus tend to use the present tense.

In the specialised corpus there are 2272 items that occur less than twenty times.
The table below shows the distribution of these words in the PIS corpus:

Frequency range	***PIS corpus***
11-20 times	131
6 – 10 times	245
5 times	106
4 times	136
3 times	232
2 times	416
1 time	1006

Table 14. Distribution of less frequent items

These numbers suggest that a learner working with this corpus will encounter a considerably high number of words only a few times or even only once. The same is of course true for patients. The latter will face certain words only once. By omitting single-occurrence items the number of target lexical items can be reduced

dramatically. However, in the case of this study elimination would not be possible since subject-specific terms also occur only once, and these are of relevance to the learners. Broadly speaking, corpus design has an enormous impact on the results of computer-based frequency counts. If a lexical item occurs only once then it does not automatically imply that it is not worth teaching (Flowerdew, J. 2001: 78). Therefore, a lot depends on the individual researcher and his/her background knowledge.

5.4 Vocabulary analysis according to range

The corpus contains 21537 tokens or 2434 types. Out of these 2434 types, 1392 are also featured in *General Service List,* also referred to as *GSL* (West 1953), the list of the 2000 most common words in the English language. 299 word types were identified as academic words. Unlike these word types, 743 were not featured in any of the base word lists that come with the program. These can be regarded as technical and low-frequency words taken from the general language.

Vocabulary level	***Number of word types (word families)***
1st 1000 most common words	1052 (596 out of 998 word families)
2nd 1000 most common words	340 (225 out of 998 word families)
AWL	299 (207 out of 570 word families)
Not in the list	743 (??) of which 545 are technical words and 198 are low-frequency words from the general language
Total types	2434 (1028 identified word families)

Table 15. Types of vocabulary in the PIS corpus

However, another way to present these results is to show how much is covered by general, academic and technical language. Around 65% of the vocabulary in PIS is taken from the general language (together with the low-frequency words that were not categorised by the program), whereas only around 12% can be regarded as academic vocabulary. Over 22% of the vocabulary in PIS falls into the category of technical words. As the topic of the PIS was restricted to four cancer types only, it is

possible to imagine that the number and the percentage of technical vocabulary would be considerably higher if material on different cancer types had been selected.

Vocabulary	***No of word types***	***Percentage of coverage***
General language[22]	1590	65,32 %
Academic language	299	12,29 %
Subject-specific language	545	22,39 %

Table 16. Overview of the results

5.4.1 General (service) vocabulary

High-frequency vocabulary or general service vocabulary

According to Nation (2001b: 33), high-frequency words are words that are seen in a wide range of language uses, and are therefore also called 'general service words' (West 1953). A list of high-frequency words in the general language was compiled by West (1953). The *General Service List of English (GSL*) is comprised of the 2000 commonest words in the English language. Despite the fact that the list was compiled over a half-century ago, it is still used in academic research and so has to be regarded as valid (Nation 1994: 45). The results of the analysis show that 1392 word types in the PIS corpus are also featured in the *GSL*. In total, 936 word families from the general language occur in the corpus. According to Nation (2001a: 11-12), high-frequency or general service words cover around 80% of an academic text. The analysis shows that over 60% of the corpus is covered by general service words.

General service words dominate the genre. Among the commonest words are function words such as *the*, *of*, and *a*. Common content words are *doctor*, *people*, *age* or *sun*. Although general service words play an important part in PIS, academic and technical vocabulary should be taken into consideration as well. The number of word types and the percentage of coverage of academic as well as subject-specific language is striking. According to Nation (2001a), academic words constitute about 9% of an academic text, whereas technical words cover about 5% of an academic text (11-12, 18). The findings of the analysis show that academic vocabulary covers

[22] Low-frequency words taken from the general language were taken into consideration as well.

around 12% of PIS and technical vocabulary covers more than 20%, These numbers suggest that PIS are more academic and technical than academic texts.
To illustrate the dominance of general service words, I have included the following example. Although it is common practice to explain diseases in language that is easy to understand, example 1 is extraordinary since style and lexis differ slightly from other PIS in the corpus. The authors decided to explain cancer in what appears to be everyday language. There are similarities to spoken language, the most obvious one being that short sentences are used. In example 1 all the words that are featured in the *GSL* are highlighted.

(1) **The body is made up of many types of** cells. **Normally**, cells **grow**, **divide and die**. **Sometimes**, cells mutate (**change**) **and begin to grow and divide more quickly than other** cells. **Rather than dying, these** abnormal cells clump **together to form** tumours.

(The American Academy of Family Physicians: Melanoma)

Words from the *GSL* dominate this excerpt: including function words such as *the*, *of*, *and*, *to*; content words for certain nouns (*body* and *types*); verbs such as *begin*, *grow*, *divide* and *die*; and adverbs such as *quickly*. Only five word types do not fall into the category of general service words. These are *cells*, *mutate*, *tumours*, *abnormal* and *clump*. I will now present a contrasting excerpt. As in the first example, the text is dominated by general service words or high-frequency words. However, the style and lexis differ. The language is more complex, as longer sentences are used to describe how a specific type of cancer develops.

(2) Colon cancer **is caused by the** abnormal **growth of** cells **in the** lining **of the** bowel. **Usually small** lumps **called** polyps **begin to form**.
Commonly these lumps are referred to as tumours.

(CKS: Cancer of the colon, rectum or bowel)

Although the style and lexis vary in each PIS, high-frequency words (or general service words) are heavily featured in most PIS, covering more than 60% of the running text. This leaves up to 40% of the text to be explained in alternative vocabularies, which I will now turn my attention to in the following section.

Low-frequency words

Low-frequency words form the largest group of words in any given language as they account for all the words that do not fall into any of the other categories. According to Nation (2001a: 18), they account for around 5% of an academic text. Low-frequency words are proper nouns, technical words from other subject areas, words that are not commonly used in a language and/or words of 'moderate frequency' that do not fall into the category of high-frequency words. It should be highlighted that there is a blur between the lines as far as high- and low-frequency words are concerned, with much depending on the design and nature of a corpus. In reference to low-frequency words, it is possible that what can seem like a technical word to one individual may be viewed as a low-frequency word by another individual. Frequency and usefulness, in this sense, are dependent on the observer and the field s/he engages in. In addition, low-frequency words can be words that are not commonly used, because they might seem to be inappropriate in a specific context (Nation 2001a: 19-20). Out of the 743 word types that were not categorised by the program RANGE, 198 can be considered low-frequency words. The majority of these words do not fall into the category of high-frequency words, although they are from the general language. Consider the following example:

(3) Melanoma is less common but can occur among **Maori** along with **Pacific** and **Asian peoples** in **New Zealand**.
(The Cancer Society of New Zealand: Malignant Melanoma)

In all PIS a reference is made to the (number of) people affected by a specific disease in the relevant country or region. Therefore, the names of ethnic groups, countries and regions are featured in the corpus, but at the same time are not featured in the *GSL*. Neither are adjectives from everyday English, such as *chaotic*, *bumpy*, *huge*, *sticky* and *tiny*, which can be regarded as low-frequency words. It is more likely to encounter these adjectives in spoken rather than written language or style. Also, in the case of adjectives such as *bumpy* and *sticky* the age of the *GSL* clearly must be taken into account.

Words related to nutrition were not categorised by RANGE either. It can be argued whether words such as *tomatoes, muesli*, *wholegrain* and *wholemeal* are technical,

general (or both) depending on the context. Accordingly, the majority of the low-frequency words in the corpus are used in the general language. They can be categorised as low-frequency words as they simply do not come under the category of high-frequency words or general service words.

5.4.2 Academic vocabulary

Academic vocabulary is particular to the areas of knowledge in academic settings. Among the various terms used to refer to 'academic vocabulary' (Coxhead 2000) is 'generally useful scientific vocabulary' (Barber 1962 referred to in Nation 2001a: 188). In an attempt to introduce a list containing the most common academic word families, Coxhead (2000) generated the *Academic Word List* (in short *AWL*). This list is also contained in RANGE, the software that was used to compare the words of the PIS corpus to the *AWL*. The analysis shows that 840 tokens, 299 types or 207 word families of the PIS corpus were classified as academic. Academic words therefore account for more than 10% of the running words in the corpus.

Word list	***Tokens***	***Types***	***Families***
Academic WL	840	299	207

Table 17. Outcome RANGE AWL

These numbers indicate that out of the 570 word families comprised in the *AWL* list, 207 were identified in the PIS corpus. The most common academic word families in the PIS corpus are presented in the table below:

normal	factor	aware	flexible	affect
area	period	available	positive	institute
research	medical	contact	incidence	major
occur	remove	physical	team	specific

Table 18. The most common academic word families found in the PIS corpus

As far as the word family *normal* is concerned, the word type *normal* was featured in the corpus most often, with 21 occurrences (a rate of 0.97 per 1000 words). *Abnormal* on the other hand, was used only 9 times, and the derived forms of *normally* 4 times and *abnormally* just once (an average of 0.41 and 0.18 vs. 0.04 per 1000 words).

To show how academic vocabulary is incorporated into the texts, a short excerpt of a PIS is offered. In the example all the academic words that are featured in the *AWL* are highlighted.

> (4) Although the number of people **affected** with CRC[23] is **declining**, it is still the second most common form of cancer in men and the third most common form of cancer, after breast and lung cancer in women. In 2007, an **estimated** 20,800 Canadians will be diagnosed with CRC and 8,700 will die of it. On average, 400 Canadians will be diagnosed with CRC and 167 will die of CRC every week.
>
> (Health Canada: Screening for colorectal cancer)

The word types *affected, declining* and *estimated* occur in the example. *Affect* is an academic word family that with its inflected forms occurs at a rate of 0.88 per 1000 words (19 occurrences). The word families *estimate* and *decline* are less common. *Estimate* is however used regularly when referring to the number of people diagnosed with a specific cancer.

Other common academic words families are *factor*, *area* and *remove*. The *factor* word family, in the plural and singular form occurs 64 times in the corpus (a rate of 2.97 per 1000 words). The commonest type of the word family *area* is the singular form, which occurred 20 times (the plural form was featured 17 times), a rate of 0.92 and 0.78 per 1000 words. The word family *remove* is also featured in the corpus. It seems that the verb form *remove* (10 occurrences) and its inflected forms *removed* (15 occurrences) and *removes* (2 occurrences) is generally preferred over the noun *removal* (6 occurrences). Considered together, all the various inflections of

[23] CRC stands for colorectal cancer.

the verb form account for a rate of 0.78 per 1000 words, while the noun only has a rate of 0.27. Nation (2001a: 11-12) has proposed that about 9% of the vocabulary in an academic text is in fact academic vocabulary. In PIS academic words cover around 12% of the vocabulary, and so are arguably more academic than academic texts.

5.4.3 Technical vocabulary

Particularity to and usefulness for a specific area of knowledge characterises technical vocabulary (Nation 1994: 47). In the literature a range of terms is used to refer to technical vocabulary, including 'specialist vocabulary' (Kennedy & Bolitho, 1984 referred to in Chung & Nation 2004: 252). So far, little research has been conducted on technical vocabulary, and even less on how to identify it. One of the first theorists to explore the topic was P. Nation and T.M. Chung (2004). Together they have summarized and evaluated different approaches of identifying technical vocabulary.[24]

One possibility to identify technical vocabulary is to use a rating scale in order to decide whether particular words fall into the category of technical vocabulary, and whether these words have specialised meanings in a specific context. This approach relies heavily on the researcher's background knowledge of the relevant area, their skill to rate words, and their intuition. Good subject knowledge is therefore a requirement (Chung & Nation 2004: 253). According to Chung & Nation (2004: 261), this approach is the most accurate, but also the most laborious and time-consuming.

Technical vocabulary can also be identified either by (a) a specialist of that specific field (b) using a specialist dictionary compiled by subject experts or (c) using clues in the relevant text (Chung & Nation 2004: 252). Chung & Nation's (2004: 252ff) evaluation of approach (c) revealed that clues in texts are difficult to detect since writers often do not see the need to highlight technical words.

[24] For more information see Chung & Nation (2004).

As for the dictionary approach, the quality of the relevant dictionary can be decisive (Chung & Nation 2004: 261). In their evaluation Chung & Nation (2004) did not deal with approach (a).

Another possibility for identifying technical words is to use a computer-based approach. This requires the use of computers, computer-accessible corpora and analysis programs. All of the aforementioned facilitate the analysis process, as word frequency data of a specific subject can then be generated easily, and compared to frequency data of another field or fields (Chung & Nation 2004: 252). According to Chung & Nation (2004: 262), the computer-based approach has disadvantages in that it is "not inclusive enough", and therefore makes it difficult to generate "a definite list of [technical] terms".

For this study, the dictionary-based approach was chosen. All the words that were not classified as general service words or academic words by RANGE were checked against three online editions of dictionaries: '*Merriam-Webster's Medical Dictionary'* (2008), '*Concise medical dictionary'* (2003) and '*Dictionary of Nursing'* (McFerran 2003) to confirm their technical status. All of these dictionaries are well-known and reliable reference sources that deal with many areas, including emergency medicine, anatomy, nutrition, diagnosis etc. Furthermore, they are addressed to different target audiences. *Merriam-Webster's Medical Dictionary* (2008) is intended for a broader group of foreign language learners, while the target audiences of the other dictionaries are more specific, e.g. medical students, doctors and nurses. The analysis showed that out of the 743 word types that were not categorised by RANGE, 545 are particular to the field of medicine (more specifically, to oncology). In table 19, the most common technical words are presented.

cancer	breasts	nodes	diagnosis
breast	colorectal	tissue	mammograms
prostate	lymph	rectum	urine
bowel	diagnosed	nipple	x-ray
melanoma	radiation	biopsy	GP (general
symptoms	therapy	colon	practitioner)
cells	chemotherapy	mole	cell
surgery	radiotherapy	scan	rectal
cancers	PSA (prostate-	UV	tumours
hormone	specific antigen)	CRC	glands
		(colorectal cancer)	menopause

Table 19. The most common technical words

All of the words presented in table 19 are technical words, although the degree of technicality varies in each case, for example, the technical words *breast* and *biopsy*. It is possible to categorise technical words according to their particularity to a specific field. Nation (2001a: 198-199) established a distinction of technical vocabulary into four categories. The first categories consists of words that seem to be particular to one specific field and are rarely found outside that field; such as *jactitation* or *per curiam* are unique to law (Nation 2001a: 198-199). The second category is comprised of words that are used within and outside a specific field. Examples would be the words *cite* or *caution* in law (199). Words belonging to the third category are used inside and outside a specific field. Their specialised meanings in one field can be derived from those of another field. An example would be the word *accused* in the field of law (Nation, 2001a: 199). The fourth category includes words that are more frequent in a specific field. Although they do not have specialised meanings, "someone knowledgeable in the field would have a more precise idea of its meaning" (Nation 2001a: 199).

The distinction of technical words according to these four categories is not completely straightforward, as the borders between categories are not clear-cut as words often fall into more than one category. For instance, the word *node* is subject-specific, but is also a word that is used in disciplines such as aviation and

telecommunication. Its meaning is therefore altered accordingly. Nevertheless, a categorisation was still attempted:

Category 1: *Subject-specific words* Cancer, magnetic resonance imaging (MRI), computer tomography (CT), ultrasound, biopsy, chemotherapy, radiotherapy, tumour, leukaemia, abdomen, lymph, plasma, endoscope, scalpel, …
Category 2: *Technical words used inside and outside a specific field* Tissue, node, cell, …
Category 3: *Words used inside and outside a specific field* Tube, spread, scan, needle, …
Category 4: *Words with no specialised meaning that are used in a particular field* Sun lotion, sunscreen, sunbed …

Table 20. Categorisation of technical vocabulary

In PIS the number and range of technical words is striking. Subject-specific words (category 1) are of special interest, since they include words that refer to the human anatomy (*abdomen*, *lymph*), diseases and conditions (*nausea*), instruments (*endoscope, scalpel*), diagnostic procedures (*ultrasound, biopsy*), and words that name therapies and treatments (*chemotherapy, radiotherapy*). Subject-specific words also cover abbreviations (*MRI – magnetic resonance imaging*).

Words that fall into the second and third categories and which thus have specialised meanings occur in the corpus as well. An instance of a word with specialised meaning is *tissue*, which has different meanings in general and subject-specific language. In general language, it is used to refer to a cloth or a handkerchief, but in medicine it refers to "an aggregate of cells usually of a particular kind […]" (*Merriam-Webster's Medical Dictionary*, 2008). Like *tissue,* the word *spread* denotes different concepts in different disciplines. In cooking, *spread* has a different meaning as in finance or medicine. In medicine *spread* is also used in the sense of distribution, and so its meaning can be derived from that from another field (namely, finance). To illustrate how *spread* and *tissue* are used in PIS an example sentence is

presented below. In the example all the technical words have been highlighted in bold. Although example 5 only consists of one sentence, this one phrase is still comprised of four technical expressions.

(5) If the **cancer cells** have **spread** into the surrounding **breast tissue**, this is called **invasive breast cancer**.

(CKS: Breast cancer)

The most obvious technical terms are either Latin- (*abdomen, cancer, carcinoma, cell, tissue, tumour*) or (Ancient) Greek-based (*diabetes, colon, colostomy, stoma*). Due to their declinations such as *-a, -ae, -us, -i, -um, -u, -ua, -es* in Latin, or *–on* in Greek, they are easier to detect. In some cases the writers of texts use specific clues to indicate that the word is different than others (Chung & Nation 2003: 112). These can include definitions of words in the running text, the use of italics or bold, and the use of words in labelling parts of a diagram (ibid.). In PIS definitions or abbreviations of technical words are part of the running text. Very often these are also offered in brackets (see example 6), and so brackets can often serve as clues for identifying some technical words in the genre of PIS.

(6) Treatment can cause serious problems, including **impotence (inability to get or keep an erection)** and **incontinence (loss of urine)**.

(American Academy of Family Physicians: Prostate cancer)

Difficulties in defining technical vocabulary made it impossible to solve the question of technical vocabulary size. Nation (2001a) assumes that technical vocabulary of e.g. geography or applied linguistics is less than a thousand technical words (Nation 2001a: 18). Although this study is small-scale, as the focus has been on four types of cancer, over 500 word types could be identified as technical. A much larger study would be needed to address the issue of technical vocabulary size in the field of oncology.

5.5 Vocabulary in discourse: multi-word units

So far, I have looked at the different vocabulary types that are found in PIS. When dealing with vocabulary (analysis) it is common practice to focus on the word as the relevant unit. However, this emphasis on words neglects other units that are also worth exploring. These are called multi-word items (Moon 1997: 43), or multi-word units (MWUs) (Schmitt 2000: 97). MWUs are items that consist of a "sequence of two or more words (a word being simply an orthographic unit)" (Moon 1997: 43). MWUs are semantically as well as syntactically inseparable. One should think of the components of multi-word items as

> acting more like the letters of an alphabet – each one contributing to the recognizable shape of the higher unit, but not necessarily adding a clearly defined meaning. (Sinclair 1997: 35)

It is thought that they are the results of fossilisation and word-formation (Moon 1997: 43). Three criteria have been identified that define MWUs and that make it possible to differentiate them from other lexical strings (ibid.). These criteria are institutionalisation, fixedness and non-compositionality. 'Institutionalisation' is the extent to which a multi-word item is conventionally used in a language, and whether it is actually regarded as an inseparable unit by a language community. Fixedness refers to the degree the sequence of multi-word items can or cannot vary. Non-compositionality is the third criterion in defining MWUs (Moon 1997: 44, Schmitt 2000: 97). Some MWUs have specialised meanings that their individual components do not convey, whereas others are more transparent (ibid.). Among the types of multi-word units are compounds, phrasal verbs, idioms, fixed phrases and so called prefabs or preconstructed phrases (Moon 1997: 45-48). In regard to language teaching, it is possible to say that multi-word items have been neglected for a long time – the only exceptions being phrasal verbs. These have attracted attention since they are commonplace in the English language (Sinclair 1997: 35). In the following sections I will begin by focusing on compounds, as they are commonly used to describe medical concepts in PIS, and then move onto lexical phrases.

5.5.1 Compounds

Compounds are by far the largest and most flexible group of multi-word units (Moon 1997: 44-45). A compound typically consists of two or more orthographic words that are combined to form a single lexeme (Moon 1997: 44, Schmitt 2000: 99). As far as the characteristics of compounds are concerned, it is possible to argue that the sequence of the component words is usually regarded as fixed or frozen. In addition, the degree of institutionalisation and compositionality of compounds can vary as the process of compounding is a highly productive means of word formation (ibid.). In PIS compounds play an interesting role. Technical words, for example, do not only include one-word items such as *biopsy*, but also multi-word items such as *magnetic resonance imaging (MRI)* or *computer tomography (CT)*. Two-word, three-word and even four-word compounds are common in PIS. Most compounds function as nouns.

Two-word compounds
skin cancer, bowel cancer, colon cancer, malignant melanoma, tissue sample, family history, medical history, general practitioner (GP), clinical oncologist, Klinefelter's syndrome, Duke's classification, Gleason score, hormone therapy (HT), computerised tomography or computer tomography (CT), ultrasound scan, barium enema, …
Three-word compounds
Basal cell carcinoma, squamous cell carcinoma, irritable bowel syndrome (IBS), sun protection factor (SPF), body mass index (BMI), nonsteroidal anti-inflammatory drugs (NSAIDS), prostate-specific antigen (PSA), familial adenomatous polyposis (FAP), hormone replacement therapy (HRT), fine needle aspiration (FNA), digital rectal exam (DRE), magnetic resonance imaging (MRI), sentinel node biopsy, …
Four-word compounds
lobular carcinomas in situ (LCIS), ductal carcinoma in situ (DCIS), hereditary nonpolyposis colon cancer (HNPCC), faecal occult blood test (FOBT), classical atypical mole syndrome (CAMS), …

Table 21. A selection of compounds

The table shows a selection of compounds that were found in the specialised corpus. Among the compounds in the PIS corpus are two-word, three-word and even four-word compounds that refer to diseases, diagnostic procedures, therapies and other medical concepts. How these compounds are used in PIS is illustrated below, in an example on the topic of breast cancer taken from the section of the PIS on treatment options. In the excerpt, advantages and disadvantages of a therapy called hormone replacement therapy are discussed.

(7) Talk to your doctor about the risks and benefits of **hormone replacement therapy (HRT)**. **HRT** can relieve symptoms of menopause and reduces the risk of osteoporosis and **colon cancer**. However, **HRT** increases the risk of **breast cancer** and **heart disease**.
(Health Canada: Breast cancer)

In example 7 there are several compounds referring to medical concepts. In this specific case they denote a kind of therapy or disease. Whereas one of the compounds consists of three words (*hormone replacement therapy*), the others are comprised of two orthographic words (*colon cancer, breast cancer and heart disease*). All of the compounds are used as single lexemes.
The compound *hormone replacement therapy* is used to refer to a therapy. According to the *Merriam-Webster's Medical Dictionary* (2008), this therapy involves "oestrogen and synthetic progestin to be administered in order to relieve the symptoms of e.g. menopause". In the example the compound is used to introduce a treatment form. The abbreviation of the therapy is also mentioned. Once the full name of the therapy has been mentioned, the writer resorts to the abbreviation – presumably to make the text shorter.

Another example of a compound is *Dukes' classification*, which is used to refer to a staging system developed by Cuthbert Esquire Dukes. It is an instance of onomastic evidence (Müller 1993), which means that a proper name has been used in coining the term. Coining new terms is not a rare phenomenon in the sciences since new findings create the need to name scientific concepts (Müller 1993: 93ff). This, of course, is also the case in medicine. *Dukes' classification* is a compound referring to a staging system that is used to stage colon cancer (CKS: Bowel cancer), and follows

the structure <name + ‘s + substantive> (Müller 1993: 93). This structure is often used when naming concepts of a specific field, such as diseases, viruses, organs, syndromes, vaccines etc. (Müller 1993: 95). Examples from the corpus include *Crohn’s disease* and *Klinefelter’s syndrome.* In all these examples the surnames of scientists were used to coin the terms. However it is also possible to use other proper names; examples taken from the corpus are *Health Canada*, *National Health Service (NHS)* or *National Cancer Institute*. These compounds refer to national organisations and institutions. As all the examples above illustrate, compounds contribute in a significant way to medical termini and nomenclature.

I will conclude by discussing the role of Latinisms. In medical discourse Latinisms such as *post mortem* play an important role (Müller 1993: 82). These can be compounds or word components of compounds. In the genre of patient information sheets, Latinisms play a minor role – only *in situ* occurs in the corpus. It is featured three times, but two out of those three it functions as a word component of a term describing a specific type of breast cancer (*ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS))*. Only once is the Latinism itself introduced and explained:

> (8) Some breast cancers are found when they are **‘in situ’**. This means they have not spread outside the milk duct or lobule where they began.
>
> (The Better Health Channel: Breast cancer)

Patient information sheets are aimed at a layperson audience, and so Latinisms seem to be avoided.

5.5.2 Lexical phrases

Lexical phrases are another category of MWUs. Coined by Nattinger & DeCarrio (1992), the term 'lexical phrase' refers to a lexico-grammatical unit (1992: 1). Among the expressions used for lexical phrases are formulaic expression or chunks of language (ibid.). These chunks of language can vary in length. Therefore, there are short ones such as *a _ ago* and longer ones such as *if I X, then I Y* (Nattinger & DeCarrio 1992: 1). The composites of lexical phrases can be other MWUs, or simply "any string of words that are commonly used together" (Schmitt 2000: 101). In regard to other characteristics of lexical phrases, it is possible to say that they can be placed somewhere between the poles of lexis and syntax, as they have combined both form and function (Nattinger & DeCarrio 1992: 1). According to Schmitt (2001: 101), formulaic expressions are linked to functional usage. To illustrate this I will refer to the example offered by Schmitt (2000), *to make a long story short* is commonly understood as the introductory phrase of a summary. Therefore, one use of lexical phrases is related to language functions such as summarising and therefore is also linked to language use and competence (Schmitt 2000: 101).

In regard to teaching, lexical phrases play an important role in vocabulary acquisition, since they constitute an important part of a native speaker's vocabulary, and are often used to a great extent. Lexical phrases are also therefore easily memorized (Schmitt 2001: 101). Moreover, lexical phrases are essential for organising a spoken or written text as they can function as macro- or micro-organisers (Schmitt 2001: 105-108). It is therefore an advantage for language learners to be familiar with lexical phrases.

Another point about the teaching and learning of lexical phrases in the language class is that lexical phrases are functional. In the previous paragraph it has been mentioned that these formulaic expressions are linked to functions such as summarising (*to make a long story short, to sum up*). Therefore, a great amount of time should be spent on teaching these chunks of language. Finally, another issue involving the introduction and use of lexical phrases in the language class is the way that new vocabulary is processed in the brain. In the literature (e.g. Dudley-Evans & St. John 1998, Schmitt 2000) it has been suggested that language learners process and store chunks of languages rather than individual words. Given that learners process and

store in units, it could be argued that lexical phrases ought to have a more prominent role in the teaching of vocabulary. I would like to take the opportunity to offer some examples of lexical phrases taken from the PIS corpus. The first example shows the typical frame *the ___er, the ___er*. The slots in this frame can be filled with adjectives. In the example the frame is used as follows:

(9) **The earlier** a cancer is found, **the easier** it is to treat **and the more** likely it will be curable.

(The Better Health Channel: Bowel cancer)

In the example the lexical phrase is used to organise cause and effect. In particular, the effects of an early diagnosis on the treatment and the chance of cure are outlined. The use of the frame *the ___er, the ___er* (and in this case) *and the more* ___ is a simple way to express cause and effect. Moreover, it contributes to the coherence of the sentence.

Another lexical phrase in PIS is *if X, then Y*, which is even commoner than the one previously introduced. *If X, then Y* also expresses cause and effect. This is shown in the example below, which outlines the procedure following the identification of breast changes. Unlike the proposed scheme above, *then* is left out.

(10) If you have a change in your breast, you may have several tests.

(The Cancer Society of New Zealand: Breast cancer)

In PIS the formulaic expression *if X, then Y* is also used in order to give advice. Modal verbs such as *could* and *should* play key roles. In the example the patient is urged to consult her/his doctor, if s/he feels to be at greater risk of developing melanoma.

(11) If you are at high risk of developing melanoma, you should discuss with your doctor what methods or protection and checking of your skin are appropriate for you.

(The Cancer Society of New Zealand: Melanoma)

In conclusion, lexical phrases play an important part in the micro- and macro-structuring of the genre. In addition, they contribute to the realisation of the communicative purpose of genre and specific genre moves. This will be discussed in the following chapter, which is focused on the move structure of PIS. In regard to teaching, it can be said that lexical phrases "provide (learners) with a number of options for expressing moves when teaching speaking or writing" (Dudley-Evans & St. John 1998: 86). However, Dudley-Evans & St. John (1998: 86) also stress the fact that it should not be the aim to promote "the unthinking learning of set phrases" to express communicative purposes. In summary, lexical phrases can provide insights into the way communicative purpose is realised.

In this section of this paper I have looked at vocabulary in discourse. It has been shown that multi-word units are common in PIS. In particular, compounds are used to refer to medical concepts such as diseases. Furthermore, it has been highlighted that proper nouns are used as constituents of compounds. In the final part of the section the focus has been on lexical phrases, their role in the language, in language teaching and in the genre of PIS. It has also been highlighted that an investigation of lexical phrases can be fruitful in combination with a genre analysis.
Before turning to the move structure analysis of PIS, I would like to provide a summary of the main findings of the vocabulary analysis.

5.6 Summary of findings

The results of the analysis show that vocabulary from the general language dominates the vocabulary found in PIS, accounting for more than 60%. Therefore, the findings of the study confirm the suggestion of Nation (2001a) that high-frequency words constitute the main part of a text and that they occur in a wide range of language uses. The probability of learners encountering high-frequency words is therefore extremely high, and so a "considerable time should be spent on them by teachers and learners" (Nation 2001a: 16). Notably, Nation (2001a: 11-12) contends that high-frequency words form about 80% of an academic text. In PIS however, they only form about 60%. I do agree with Nation (2001) that attention should be given to high-frequency words from the general language in the language classroom. Nevertheless, the number and multitude of academic as well as technical vocabulary in PIS are striking. Academic and technical words form 12% and 22% of PIS, as

opposed to 9% and 5% (in academic texts) as suggested by Nation (2001a: 11-12, 18). According to these figures, PIS are more academic and technical than academic texts. In terms of technical vocabulary, the findings of the study suggest that the core vocabulary in the field of oncology centres on cancer, cell, anatomy, risk groups, prevention, diagnostic procedures, treatment and therapy options, instruments and drugs. The technical words can be single-word items (*scalpel*) or compounds (*magnetic resonance imaging – MRI*). Whenever a technical word such as *chemotherapy* or *faecal occult blood test (FOBT)* or its abbreviation is introduced in PIS, it is defined and/or described. Some medical terms are Latin- and Greek-based words. However, Latinisms are hardly found in patient information sheets. This can be explained by the fact that these sheets are addressed for a layperson readership, and the use of Latinisms would therefore make it more difficult for laypeople to access these texts.

The findings of the analysis suggest that multi-word units are used in discourse.
Types of multi-word units common in PIS include compounds and lexical phrases. The former are mainly used to refer to medical concepts such as diseases (*breast cancer*), therapies (*hormone replacement therapy*) and so on; the latter are essential in any kind of text, as they contribute to the macro- and micro-structure of a text. Moreover, it has been stressed that the investigation of lexical phrases of a specific genre can be insightful in terms of teaching and learning a genre. In the following chapter I will explore the generic structure of PIS. Hence the communicative purposes and language functions inherent in the individual moves will be considered.

6 The macrostructure of patient information sheets

In the previous chapter the vocabulary in PIS was investigated. Now I will focus on the macrostructure of the genre. The way a text-genre is built up can reveal information about the cognitive aspects of language organisation. Writers in a certain field seem to organise what they would like to convey in a rather consistent way (Bhatia 1993: 29), and thus contribute to the establishment of prototypical forms of genres (Upton & Connor 2001: 317). The idea that these genres can be investigated by teachers and learners (and also produced by learners) is essential to the ESP approach to genre analysis.

A genre consists of various elements. Hasan (1984, 1989) has called them 'elements of text structure' or 'elements'; Swales and Bhatia, following Hasan (1984, 1989), however, preferred the term 'moves'. 'Move' is now the most commonly used term (Henry & Roseberry 2001: 94). A move is seen as contributing to the overall communicative purpose of a genre (Bhatia 1993: 30). This is illustrated in the figure below. Comments on teaching and learning will be made in the final section (7).

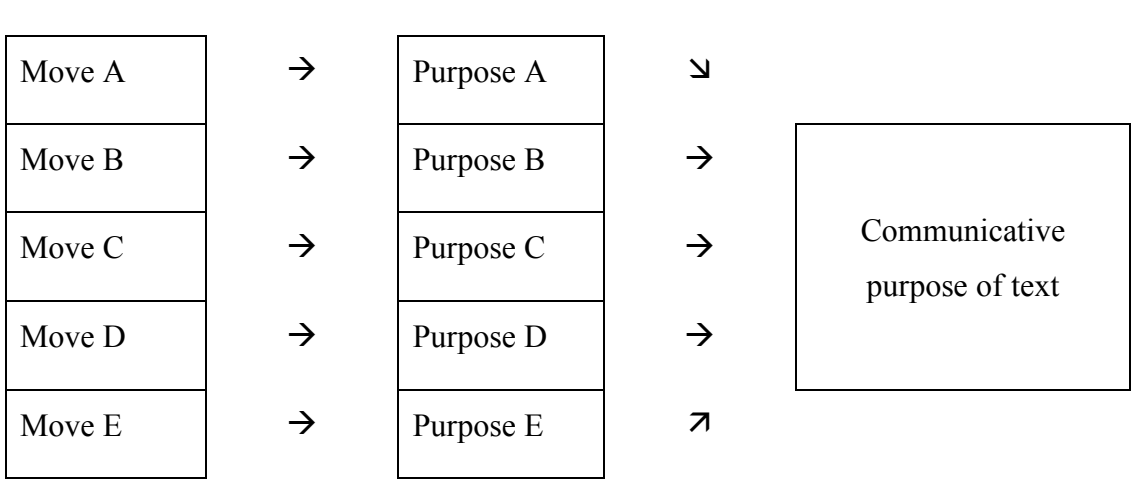


Figure 4: Moves and the communicative purpose of a text (Henry & Roseberry 2001: 95)

Moves have different characteristics. First of all, they serve the overall communicative purpose of a text. Secondly, they do not have to "coincide with paragraphs" (Bhatia 1993: 56), which sometimes makes it difficult for researchers to identify them. Thirdly, moves can be categorised according to their frequency. There

are optional, core and obligatory moves. Henry & Roseberry (2001: 95) suggest that moves that appear in less than 50% of the texts of a corpus should be regarded as optional moves. Those that appear in 50-90% of the texts can be described as core moves. Obligatory moves are those appearing in 90-100% of the compiled texts (Hüttner 2007). Another characteristic of moves is that they can consist of sub-units known as steps. These, like moves, can be categorised according to their frequency. A step can contribute to the overall communicative purpose of a genre text.

7.1 Communicative purpose(s) of PIS

Patient information sheets serve a variety of communicative functions. The primary aim is to be informative. Patient education/information materials are given to patients or even members of the patients' families prior to or after any diagnostic procedure, surgery or treatment, in order to offer information, address and clarify issues concerning an illness, a type of treatment etc. (Krois-Lindner 2007). Often patients are too shocked, shy or unprepared to ask questions during consultations; and so questions may well be addressed later. PIS enable patients and/or their family members to revisit the information multiple times. Since PIS are addressed to laypeople rather than specialists, it is important for them to be informative, in the sense that they should clarify matters rather than complicate them (Johansson 2005: 220). It is important to note that a patient often makes decisions e.g. on treatment (in terms of 'informed consent'[25]) that have been based on the information (i.e. PIS) given to them.

Another function of patient information sheets is that they protect hospitals, doctors (and so) from legal lawsuits. In Austria, PIS are accompanied by consent forms. Whenever a patient decides to undergo surgery, s/he has to sign such a form.

When the patient legally consents to treatment, they also declaring that they have read the entire PIS and are also aware of possible complications and side effects of the procedure (Krois-Lindner 2007).

The needs of a layperson audience must be taken into consideration when designing and writing PIS. In the genre it is necessary to introduce, describe and explain medical concepts such as diseases, diagnostic procedures, instruments and

[25] For more information see section 4.1.

sometimes medication. One should bear in mind that the major purpose of written professional discourse is to convey information "clearly, briefly, directly, and sincerely" (Scollon & Scollon 1995: 94).

7.2 General observations

There is a multitude of expressions referring to patient information sheets (PIS). Terms include patient information leaflets, (health) fact sheets, patient education sheets or patient education material[26]. There are also brochures, however, these usually exceed PIS and leaflets in length as they consist of more than 30 pages (and so were not taken into consideration for this study). This multitude of expressions could imply that there is no convention regarding the genre name; and probably no convention on the genre itself. During the course of my research I have found PIS on different types of diseases, treatments options, medication and studies, which can be regarded as sub-genres of PIS. Therefore, the genre investigated in this study is a sub-genre, since the focus was on specific cancer types. But here again, some PIS just focused on a specific diagnostic procedure.

The analysis of the move structure showed that there is agreement on the overall generic structure. Although the individual steps vary from PIS to PIS (even with those of the same provider), PIS dealing with specific types of cancer are usually comprised of sections on the disease itself, signs or symptoms, causes and risk factors, diagnostic procedures, treatment options and prevention. Another common move is to refer to other sources of information.

Another interesting point is that PIS are made available online and/or in printed formats. Whereas some PIS providers just publish their intellectual property online, others also distribute their material in form of handouts or leaflets. The use of different communicative channels and the existence of various names for PS imply that PIS are a hybrid genre rather than a homogenous one.

[26] According to my research.

7.3 Move structure analysis

The move structure analysis is covered different stages. The initial stage of the analysis is comprised of the reading of the PIS in order to gain an overview, and a general feeling about the generic elements and strategies of the genre. As a second step, the different moves were identified. The basic skeleton of PIS was determined rather quickly; the identification of individual moves and steps, on the other hand, turned out to be rather difficult. It has to be noted that the determination of moves and steps depends on the researcher's intuition and interpretation. I identified moves and steps according to the following categories: optional, core and obligatory. In this study, the optional moves do not play an important role, since the scope of the paper is very focused. Only extraordinary steps and moves that fell into the category 'optional' are mentioned briefly. Core and obligatory moves are of interest and are thus outlined in detail.

As a final step of the move structure analysis, words, phrases and example sentences were extracted from the corpus. This part of the study was conducted with the aid of a computer program. The software *WordSmith*, in more detail the program *WordSmith concordance program*, was used to see how specific words or phrases are used. The aim of the move structure analysis was to identify core and obligatory moves of the genre and the move order. In the following sections, the allowable move order and the individual moves and steps are presented and outlined.

Move 1: Introducing the institution or the project
Step 1.1: Stating the name
Step 1.2: Presenting the logo and/or catch phrase

Move 2: Identifying the genre

Move 3: Offering general information about the disease
Step 3.1: Signalling the beginning of a new section by stating the name of the disease
Step 3.2: Offering a definition of the body part and/or disease
Step 3.3: Describing body part and its function/ type of cancer/ procedure and what it does (in the body)
Step 3.4: Referring to the (number of) people affected

Move 4: Introducing symptoms or warning signs
Step 4.1: Signalling the beginning of a new section by stating 'symptoms' or 'signs'
Step 4.2: Referring one/main warning sign(s) /symptom(s)
Step 4.3: Listing symptoms/warning signs
Step 4.4: Encouraging patient to go and see/ discuss this issue with his/her doctor

Move 5: Introducing causes or risk factors/ groups
Step 5.1: Signalling the beginning of a new section by stating 'causes', 'risk factors' or 'risk groups'
Step 5.2: Stating that causes/risks are (not clearly) known
Step 5.3: Listing or outlining the factors that increase the risk
Step 5.4: Referring to risk factors of a specific group

Move 6: Introducing diagnostic procedures
Step 6.1: Signalling the beginning of a new section by stating 'diagnosis'
Step 6.2: Introducing diagnostic procedure A
Step 6.3: Describing function and/or procedure of A
This is combined with Step 6.4
Step 6.4: Outlining exact procedure and instruments involved
Step 6.5: Introducing diagnostic procedure B (Steps 6.3 and 6.4 are repeated)
Step 6.6: Introducing diagnostic procedure C (Steps 6.3 and 6.4 are repeated)
Step 6.7: Explaining that if cancer is diagnosed, other tests will have to be carried out

Move 7: Introducing treatment options
Step 7.1: Signalling the beginning of a new section by stating 'treatment'
Step 7.2: Introducing (one/main) treatment option(s) by listing them

Move 8: Introducing the topic of prevention
Step 8.1: Signalling the beginning of a new section by stating 'prevention' or 'how to reduce/minimize risks'
Step 8.2: Giving advice on how to reduce risks of developing a specific cancer

Move 9: Offering more information
Step 9.1: Signalling the beginning of a new section by stating 'other organisations', 'selected links', or 'want to know more?'
Step 9.2: Offering contact information and/or referring to organisations/ articles/ websites and/or doctors

Move 10: Stating copyright statement

Table 22. Move structure of patient information sheets

Before turning to the individual moves and steps, I would like to briefly deal with the role of headings. Generally speaking, headings are used as a means of structuring PIS, so that the layout is more attractive and it is easier for the layperson readership to identify the different sub-topics and sections.

7.3.1 Move 1: Introducing the institution or the project

Move 1 is essential as it is featured in all the PIS that were investigated.

The institution or the project is introduced in two steps. Firstly, the name is stated; then the logo and/or the catch phrase are presented. In doing this, PIS providers ensure that their products are identified as their intellectual property. Another function of this move is to display that the institution is respectable and its publications are reliable. In addition, there is a promotional aspect to this action. Logos can be eye-catching, and thus they can be memorised and recognized easily. The same is true for catchphrases. Interestingly, the catchphrases in PIS mirror the institution or project's missions or intentions. *Health Canada*, for example, uses the slogan "It's Your Health" (Health Canada: Breast cancer). This catchphrase is used to address the reader directly and encourage him/her to take responsibility for his/her individual health in terms of patient empowerment. In the example below the realisation of Move 1 is presented. *The Better Health Channel* is introduced with the name and logo. The apple could be seen as a symbol of good health as in "an apple a day keeps the doctor away"). The catchphrase "healthier living, online" clearly refers to the project's aim of providing information on health and prevention via the Internet (Better Health Channel: Prostate cancer).

Figure 5: Realisation of Move 1 (Copyright: State of Victoria – Better Health Channel)

7.3.2 Move 2: Identifying the genre

In over 50% of the analysed PIS, the genre was identified. By identifying and naming the genre, the authors of PIS identify the target audience as well as the primary function of the genre, thus providing information for patients. This move illustrates that there is a range of expressions used for PIS. Whereas some authors of PIS use the term 'patient information sheet', others label the genre 'fact sheet' or 'patient information leaflet'. This multitude of expressions shows that there is no common agreement on the name of the genre and therefore possibly on its conventional features.

7.3.3 Move 3: Offering general information about the disease

The topic of the relevant PIS is introduced through the third move. Taking the form of a heading, the name of a disease, and in some rare cases the name of a specific diagnostic procedure is highlighted for the readers. What follows is a general introduction to the topic by offering basic information such as a definition of the body part or the disease, the function of the body part or a description of what cancer cells do in the body. Eventually some facts and figures concerning the (number of) population suffering form a specific cancer are offered. Steps 3.2 and 3.4 are obligatory, whereas Step 3.3 is core as they appear in 90% and 50% of the PIS that were investigated for this study. The overall aim of this move is to offer a helpful and concise introduction to the topic in order to enable the reader to gain a broad synopsis of the issue at hand. The structure and realisation of Move 3 (except Step 3.1) is presented below:

Move 3: Offering general information about the disease
Step 3.1: Signalling the beginning of a new section by stating the name of the disease
Step 3.2: Offering a definition of the body part and/or the disease
Step 3.3: Describing body part and its function/ type of cancer/ procedure and what it does (in the body)
Step 3.4: Referring to the (number of) people affected

Table 23. The Structure of Move 3

(12) The prostate gland is part of the male reproductive system. It produces some of the fluid that makes up semen. Prostate cancer affects one in 11 Australian men and is common in the over-65 age group. Around 3,000 Victorian men are diagnosed with prostate cancer every year. Many cases are not life threatening because the cancer may be slow growing and usually occurs in older men.

(Better Health Channel: Prostate cancer)

In the first two sentences Steps 3.2 to 3.4 can be identified. The first sentence offers a definition of the prostate and the second refers to its function in the human body. In the third, fourth and fifth sentences a specific risk group is identified and the number of newly-diagnosed cancers per year is explained. In the last sentence a reference to the chance of cure is made. This generic element, however, turned out to be an optional one as it was featured in only 35% of the PIS that were analysed. In the following sections, the individual steps are explored in more detail, with prominence given to the language functions of defining.

Step 3.1: Introducing the disease/diagnostic procedure by stating the name

This step is not dealt with in detail as it is self-explanatory.

Step 3.2: Offering a definition of the body part and/or the disease

Before exploring the PIS for more examples of definitions, it is necessary to introduce a theoretical framework for definitions. In the 1970s and 1980s there was a growing interest in EST courses and materials (Swales 1971: 66ff). Various publications focused on definitions and their role in the sciences, including some of those written by Swales (1971). According to Swales (1971: 66ff), it is easier to state what a definition is not. Giving an example is not a definition, and so definitions cannot be substituted by examples. Another characteristic of definitions is that there is an information shift from the general to the particular. Within a definition the general properties should be mentioned before the specific ones (ibid.). All these observations lead to the following scheme for general definitions:

$T = G + (d_a + d_b + dc + \ldots\ldots d_n)$
Where T equals the thing to be defined Where = equals be Where G equals a general word class word Where d_a, d_b, etc. are the properties which distinguish T from the other members of the general class
Example: A catalyst (T) is a substance (G) which alters the rate at which a chemical reaction occurs (da) but is itself unchanged at the end of the reaction (db).

Table 24. Scheme for general definitions (Swales 1971: 66-67)

Another common and simple way of expressing definitions is through the following formula (Swales 1971: 67-74):

An x/y (countable/uncountable noun) is a /an … general class word + wh-word…	
A dentist is a person who takes care of people's teeth.	
Summary of the definition formula An x/y is a class-word	Which is verb + -ed … Verb + -ed … For verb + -ing … Wh-word verb + s … Verb + -ing … Preposition wh-word … With noun-phrase … With the property of verb + -ing

Table 25. Another scheme for definitions (Swales 1971: 67-74)

As far as the grammar in definitions is concerned, the indefinite article is used before countable nouns, and it is left out before uncountable ones. It is the indefinite article that is used to indicate that "any representative of this term will fit the assigned class" (Swales & Feak 1994: 37). The second part of a definition sentence is often a full or reduced restrictive relative clause (38).

Swales (1971) differentiates between general, specific and expanded definitions. So far I have dealt with general definitions. For Swales (1971: 75), specific definitions deal with a specific type of a thing (T) rather than the general thing (t).

T + t = {t (thing) or G (general class word)}+ da + db + … dn
An equilateral triangle is a triangle with all three sides equal in length. An equilateral triangle is a plane figure with all three sides …

Table 26. Scheme for specific definitions (Swales 1971: 75)

Turning now to expanded definitions, it is possible to say that they exceed the length of general and specific ones as these are usually comprised of just one sentence. Definitions can be expanded in order to compile examples, use(s), and parts or types (Swales 1971: 77-80). Swales & Feak (1994) use the term "extended definition" instead of "expanded definition". Some possibilities of expansions/extensions are outlined below:

Definition formula + {	Common examples are a, b, c and d. Typical examples are a, b, c, and d. Main types are a, b, c and d. Such as a, b, c, or d. Therefore, it is used … As a result, on of its main uses is …. It consists of … main parts: … Its main components are …

Table 27. Scheme for expanded definitions (Swales 1971: 77-80)

Swales (1981: 107) summarises the findings on definitions that were made in the 1970s and 1980s as follows:

Content:	Term = Class + Sum of differences
Form:	A is B which is C (the relative clause introduces the sum of differences)
Function:	To carry the reader from the known to the unknown by explaining new terms or by re-defining old ones
Distribution:	… in the standard form A is B which …C they are most common in text books…
Example:	"A flower is a group of specialized leaves at the end of a short branch, containing one or more of fours kinds of members. These members are sepals, petals, stamens and carpels […]"

Table 28. Summary of the findings on definitions (Swales 1981: 107)

So far, I have dealt with general, specific and expanded definitions. However, there is another formal definition called 'nominal definition' or 'naming statement'. The basic components are similar but the ordering of the components is different. Most EST researchers and writers exhibit a tendency to see definitions and naming statements as equal in terms of function (Swales 1981:108).

Definitions	A is B which is C
Naming statements	B which is C is known as A, B which is C is called A

Table 29. Formula of definitions and naming statements (Swales: 1981:108)

In law, for example, definitions have a different structure. Swales (1981: 110) offers an example to illustrate this. Giving False Evidence is defined as follows:

> Whoever, being legally bound by an oath or by any express provision of law to state the truth of being bound to make a declaration upon any subject makes any statement which is false and which he knows or believes to be false or does not believe to be true, is said to give false evidence. (Swales 1981: 110)

Ingredient 1	+ Ingredient 2+	Ingredient 3	= False Evidence
Legally binding	+ False statement	+ *Mens Rea* Statement	= False Evidence
↧↓↴		↧↓↴	
a or b or c		a or b or c	

Table 30. Definitions in law (Swales 1981: 110)

A definition in law is typically a rather long sentence comprised of different elements that describe the offence. Furthermore, the definition is towards the end of the sentence. In law definitions tend to be naming statements. In terms of 'naming verbs', these naming statements contain 'said to' rather than 'called' or 'known as' (110). In his article, Swales attempts to show that:

> a common and commonly-identified communicative act such as defining can have different discoursal functions in different academic subjects such as science and law; and that such a difference in function may well show up in differences in the linguistic forms through which definitions are typically expounded. (Swales 1981: 111)

According to Swales, these findings suggest that the perceived uniformity of acts such as 'describing', 'classifying' etc. is not so solid as their functional and formal realisation in the various disciplines seems to differ. I will explore whether this is indeed the case.

In PIS definitions occur. These vary in range and form, thus there are general, specific and expanded definitions. When it comes to defining specific cancer types, specific definitions as postulated by Swales (1971: 75) are preferred. Presented below are three definitions of bowel cancer (also known as colorectal cancer). The standard form of definitions *A is B which is C* can be identified, although it is usually slightly altered as e.g. elements are omitted or added.

(13) Bowel cancer **is** cancer of the colon and rectum, or colorectal cancer.
(Cancer Society of New Zealand: Bowel cancer)

(14) Colorectal cancer **is** a malignant tumour **that** develops over a period of time (at least 10 years) on the bowel wall, before invading the wall and moving on to other organs.

(Health Canada: Screening for colorectal cancer)

(15) Colorectal cancer – **commonly known as** colon cancer, or bowel cancer – **is** any cancer **that** affects the last section of the digestive system. This usually means the colon (large bowel) or rectum (back passage).

(NHS: Cancer of the colon, rectum or bowel)

Example (15) shows the standard form *A is B which (that) is C*. In addition, this definition is an expanded one as the parts of the digestive system are mentioned. Interestingly, technical terms are often explained by offering synonyms in brackets as in example (15). Within this definition there is a naming statement.

Definitions following the formula *A is B which is C* are most common in text books (Swales 1981: 107) where they appear in exactly the proposed form. In PIS, slightly altered formulae can be found.

It is not just definitions that are featured in PIS, but also naming statements. These can also vary in form, just as definitions do. I will illustrate this with an example, in which an altered version of the form *B which is C is known as A* is highlighted:

(16) Cancer of the bowel **is** also **known as** colorectal cancer.

(Better Health Channel: Bowel cancer)

Definitions and naming statements used can include: *be (is)*, *call (is called)*, *mean (means), know (is known)* and *refer (is referred to as)*. *Define* or *defined as* are not featured in the corpus. In terms of tenses, the present simple tense is often used to define and name body parts, diseases, diagnostic procedures and treatment options as the tense is typically used for describing general facts. In regard to the scheme or formula of definitions and naming statements, standard forms are usually slightly altered.

In patient information sheets definitions and naming statements play an important role. As the main topic of the PIS in this study is cancer, it is usually the case that cancer types and medical terminology concerning e.g. diagnostic procedures are introduced and defined for a laypeople readership.

Step 3.3: Describing body part and its function/ type of cancer or procedure and what it does (in the body)

Alongside the language function of defining, describing plays an important part in PIS. General characteristics of body parts such as function and size are introduced, for this purpose, the present simple tense is used. As far as verbs are concerned, it is possible to say that durative verbs (*to be, keep, have*), inchoative verbs (*to get, go, fall, become*) and causative verbs (*to bring, make, put, take and place*) are used for descriptions. Durative verbs are used to refer to a condition or a process that will terminate at some point, while inchoative verbs describe a change in condition. Causative verbs are used to describe consequences of an action (Müller 1992: 73). In examples (17) and (18) I will illustrate how the prostate is described. The durative verb *to be*, the causative verb *to make* as well as the verbs *to mix*, *to produce* and *to make up* are used to describe the size of the prostate and its function in the body:

(17) The prostate gland **is** part of the male reproductive system […]. The prostate **makes** a fluid that **mixes** with sperm and other fluids during ejaculation. A normal prostate **is** about the size of a walnut.
(The American Association of Family Physicians: Prostate cancer)

(18) The prostate gland **is** part of the male reproductive system. It **produces** some of the fluid that **makes up** semen.
(The Better Health Channel: Prostate cancer)

Although subject-specific vocabulary generally consists of technical verbs such as *to medicate*, *to lance*, *to secrete, to abort* and *to alleviate* (Müller 1993: 73), these are not found in the PIS corpus. In order to determine whether medical verbs are used in PIS at all, a study on a larger scale is needed. From what I have observed so far, it can be said that verbs from the general language outperform subject-specific verbs.

In addition, authors of PIS seem to use verbs that differ in formality. Therefore, formal verbs (*to produce*) and less formal verbs (*to make*) can be found in the corpus.

While in other sections of PIS the personal pronoun *you* and the possessive pronoun *your* are used to address the reader directly, this strategy is less commonly employed when discussing the function and characteristics of a specific body part.
Expressions for referring to spatial location are instead used. In table 31 some indicators of spatial location are shown:

in (240x), on (45x), under (10x), around (6x), behind (2x), below (2x), between (2x), in front of (2x), above (1x), beneath (1x)

Table 31. Indicators of spatial location

(19) The prostate sits **under** the bladder, **between** the base of the penis and the front of the back passage (the rectum). The tube (urethra) which takes urine from the bladder to the penis passes through the prostate.
(Cancer Society of New Zealand: Prostate Cancer)

It should be mentioned that the prostate description described in example (19) was accompanied by an illustration. One surprising element is that only in a minority of the PIS (20%) illustrations are used as visual aids to inform the reader about body parts or diagnostic procedures[27].

Step 3.4: Referring to the (number of) people affected
In 90% of the PIS studied, a reference is made to the number of people affected by a specific type of cancer. This step is often described in different ways. In the majority of PIS the number of newly-diagnosed people per year is simply stated. Only in a limited number of cases is a percentage introduced. The latter is often used when a reference is made to a specific risk group.

[27] For more information see section 7.4.

(20) Prostate cancer **affects** one in 11 Australian men and **is common in** the over-65 age group. Around 3,000 Victorian men **are diagnosed with** prostate cancer every year. Many cases are not life threatening because the cancer may be slow growing and usually **occurs in** older men.

(The Better Health Channel: Prostate cancer)

(21) Cancer of the breast becomes **more common with** age with approximately 80% of breast cancers **occurring in** women over the age of 50.

(NHS: Cancer of the breast, female)

When referring to specific risks groups then the following patterns tend to be used:

Cancer X is most common in … (e.g. women, older people, men over 65) is the most common cancer in … (country or risk group) occurs in … affects … is diagnosed in …
People are affected by cancer X are diagnosed with cancer X

Table 32. Lexical phrases frequently used in Step 3.4

7.3.4 Move 4: Introducing symptoms or warning signs

Move 4: Introducing symptoms or warning signs
Step 4.1: Signalling the beginning of a new section by stating 'symptoms' or 'signs'
Step 4.2: Referring one/main warning sign(s) /symptom(s)
Step 4.3: Listing symptoms/warning signs
Step 4.4: Encouraging patient to go and see/ discuss this issue with his/her doctor

Table 33. The structure of Move 4

The overall communicative purpose of Move 4 is to provide information on symptoms or warning signs in order to enable readers to recognise these signs/symptoms, and thus to they can take action and consult a doctor for clarification. Move 4 appears in 18 out of 20 PIS and therefore can be categorised as obligatory. Step 4.1 is obligatory as well. It typically takes the form of a heading. Keywords such as 'symptoms' or '(warning) signs' are used in headings to introduce the section on symptoms or warning signs, but in some cases questions are also used, for instance, 'What are its symptoms?' (Cancer Society of New Zealand: Bowel cancer). Steps 4.2 and 4.3 can also occur at the same time. A reference to one or more of the main warning sign(s) and symptom(s) was found in 55% of the texts. In the relevant cases one warning sign or symptom was highlighted before a list of all the signs or symptoms was offered. In 80% of the PIS, symptoms or early warning signs of a specific type of cancer are just listed. The lists are typically phrased in bullet points. Whenever there is a consensus on symptoms, then they are introduced as in example (22), by using forms of *be* or *have*. In other cases, when authors of PIS are less explicit, modal verbs are employed to express tentativeness and possibility. Modal verbs are therefore used to downplay the reader's expectations. Modal verbs used for introducing symptoms include *might, may* and *can*. In three instances *if you have (the following symptoms)* is used.

(22) The most common symptoms of bowel cancer **are** […].

(The Better Health Channel: Bowel cancer)

(23) Women with breast cancer **can** have many different symptoms. These include:

- a lump or lumps in the breast,
- change in the size, shape or skin of the breast,
- changes in the nipple,
- blood-stained discharge from the nipple, and
- pain in the breasts (this is quite rare).

(NHS: Cancer of the breast, female)

In seven PIS I discovered an extraordinary step. After identifying and listing symptoms, the authors downplayed the aforementioned by saying that the symptoms do not necessarily imply cancer X, and that other diseases are likely to cause similar symptoms.

(24) All of these symptoms **can** also be caused by conditions other than prostate cancer. <u>You</u> **should** discuss them with <u>your</u> doctor.

(The Better Health Channel: Prostate cancer)

In example (24) there is a realisation of Step 4.4 "Encouraging patient to go and see/ discuss this issue with his/her doctor". This step is featured in 55% of the PIS that were investigated for this study. By using the personal pronoun *you* and the possessive pronoun *your* the reader is directly addressed, and so a relationship between reader and author is established. In addition, the reader is advised, encouraged and sometimes urged to visit his/her doctor in order to clarify possible issues about that topic. Common ways of realising Step 4.4. are presented below:

Consult your doctor, if …
See your doctor (right away), if …
You should see your doctor, if …
You should discuss … with your doctor
Talk to your doctor about …

Table 34. Lexical phrases frequently used in Step 4.4

7.3.5 Move 5: Introducing causes or risk factors/ groups

Move 5: Introducing causes or risk factors/ groups
Step 5.1: Signalling the beginning of a new section by stating 'causes', 'risk factors' or 'risk groups'
Step 5.2: Stating that causes/risks are (not clearly) known
Step 5.3: Listing or outlining the factors that increase the risk
Step 5.4: Referring to risk factors of a specific group

Table 35. The structure of Move 5

The introduction of causes, risk factors or even risk groups is an integral part of the genre of PIS. Causes are usually introduced to convey a complete picture of the disease. Thereby patients are enabled to understand and cope with their disease. In addition, they are empowered to change their lifestyles on the basis of their newly-gained knowledge. Move 5 is an obligatory move as it is featured in 90% of the PIS. As a first step, authors usually write a heading that introduces and signals a topic change. Another step is to state that causes or risk are not clearly known. This generic element appears in 55% of the texts. Step 5.3 is more common. Risk factors or causes are either listed or part of the running text. Step 5.4 is a core step. In about 50% of the PIS a reference to risk factors of a specific group is made. The borders between Steps 5.3 and 5.4 are not clear-cut, as a combination of these two elements can often be found. In example 25, Steps 5.3 and 5.4 are combined; and in addition, the form of a bulletin list was chosen by the author(s) to refer to the different causes and risk factors. In the first sentence *are thought to* is used to downplay the reader's expectations. In the following segment the individual risk factors are outlined. In brackets more information or explanations are offered. This strategy is common in PIS and is not particular to a specific move.

(25) Various factors are thought to increase the risk of breast cancer. These include:

- Having a close relative with a history of breast cancer (about 10% of cases are thought to have a genetic basis)
- Early start to menstruation (before the age of 11) and late menopause (after age 54)
- Having a first child late […]

(NHS: Cancer of the breast, female)

7.3.6 Move 6: Introducing diagnostic procedures

Move 6: Introducing diagnostic procedures
Step 6.1: Signalling the beginning of a new section by stating 'diagnosis'
Step 6.2: Introducing diagnostic procedure A
Step 6.3: Describing function and/or procedure of A This is combined with Step 6.4
Step 6.4: Outlining exact procedure and instruments involved
Step 6.5: Introducing diagnostic procedure B (Steps 6.3 and 6.4 are repeated)
Step 6.6: Introducing diagnostic procedure C (Steps 6.3 and 6.4 are repeated)
Step 6.7: Explaining that if cancer is diagnosed, other tests will have to be carried out

Table 36. The structure of Move 6

Move 6 appears in 90% of the PIS that were analysed in this study. The overall communicative purpose of the move is to introduce and outline different diagnostic procedures to the layperson readership. Procedures, instruments and agents involved are outlined. Move 6 contains a range of interesting language functions, such as describing. Within the written genre of PIS and the spoken genre of doctor-patient consultation the language function of description plays a central role. Authors of PIS describe objects (instruments), body parts, processes (diagnostic procedures, processes in the human body) to their readers via the medium text, whereas doctors describe the same issues to their patients during face-to-face consultations. It is within a doctor's or a PIS author's best interests to describe things as clearly as possible. It is also within the patient's best interests to find out how specific tests and processes in the human body work. Not much has been written about this language function, although describing is something that is done in many professions and therefore should be of interest to ELT and ESP researchers and teachers. Pettinari (1988) has explored reporting styles, and argues:

> It is important to note, […], that, in procedural discourse, varieties of complexities exist, from […] straight-forward step-by-step description of how a procedure should be done to more complex composites of process plus other rhetorical elements such as description, classification, evaluation and/ or definition. (Pettinari 1988:98)

Pettianri (1988) focused on changing reporting styles of first and last year surgical residents in operative reports. In her point of view, reports are narratives of procedures and descriptions and thus comprise an additional rhetorical element. Anatomical observations are usually described (Pettinari 1988: 98). Interestingly,

patient information sheets and operative reports share a number of common features. Both have narrative elements and comprise descriptions of e.g. body parts. They differ insofar as operative reports concern specific past events and processes, whereas patient information sheets offer general information on processes. Information about descriptions of processes realised in PIS is presented below:

(26) A specialist may perform a virtual colonoscopy (also known as a CT colonography). **This** uses X-rays to build up a series of images of your colon. A computer **then** organises **these** to create a detailed picture **that** may reveal the presence of polyps, and anything else that appears unusual on the surface of your colon.
Having an X-ray can **also** diagnose rectal cancer. A fluid called barium will be placed into your bowel, via your back passage, in order to highlight on an X-ray any abnormalities in tissue growth. **This** procedure is known as barium enema.
If you are diagnosed with colon, cancer, you may need further tests to establish the size and position of the cancer. […]

(CKS: Cancer of the colon, rectum or bowel)

Example (26) displays almost all steps that were identified in table 36. In the first paragraph, diagnostic procedure A and its function are introduced. In addition, the exact procedure and instruments involved are outlined, before turning to diagnostic procedure B. In the last paragraph step 6.7 'Explaining that if cancer is diagnosed, other test will have to be carried out' is realised. It is therefore possible to identify a step-by-step description. The procedures are outlined chronologically, and instruments as well as agents are incorporated into the description in order to guarantee that a layperson can follow the description. In order to establish a logical consequence, cohesive devices are used.

In terms of tenses, the present simple tense is often used to describe how individual diagnostic procedures are performed. Active as well as passive constructions are used to describe and explain procedures. Only when a reference is made to the reader and the procedures is s/he likely to face the term *will be* (as in example 26).

Particular to Move 6 and other moves that focus on the description of processes is the use of discourse markers. In PIS certain discourse markers are employed to establish a logical sequence, whereas others seem to be avoided. Generally, among the different types of discourse markers, there are those that indicate causes and results and those used for numbering and ordering (Parrot 2000: 303-306).

Common discourse markers
So (17x), then (17x)
Less common discourse markers
Therefore (6x), first (1x), hence (1x), in that case (1x), next (1x)
Discourse markers not appearing at all
Consequently, as a result, thus, …

Table 37. Discourse markers used for establishing logical sequence

What is striking is that less formal discourse markers such as *so* and *then* are preferred to more formal ones such as *therefore* or *consequently*. This can be explained by the fact that formal expressions seem to be avoided by authors of PIS. Another way of establishing logical sequence is to use if-clauses, as illustrated in example (26). The reader is told that a positive diagnosis with cancer of the colon may go hand-in-hand with further tests.

7.3.7 Move 7: Introducing treatment options

Move 7: Introducing treatment options
Step 7.1: Signalling the beginning of a new section by stating ‘treatment’
Step 7.2: Introducing (one/main) treatment option(s) by listing them

Table 38. The structure of Move 7

In 55% of the PIS, treatment options are introduced. This is a surprisingly low number. However, it has to be said that PIS providers often offer separate PIS on specific treatment options such as chemotherapy or radiotherapy. In all of the PIS that had a section on treatment the topic was introduced by using a headline stating ‚treatment’ or ‚treatment options’. Another step is to list the main treatment options. Authors of PIS refer to treatment options in order to highlight the fact that cancer can

be treated in a number of different ways. If the different treatment options are outlined in a more detailed way, the steps are similar to those employed in Move 6 'Introducing diagnostic procedures'.

Step 7.1: Signalling the beginning of a new section by stating 'treatment options'
Step 7.2: Introducing treatment option A
Step 7.3: Describing function and/or procedure of A This is combined with Step 6.4
Step 7.4: Outlining exact procedure and instruments involved
Step 7.5: Introducing treatment option B (Steps 6.3 and 6.4 are repeated)
Step 7.6: Introducing treatment option C (Steps 6.3 and 6.4 are repeated)
Step 7.7: Suggesting that this topic will be dealt with during doctor-patient consultation

Table 39. Optional steps

To illustrate how a treatment option is described, I will offer an example. In terms of language functions, tenses and cohesive devices, Move 7 is similar to Move 6. This is also true for vocabulary range. In Moves 6 and 7 the number of technical words and especially subject-specific words (category 1) is striking. Academic words are also used. In example (27) the subject-specific words are highlighted and the academic words are underlined.

(27) **Colorectal cancer** is usually treated by a **surgical operation** to remove the affected segment of the **bowel**. The free ends are then joined together, if this is possible. It may be necessary to remove the **tumour** and some of the surrounding **tissue**. When the **cancer** is very low in the **rectum**, the **anal canal** must also be removed to prevent the **cancer** coming back. In that case an **artificial opening** (**colostomy**) is necessary, bringing the **bowel** out through the front wall of the **abdomen**.

(CKS: Bowel cancer)

In example (27) different technical word types occur. Two academic word families appear in the example. As the example sentence deals with surgery, the verb *to remove* is used. Other verbs that are used in the 'treatment section' are presented below. Since procedures are described, the passive voice is often used.

Something is treated by (e.g. surgery, chemotherapy)
given intravenously
removed
placed (directly) in (e.g. the prostate)

Table 40. Verbs frequently used in Move 7

7.3.8 Move 8: Introducing the topic of prevention

Move 8: Introducing the topic of prevention
Step 8.1: Signalling the beginning of a new section by stating 'prevention' or 'how to reduce/minimize risks'
Step 8.2: Giving advice on how to reduce risks of developing a specific cancer

Table 41. The structure of Move 8

The topic of cancer prevention is raised in the majority of PIS that were explored (70%). It is usually introduced by a heading that states either ‚prevention' or ‚reducing/ minimizing the risk'. However, it should be highlighted that this section also seems to be dependent on the type of cancer. In the case of PIS on the subject skin cancer, lists with bullet points are used to structure the advice on how to reduce the risk of skin cancer and protect the skin from the sun, while in other PIS these moves are combined with Move 6 'Introducing diagnostic procedures'. Prevention then is related to cancer screening. To illustrate how Move 8 is realised, I will present an excerpt of a PIS on skin cancer:

(28) […]

- **Avoid** the use of tanning lamps
- **Never let** young children stay in the sun for long periods, even when wearing sunscreen. […]

(Health Canada: Preventing skin cancer)

In PIS, advice is given in different ways. Although *you* is frequently used to address the reader directly, it is sometimes omitted. *You* is instead implied. Authors use imperatives to give advice and orders, and also to make suggestions (example 28). If *you* is used, it is used in combination with modal verbs. Modal verbs associated with the language functions of advising and suggesting are *could* and *should*, but also *may* and *can* (Parrott 2000: 120).

7.3.9 Move 9: Offering more information

Move 9: Offering more information
Step 9.1: Signalling the beginning of a new section by stating 'other organisations', 'selected links', or 'want to know more?'
Step 9.2: Offering contact information and/or referring to other organisations, articles, websites and/or doctors

Table 42. The structure of Move 9

Authors of PIS often offer additional information to the reader. In all the PIS that were explored, Move 9 was identified. It comprises two steps, the first one signals the reader at the beginning of section and while the second one highlights different sources of information. Patients and family members can find contact information and/or websites of other organisations. Strategies for offering more information are (a) listing contact information of the institution, (b) offering a list of other institutions and their contact information such as addresses, telephone number or websites, (c) referring to other publications or (d) advising the reader to get into contact with his/her general practitioner, who in turn will be able to offer more information or to clarify issues. Phrases used to express Step 9.2. include *for further information see [...]* and *to find out more, contact [...].*

7.3.10 Move 10: Stating copyright statement

Finally, I will raise the issue of intellectual property. Copyright statements are provided in order to identify the copyright holder.

7.4 A word on optional moves and steps

As has been mentioned in the opening section of this chapter, only core and obligatory moves are dealt with in more detail since the scope of this paper is limited. However, during the course of the analysis I have identified interesting optional moves or rather steps that are well worth discussing. One such step is to offer and provide visual aids alongside definitions or descriptions. Notably, this step was only found in 20% of the PIS that were investigated. This is a very low percentage. I have come to the conclusion that this should be an obligatory step, since visuals enable laypeople to follow definitions and descriptions more easily, and thus enhance comprehension. In order to exemplify how visual aids are designed and used, I have included one illustration. It is taken from the PIS on bowel cancer provided by the *CKS* (*NHS*). The illustration shows the stomach and the surrounding digestive system.

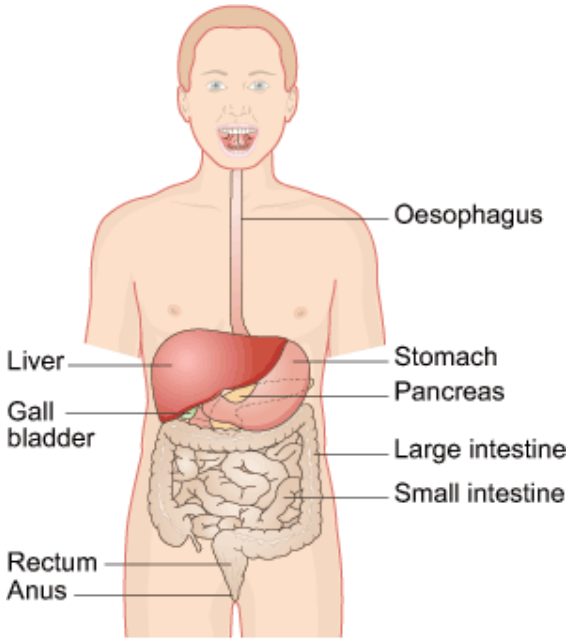


Figure 6: Realisation of "Offering visuals to enhance comprehension" (CKS: Bowel cancer)

The illustration is a realisation of what I refer to as "Offering visuals to enhance comprehension". In the picture the different parts of the digestive systems are labelled in order to show the patient (a) what body parts are involved when digesting and (b) where these are located in the body. In addition, it is implied that cancer of the bowel (stomach) and the treatment of it also may have effects on other parts of the digestive system. Illustrations such this allow laypeople to follow definitions and

descriptions more easily, therefore I have concluded that the step "Offering visuals to enhance comprehension" should be included in PIS.

Another interesting optional move was found in the PIS provided by *The Better Health Channel*. After Move 9 „Offering more information" the main facts were summarised. This move is beneficial for the reader since it captures and stresses the key issues in just a few words. Furthermore, the reader can return to the summary at a later point in time, without having to read the whole sheet again. In terms of the structure of this move, the beginning of the new move is often signalled with a headline, in this case, "Things to remember". I will now present an example of this type of lists, which contains the most important facts. The optional move "Summarising the key issues, as I have termed it, is shown in the example below:

(29) Things to remember

- Breast cancer affects one in 11 Australian women.
- It is important for all women to get to know the normal look and feel of their breasts.
- Although most breast changes aren't caused by breast cancer, you should always consult your doctor if you notice an 'unusual' change.
- Treatment options for breast cancer include surgery, radiotherapy, chemotherapy and hormone treatment.

(The Better Health Channel: Breast cancer)

Finally, I would like to refer to another optional move that I consider useful. This is the move "Providing questions patients may want to ask during consultation". When a patient is diagnosed with cancer, s/he is perhaps confused and under considerable emotional stress and therefore is sometimes too shocked, irritated or even intimidated to pose questions concerning her/his condition. Therefore, it is extremely helpful to offer questions they might wish or even should ask during a separate consultation. Although this move is extremely useful, it has been classified as optional since it was only featured in one (!) PIS investigated for this study. In the following segment I will demonstrate how this move is realised in the PIS that featured it.

(30) **Questions you may wish to ask**

1. What type of cancer do I have?
2. How extensive is my cancer? What stage is it?
3. What treatment do you advise for my cancer and why?
4. Are there other treatment choices for me?
5. What are the risks and possible side effects of each treatment?
6. What will the scar look like?

[…]

If you receive answers you do not understand, feel comfortable in saying: 'Can you explain that again.'

'I am not sure what you mean' or

'Would you draw a diagram or write it down'.

(Cancer Society of New Zealand: Breast cancer)

In terms of structure, the headline signals the beginning of a new section. There then follows a list of questions, which are ordered according to various topics, such as diagnosis (*What type of cancer do I have?*), decision on treatment (*Are there other treatment choices for me?*) and life after treatment (*How frequent will my check-ups be and what will they involve?*) etc. (Cancer Society of New Zealand: Breast Cancer). Directly after the section on questions the patients are encouraged to intervene and interrupt a consultation as soon as they feel unable to understand the doctor's answer or explanation. For this purpose, sample sentences (*Can you explain that again*) are offered (Cancer Society of New Zealand: Breast cancer). To sum up, all the moves that were introduced in this section share two characteristics. Firstly, they are featured in only a small percentage of the PIS in the corpus, and secondly, they should occur in a greater amount of these texts since they are extremely useful and can offer patients guiding questions for their consultations with doctors. They also enhance comprehension and provide a summary of the most important issues.

7.5 Summary of findings

In this chapter I have looked at the communicative purposes and the macro-structure of PIS. As far as the overall communicative purpose is concerned, it can be said that PIS should be informative. Firstly, they should be informative so as to provide patients enough information to enable them to make a decision about a specific diagnostic procedure or treatment. Secondly, PIS are informative in order to prevent the hospital from being sued. In signing a consent form, a patient usually agrees to a specific diagnostic procedure or treatment and thereby declares that they have read the PIS thoroughly and are therefore aware of possible risks and dangers.

In terms of the move structure, the results of the analysis imply that people who are at risk of a specific cancer should be informed about the disease, the early warning signs and symptoms, the different diagnostic procedures and treatment options. Further references to risk reduction and further sources of information should also be included in PIS or consultations. In the course of the move structure analysis, references were made to the language functions found in some moves. Among the most prominent language functions are defining, explaining, describing, informing (on body parts, body functions, diagnostic procedures, instruments and other equipment) and giving advice (on how to reduce risks and change lifestyle). The analysis also showed that illustrations, summaries and questions sections are not common in PIS. However, visual aids as well as other guiding sections should be ascribed value as they enhance understanding.

Additionally, the findings of the analysis suggest that different strategies are used to engage the reader. Strategies include directly addressing the reader and also downplaying the reader's expectation e.g. regarding symptoms. A useful and effective strategy is that of addressing the patient directly. This is done for several reasons. Firstly, by addressing the reader directly, a bond between the patient and the author develops; secondly, the patient is more closely engaged; and thirdly, the patient is enabled to realise that s/he is responsible for her/his health and wellbeing, and is thus encouraged to actively participate in the healthcare process. Another interesting strategy found in PIS is the one of downplaying the reader's expectations. This clearly goes hand-in-hand with what has been previously written about avoiding lawsuits. In addition to this, patient disappointment and anger are prevented.

7 Implications for learning and teaching

In the previous chapters, the focus has been on Bhatia's (1993) sixth step of analysing genre. Therefore, the both the vocabulary and generic structure of PIS were investigated. Valuable results were gained. However, the question that should be posed now is how to exploit this data for teaching and learning purposes. Therefore, I will now consider the implications for learning and teaching. According to Henry & Roseberry (2001: 117),

> [i]nformation from genre analysis […] can form the basis for two kinds of teaching materials: lessons, and source materials, which students can use as a source of more data.

Henry & Roseberry's (2001) "lessons" are sessions that are primarily aimed at familiarising students with a particular genre. Drawing from their own experiences of teaching genres to students, Henry & Roseberry (2001) outline the procedure they followed when teaching genre-based lessons. Lessons, according to them, can be divided into four parts: (1) introducing the genre, (2) reading and analysing the moves, (3) learning the language of moves and (4) constructing the genre (Henry & Roseberry 2001: 117). When introducing a genre to students it is important to provide some background information, including the social and cultural setting of the genre, the producers and the receivers of it, as well as the main purpose of the genre (ibid.). Following the introduction is a stage titled 'reading and analysing the moves'. According to this, students should read, recognise and analyse moves. If students are not familiar with the principles of move-structure analysis, they should be providing with a table of moves. If they are already knowledgeable, students should be encouraged to identify moves themselves. The results of their analyses then have to be discussed and compared. Students can then be asked to look at how moves are realised. In the third step, the focus should be on the language. Therefore, students could be asked to do group work in order to find out which words and tenses are commonly used and how e.g. politeness is expressed. This step is important since it prepares student for the last step, which is "constructing the genre". Here, students are asked to produce their own examples of the relevant genre (Henry & Roseberry 2001: 117-120). Interestingly, this four-part procedure advocated by Henry & Roseberry (2001) shares characteristics with the 'Martin Model' of genre introduced

in section 2.2. Broadly speaking, the four-part procedure used by Henry & Roseberry (2001) as well as the other model can be adapted for the teaching of PIS. However, two aspects should be taken into account. Firstly, genre conventions can change. This means teachers should

> present such genres not as specific norms to be conformed to, but as more general points of reference within which room for manoeuvre is possible. (Widdowson 2003:70)

Secondly, the fourth part "constructing the genre" has to be realised differently when working with the genre of PIS. Aspiring medical personnel are not often in a position where they have to produce PIS. Nevertheless, they are likely to use language functions that can be found in the moves of the genre, such as defining, describing, explaining or giving advice when communicating with patients. Therefore in this context "constructing the genre" does not mean writing PIS but rather simulating doctor-patient communication and situations where procedures have to be defined and explained. Of course, the learners' awareness should be raised in regard to differences between written (PIS) and spoken (consultations) genres and also to the fact that simplified language should be used when communicating with a patient.

In order to illustrate how moves or rather language functions can be explored and practised in the language classroom, I will present an excerpt from an in-print course book on Medical English, *English in Medicine* (Glendinning & Halmström 1998). In this excerpt, the language function of explaining is introduced to the learner. Special emphasis is attached to the fact that non-specialist language should be used when explaining diseases to patients. As a first task the language learners are encouraged to compare an explanation taken from a textbook to one taken from a patient-doctor consultation. Attention is then drawn to the differences between those explanations, and in particular to the way that the doctor describes the medical concept. As a follow-up exercise the learners are then asked to produce explanations.

Language focus 16

When explaining a diagnosis, a patient would expect you to answer the following questions:

1 What's the cause of my problem?
2 How serious is it?
3 What are you going to do about it?
4 What are the chances of a full recovery?

In Unit 7, we will deal with questions 3 and 4. Here we will look at some of the language used to answer questions 1 and 2.

In explanations it is important to use straightforward, non-specialist language with only such detail as is important for the patient's understanding of the problem. The language of the textbooks you may have studied is clearly unsuitable for patient explanation. Compare this extract with the recorded explanation in Task 7.

> Herniation of part of a lumbar intervertebral disc is a common cause of combined back pain and sciatica ... Part of the gelatinous nucleus pulposus protrudes through a rent in the annulus fibrosus at its weakest part, which is postero-lateral ... If it is large, the protrusion herniates through the posterior ligament and may impinge upon an issuing nerve to cause sciatic pain.
>
> (J. C. Adams, *Outline of Orthopaedics*, 10th ed. (Edinburgh: Churchill Livingstone, 1986), p. 217.)

You can make sure your explanations are easily understood by avoiding medical terminology where possible and defining the terms you use in a simple way. Note how the doctor describes a disc:

–The disc is a little pad of gristle which lies between the bones in your spine.

Task 8 ✎

Write simple explanations for patients of these terms. Compare your explanations with those of other students.

1 the pancreas
2 the thyroid
3 fibroids
4 emphysema
5 arrhythmia
6 bone marrow
7 the prostate gland
8 gastro-oesophageal reflux

Figure 7: 'Explaining' in English in Medicine (Glendinning & Halmström 1998: 68-69)

As mentioned at the beginning of the chapter, the materials gained from a genre analysis can also serve as a basis for source materials "students can use as a source of more data" (Henry & Roseberry 2001: 117). The underlying idea of this approach is to let the learners explore the genre themselves with the aid of computer software. According to Henry & Roseberry (2001), two principles should be considered when

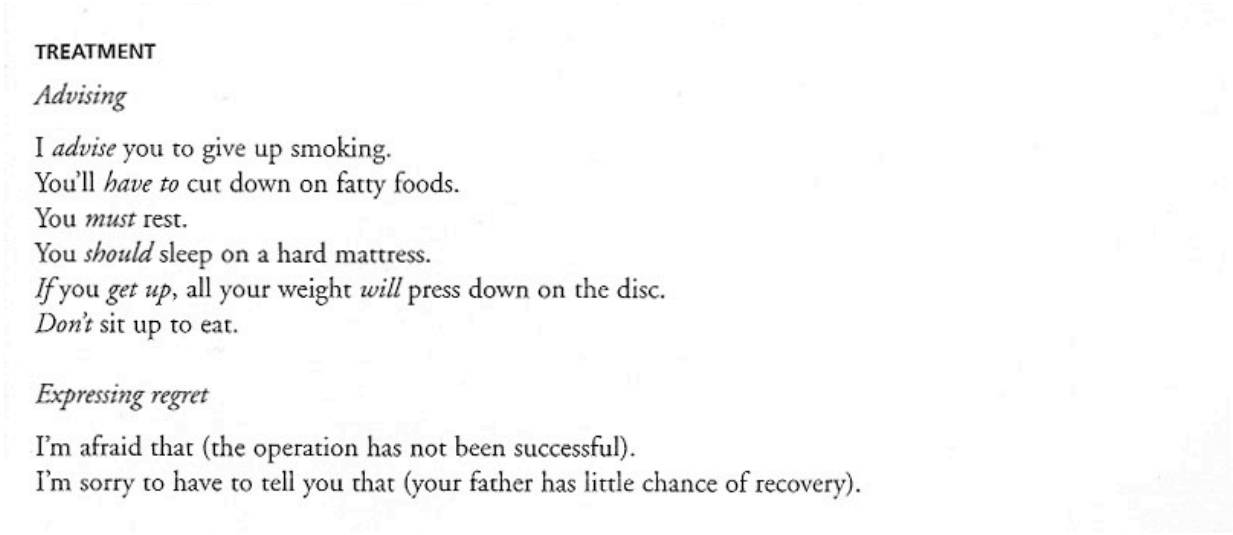
deciding to let students investigate genres. First of all, the data has to be organised and/or selected according to set criteria, in order to make it more manageable and less confusing for the learners. More importantly, intimidating learners by introducing all the data in printed form should be avoided. It is advised instead to provide the data electronically before issuing student with the printed version.
The data in printed form could be introduced at a later time, when the focus is on specific aspects of the data (Henry & Roseberry 2001: 121). Henry & Roseberry (2001) suggest working with a computerised branching program. This enables the researcher to organise the genres into moves and steps, and to show how the individual generic elements are realised. The learners then explore the genre, in the sense that they investigate a specific move by clicking onto it in order to see how this move is realised by authors of the genre (Henry & Roseberry 2001: 120-126). The data therefore functions as a source of information for the learners. Henry & Roseberry (2001: 125) picture the following scene:

> Students use the computer laboratory or their own computers at home, drawing on the computerised materials as a resource while studying and writing examples of the genre.

In an EMP setting, this way of exploiting corpus-based genre data would be useful, especially in regard to genres such as doctor-patient communication, medical reports and even abstracts and articles. In regard to the genre of PIS, I am not fully convinced about this, since the learners are not likely to produce PIS themselves. However, learners would be empowered to explore the differences between the written genre of PIS and the spoken genre of doctor-patient consultation.

So far, I have looked at how information drawn from genre analysis can be used in lessons and as source material by language learners. The data obtained from a corpus-based genre analysis can also be used for ESP course books and syllabus design (Flowerdew 2001: 79). For example, it can be used to design a reference section or unit in course books. In the course book *English in Medicine* (Glendinning & Halmström 1998), a summary of all the language functions covered in the book is offered in the appendix. This information is intended to function as a source of information for learners. In the following excerpt, language functions linked to the

topic of treatment are presented, and sample sentences are offered. In these sample sentences, advising verbs and modal verbs used for advising are highlighted:



TREATMENT

Advising

I *advise* you to give up smoking.
You'll *have to* cut down on fatty foods.
You *must* rest.
You *should* sleep on a hard mattress.
If you *get up*, all your weight *will* press down on the disc.
Don't sit up to eat.

Expressing regret

I'm afraid that (the operation has not been successful).
I'm sorry to have to tell you that (your father has little chance of recovery).

Figure 8: Language functions linked to the topic of treatment in English in Medicine (Glendinning & Halmström 1998: 139)

In summary, there are various ways for exploiting genre-based data for learning and teaching. Whether teachers, learners or course book designers should be exploring the genre is a matter of personal preference. The fact is that the concept of genre should be explored in the language classroom, since it:

> provides a way of looking at what students have to do linguistically – what kinds of discourse they have to be able to understand and produce in speech and writing. It also provides us with an understanding of why a discourse is the way it s, through a consideration of its social context and purpose. (Kay & Dudley-Evans 1998: 310)

In regard to the fact that genres are useful tools in the language classroom, I can say that medical genres should have a more prominent role in the teaching and learning of EMP.

8 Conclusion

In this paper I have aimed to highlight the linguistic and communicative characteristics of patient information sheets (PIS) that are relevant to learners and teachers of Medical English. In order to achieve this aim, a genre-based corpus analysis was conducted. Several PIS were therefore compiled, according to a specific set of criteria, in order to be analysed in terms of vocabulary and generic structure. The first step of the analysis consisted of a computer-assisted analysis of the vocabulary found in patient information sheets. One of the research aims of this study was to show the coverage of general, academic and technical languages in PIS. This aim was achieved insofar as it has been shown that academic as well as technical words constitute a considerable portion of the words found in PIS, although vocabulary from the general language dominates the genre. Most importantly, the findings of the vocabulary analysis suggest that PIS are more academic and technical than academic texts. The multitude and range of technical vocabulary found in patient information sheets implies that learners and teachers of English (as well of readers of these types of sheets) may have difficulties reading and understanding them. Subject-specific and academic vocabulary has to be included in the English curriculum of medical personnel, as well as expressions from general English and everyday English.

In this paper, I also gave an account of the communicative purposes and the generic structure of PIS. In terms of the communicative purpose of PIS, the conveyance of information has great value. In other words, PIS should be informative, in the sense that they should provide the patients with information that can allow them to make decisions on treatments. Another function of PIS is that they operate as a legal backing. One of the most significant findings of my investigation is the move structure of PIS. In this paper, a total of 10 moves were identified that contribute to the overall communicative purpose of the genre texts. It has been shown that these generic elements characteristically address issues such as the disease itself, its warning signs and symptoms, ways of diagnosing it, its treatment and prevention. Inherent in the individual generic elements are language functions. These, such as defining, describing, explaining and giving advice in PIS, play key roles in the teaching and learning for EMP.

Notably, the genre analysis also suggested that the genre investigated was actually a sub-genre. Patient information sheets may address studies, medication, treatments and diseases. The texts analysed in this study focused on the diseases of breast cancer, prostate cancer, skin cancer and bowel cancer; and should be regarded as belonging to the genre of patient information sheets or more specifically to the sub-genre of PIS addressing diseases. Generally speaking, I can conclude that the overall research aims were achieved, to the extent that vocabulary and the move structure of patient information sheets were analysed. The findings of the study can be used for syllabus or course material design. Since this study was performed on a small-scale, it would be of interest to conduct a larger study, which would take into consideration the different sub-genres of PIS. A larger study would also be beneficial, as the results would be more representative. The findings would also allow more generalisations.

Participants of EMP courses work in different wards and departments, and therefore engage in different fields of medicine. In addition, the motivation to learn the English language differs from person to person. There could, for instance, be a doctor wanting to improve his English in order to be able to participate at international conferences or to publish journal articles. Courses and course books on Medical English, however, often only cover ‚general' issues without taking into account specialised needs. Practice could therefore be improved by investigating and exploring various spoken and written genres in the field of EMP. Furthermore, it would be interesting to address and analyse the role of 'patient talk' in doctor-patient communication. During consultations, patients describe their symptoms, health conditions and body parts. They often use slang expressions, child talk and euphemisms, thus the role of colloquial language in doctor-patient communication is also worth exploring. Accordingly, further research and course book publications are needed in field of EMP.

References

Anderson, Robert M. 1995. "Patient empowerment and the traditional medical model". *Diabetes Care* 18, 412-415.

Anderson, Robert M.; Funnel, Martha .M.; Barr, Patricia A.; Dedrick, Robert F.; Davis, Wayne K. 1991. "Learning to empower patients, results of professional education program for diabetes educators". *Diabetes Care* 14, 584-590.

Aston, Guy. 2001. "Learning with corpora: an overview". In Aston, Guy (ed.). *Learning with corpora*. Bologna: Cooperative Libraria Universitaria Editrice and Houston TX: Athelstan, 7-45.

Austrian Federal Ministry for Education, the Arts and Culture/ General Directorate for Vocational Education and Training (ed). 2007a. *Information Sektion Berufsbildung: Englisch*. Graz. 11 June 2008 <http://www.berufsbildendeschulen.at/de/download.asp?id=26&theme=Information%20Sektion%20Berufsbildung:%20Englisch%20(2007)>.

Austrian Federal Ministry for Education, the Arts and Culture/ General Directorate for Vocational Education and Training (ed). 2007b. *Schulen für Sozialbetreuungsberufe / Lehrplan*. Vienna. 11 June 2008 <http://www.berufsbildendeschulen.at/de/download.asp?id=13&theme=Lehrpl%E4ne:%20Sozialberufe>.

Austrian Federal Ministry for Health, Family and Youth. 2008. *Ausbildungseinrichtungen in Gesundheitsberufen; Training for health professions.* Vienna. 11 June 2008 <http://www.bmfgj.gv.at/cms/site/standard.html?channel=CH940>.

Barber Charles L. (1962). "Some measurable characteristics of modern scientific prose". In *Contributions to English syntax and philology*. *Gothenburg Studies in English* (14), 21-43. (Reprinted in Swales, John. 1988. *Episodes in ESP*. Hertfordshire: Prentice Hall, 1-16).

Basturkmen, Helen. 2006. *Ideas and options in English for specific purposes*. Mahwah, NJ: Erlbaum.

Bazerman, Charles. 1988. *Shaping written knowledge: the genre and activity of the experimental article in science.* Madison: University of Wisconsin Press.

Bazerman, Charles. 1994. "Systems of genres and the enhancement of social intentions". In Freedman, Aviva; Medway, Peter (eds.) *Genre and the New Rhetoric*. London: Taylor and Francis, 79-101.

Berkenkotter, Carol; Huckin, Thomas N. 1995. *Genre knowledge in disciplinary communication – cognition / culture / power*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bhatia, Vijay. 1993. *Analysing genre*. Harlow: Pearson Education.

Bhatia, Vijay. 2004. *Worlds of written discourse: a genre-based view*. London: Continuum.

Bowker, Lynne; Pearson, Jennifer. 2002. *Working with specialized language: a practical guide to using corpora*. London: Routledge.

British National Corpus. 21 Feb. 2008 <http://www.natcorp.ox.ac.uk>.

Brown, Penelope; Levinson, Stephen. 1987. *Politeness: some universals in language usage.* (2nd edition). New York: Springer.

Carter, Ronald. 1998. "Orders of reality: CANCODE, communication and culture". *ELT Journal* (52), 43-56.

Chumley, Gilian M.; Hall, George M.; Salmon, Peter. 2002. "Patient-controlled analgesia: what information does the patient want?" *Journal of Advanced Nursing* 39 (5), 459-471.

Chung, Teresa M.; Nation, Paul. 2003. "Technical vocabulary in specialised texts". *Reading in a Foreign Language* 15 (2), 103-116.

Chung, Teresa M.; Nation, Paul. 2004. "Identifying technical vocabulary". *System* 32 (2), 251-263.

Concise medical dictionary. 2003. Online edition. Oxford: Oxford University Press. Oxford References Online. 18 June 2008 <http://www.oxfordreferences.com>.

Cook, Guy. 1998. "The uses of reality: a reply to Ronald Carter". ELT Journal 52, 57-63.

Cope, Bill; Kalantzis, Mary. 1993. "Introduction: how a genre approach to literacy can transform the way writing is taught". In Cope, Bill; Kalantzis, Mary (eds.). *The powers of literacy: A genre approach to teaching writing*. Bristol. PA: Falmer Press, 1-21.

Coxhead, Averil. 2000. "A new academic word list". *TESOL Quarterly* 34 (2): 213-238.

DIN 2330. 1979. *Begriffe und Benennungen (Allgemeine Grundsätze)*. Deutsche Normen, Berlin und Köln.

Devitt, Amy. 1991. "Intertextuality in tax accounting". In Bazerman, Charles; Paradis, James (eds.). *Textual dynamics of the professions*. Madison, WI: University of Wisconsin Press.

Dudley-Evans, Tony. 1994. "Genre analysis: an approach to text analysis for ESP". In Coulthard, Malcolm C. *Advances in written text analysis*. London: Routledge, 219-228.

Dudley-Evans, Tony; St. John, Maggie Jo. 1998. *Developments in English for Specific Purposes: a multi-disciplinary approach*. Cambridge: Cambridge University Press.

Eggly, Susan; Jusial, Joseph and Smulowitz, Jack. 1999. "Research and discussion note: the relationship between English language proficiency and success as a medical resident". *English for Specific Purposes* 18 (2), 7pp. 25 Oct. 2007 <www.sciencedirect.com/science/journal/088949063>.

Fairclough, Norman. (1995). *Critical discourse analysis: the critical study of language.* London: Longman.

Fligelstone, Steven. 1993. "Some reflections on the question of teaching, from a corpus linguistics perspective". *ICAME Journal* 17, 97-109.

Flowerdew, John. 2001. "Concordancing as a tool in course design." In Ghadessy, Mohsen; Henry, Alex and Roseberry, Robert L (eds). *Small corpus studies and ELT: theory and practice.* Amsterdam: Benjamins, 71-92.

Flowerdew, Lynne. 2005. "An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticism against corpus-based methodologies". *English for Specific Purposes* 24 (3), 11 pp. 18 Oct. 2007 <http://www.sciencedirect.com/science/journal/08894906>.

Frank, Ruth A. 2000. "Medical communication: non-native English speaking patients and native English speaking professionals" *English for Specific Purposes* 19, 31-62.

Freedman, Aviva; Medway, Peter. 1994. *Genre and the New Rhetoric*. London: Taylor and Francis.

Gavioli, Laura. 2005. *Exploring corpora for ESP learning.* Amsterdam: Benjamins.

Glendinning, Eric H.; Holmström, Beverly A.S. 1998. *English in Medicine*. (2nd edition). Cambridge: Cambridge University Press.

Halliday, Michael A.K. 1978. *Language as a social semiotic: the social interpretation of language and meaning*. London: Edward Arnold.

Halliday, Michael A.K.; Hasan Ruqaiva. 1985. *Language, context and text: aspects of language in a social-semiotic perspective*. London: Longman.

Halliday, Micheal A.K.; Martin, Jim R. 1993. *Writing science: literacy and discursive power.* London: Falmer Press.

Hasan, Ruqaiya. 1984. "Coherence and cohesive harmony". In Flood, James (ed.), *Understanding reading comprehension*. International Reading Association, Delaware, 181-219.

Hasan, Ruqaiya. 1989. “The structure of a text”. In Halliday, M.A.K and Hasan, Ruqaiya. *Language, context, and text: aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press. 52-69.

Heatley, Alex; Nation, I.S.P; Coxhead, Averil. 2005. *Instructions* (delivered with RANGE and FREQUENCY programs). 21 Feb. 2008 <http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>.

Henry, Alex and Roseberry, Robert. L. 2001. “Using a small corpus to obtain data for teaching a genre”. In Ghadessy, Mohsen; Henry, Alex and Roseberry, Robert L. (eds.). *Small corpus studies and ELT: theory and practice.* Amsterdam: Benjamins, 93-133.

Higgins, John. 1966. “Hard facts: notes on teaching English to science students”. *ELT Journal* 21 (1), 55-60.

Hoekje, Barbara. 2007. “Medical discourse and ESP courses for international medical graduates (IMGs).” *English for Specific Purposes* 26 (3), 17pp. 18 Oct. 2007 <http://www.sciencedirect.com/science/journal/08894906>.

Hornby, Albert S. 1989. *Oxford advanced learner’s dictionary of current English.* (4th edition). Oxford: Oxford University Press.

Howatt, Anthony P.R. 1984. *A history of English language teaching*. Oxford: Oxford University Press.

Hunston, Susan. 2002. *Corpora in applied linguistics*. Cambridge University Press: Cambridge.

Hüttner, Julia I. 2007. “Approaching ESP texts.” University of Vienna. English Department. 6 Dec. 2007.

Hyon, Sunny. 1996. “Genre in three traditions: implications for ESL.” *TESOL Quarterly* 30, 693-722.

Johansson, Kirsi. 2005. “Preoperative education for orthopaedic patients: systematic review”. *Journal of Advanced Nursing* 50 (2), 212-223.

Johns, Ann M. 1997. *Text, role and context: developing academic literacies.* Cambridge: Cambridge University Press.

Kay, Heather; Dudley-Evans, Tony. (1998). “Genre: what teachers think”. *ELT Journal* 52 (4), 308-313.

Kennedy, Chris; Bolitho, Rod. 1984. *English for Specific Purposes*. London and Basignstoke: Macmillan Publishers Limited.

Krois-Lindner, Amy. 2007. “World of Work 2: a look at Legal, Technical and Medical English”. University of Vienna. English Department. 12 June 2007.

Lanham, Richard A. 1983. *Analyzing prose*. New York: Continuum.

Leech, Geoffrey. 1997. “Teaching and language corpora: a convergence”. In Wichmann, Anne; Fligelstone, Steven; McEnery, Tony; Knowles, Gerry (eds.). *Teaching and language corpora.* New York: Addison Wesley Longman Inc., 1-23.

Leino-Kilpi Helena. 1999. “Keynote speaker identifies need for research and empowerment among patients. *AORN Journal* 70, 863-877.

Leino-Kilpi, Helena; Luoto, Eija.; Katajisto, Jouko. 1998. “Elements of empowerment and MS patients.” *Journal of Neuroscience Nursing* 18, 116-123.

Leino-Kilpi, Helena; Mäenpä, Inger; Katajisto, Jouko. 1999. “Nursing study of the significance of rheumatoid arthritis as perceived by patients using the concept of empowerment.” *Journal of Orthopaedic Nursing* 3, 138-145.

Macken, Mary; Kalantzis, Mary; Kress, Gunther, Martin, Jim R.; Cope, Bill and Rothery, Joan. 1989. *A genre-based approach to teaching writing, years 3-6, Book 2: Factual writing: A Teaching Unit based on reports about sea mammals.* Directorate of Studies, NSW Department of Education, in association with the Literacy and Education Research Network, Sydney.

Martin, Jim R. 1984. “Language, register and genre”. In *ECT 418, Language Studies. Children Writing: Reader*. Victoria, Deakin University.

Martin, Jim R. 1985. “Process and text: two aspects of human semiosis”. In Benson, James; Greaves, William (eds.). *Systemic Perspectives on Discourse: selected theoretical papers from the 9th International Systemic Congress 1*. Norwood, NJ: Ablex, 248-74.

Martin, Jim R. 1989. *Factual writing: exploring and challenging social reality.* Oxford: Oxford University Press.

Martin, Jim R. 1992. *English text: system and structure.* Amsterdam: Benjamins.

Martin, Jim R.; Christie, Frances; Rothery, Joan. 1987. “Social processes in education: A reply to Sawyer and Watson (and others). In Reid, Ian (ed.). *The place of genre in learning: current debates.* Geelong, Australia: Deakin University Press, 46-57.

McFerran, Tanya. 2003. *A dictionary of nursing*. Oxford: Oxford University Press. Oxford Reference Online. 18 June 2008 <http://www.oxfordreference.com>.

Medical University of Vienna. 2008. *Wahlfach Medical English*. 11 June 2008 <http://www.uv-medizin.at/english/>.

Merriam-Webster’s Medical Dictionary. 16 June 2008 <http://www.merriam-webster.com>.

Miller, Carolyn R. 1984. “Genre as social action”. *Quarterly Journal of Speech* 70, 157-178. Also published in Freedman, Aviva; Medway, Peter (eds.). 1994. *Genre and the New Rhetoric*. London: Taylor and Francis, 23-42.

Moon, Rosamund. 1997. “Vocabulary connections: multi-word items in English”. In Schmitt, Norbert; McCarthy, Michael (eds.). *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press, 40-63.

Mungra, Philippa. 2007. “A research and discussion note: The macrostructure of consensus statements”. *English for Specific Purposes* 26 (1), 79-89.

Müller, Renate. 1993. *Phraseologismen in englischen Fachtexten der Humanmedizin: eine empirische Untersuchung zur Fachphraseologie.* Frankfurt am Main: Lang.

Nation, I.S.P. 2001a. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. 2001b. “Using small corpora to investigate learner needs”. In Ghadessy, Mohsen; Henry, Alex; Roseberry, Robert, L. (eds.). *Small corpus studies and ELT: theory and practice*. Amsterdam: Benjamins, 31-45.

Nation, Paul; Kennedy, Paul. 1994. “How useful is EAP vocabulary for ESP? – A corpus based case study.” *RELC Journal* 25 (2), 34-50.

Nattinger, James; DeCarrio, Jeanette. 1992. *Lexical phrases and language teaching.* Oxford: Oxford University Press.

Nestel, Debra; Kidd, Jane. 2004. “Teaching and learning about written communications in a United Kingdom medical school”. *Education for Health* 17 (1), 27-34.

ÖBIG (Österreichisches Bundesinstitut für Gesundheitswesen). 2003. *Offenes Curriculum für die Ausbildung in Allgemeiner Gesundheits- und Krankenpflege*. Wien.

Parrott, Martin. 2002. *Grammar for English language teachers.* Cambridge: Cambridge University Press.

Pettinari Catherine J. 1988. *Task, talk, and text in the operating room: a study in medical discourse*. Norwood, NJ: Ablex Publishing Company.

RANGE and FREQUENCY: computer programs. 21 Feb. 2008 <http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx>.

Renouf, Antoinette. 1997. “Teaching corpus linguistics to teachers of English”. In Wichmann, Anne; Fligelstone, Steven; McEnery, Tony; Knowles, Gerry (eds.). *Teaching and language corpora.* New York: Addison Wesley Longman Inc., 255-266.

Schmitt, Norbert. 2000. *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Scollon, Ron. 1998. *Mediated Discourse as social interaction – a study of news discourse*. London: Longman.

Scollon, Ron; Scollon, Suzanne. 1995. *Intercultural communication: a discourse approach*. Oxford: Basil Blackwell.

Scott, Mike. 2004. *WordSmith Tools: Online manual*. 8 Jan. 2008 <http://www.lexically.net/downloads/version4/html/index.html>.

Seidlhofer, Barbara. 2003. *Controversies in applied linguistics*. Oxford: Oxford University Press.

Sinclair, John. 1987. *Collins COBUILD English Language Dictionary*. London: Collins.

Sinclair, John. 1990. *Collins COBUILD English Grammar*. London: Collins.

Sinclair, John. 1997. "Corpus evidence in language description". In Wichmann, Anne; Fligelstone, Steven; McEnery, Tony; Knowles, Gerry (eds.). *Teaching and language corpora.* New York: Addison Wesley Longman Inc., 27-39.

Sökmen, Anita. 1997. "Current trends in teaching second language vocabulary". In N. Schmitt & M. McCarthy (eds.). *Vocabulary: description, acquisition, and pedagogy*. Cambridge: Cambridge University Press, 237-257.

Stubbs, Michael. 2000. "Society, Education and Language: The last 2,000 (and next 20?) years of language teaching." In Trappes-Lomay. Hugh R. (ed): *Change and continuity in applied linguistics: selected papers from the annual meeting of the British Association for Applied Linguistics held at the University of Edinburgh, September 1999*. Clevedon: British Association for Applied Linguistics, 15-34.

Strevens, Peter. 1973. "Technical, technological, and scientific English (TTSE)." *ELT Journal* 27 (3), 223-234.

Swales, John. 1971. *Writing scientific English: a textbook of English as a foreign language for students of physical and engineering sciences.* London: Nelson.

Swales, John. 1974. "Notes on the function of attributive en-participles in scientific discourse". *Papers for Special University Purposes* (1), ELSU, University of Khatoum.

Swales, John. 1981. "Definitions in science and law – evidence for subject-specific course components?" *Fachsprache* 81 (3), 106-112.

Swales, John. 1988. *Episodes in ESP*. Hertfordshire: Prentice Hall.

Swales, John. 1990. *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Swales, John; Feak, Christine. 1994. *Academic writing for graduate students: essential talks and skills; a course for non-native speakers of English*. Ann Arbor, Michigan: University of Michigan Press.

The academic wordlist. 21 Feb. 2008 <http://language.massey.ac.nz/staff/awl>.

Upton, Thomas A.; Connor, Ulla. 2001. "Using computerized corpus analysis to investigate the textlinguistic discourse moves of a genre". *English for Specific Purposes* 20, 313-329.

Viefhues, Herbert. 1999. "Medizinsche Ethik in einer offenen Gesellschaft". In Sass, Hans-Martin (ed.), *Medizin und Ethik*. Stuttgart: Philipp Reclam jun. Stuttgart, 17-39.

West, Michael. 1953. *General service list of English. London*: Longman.

Widdowson, Henry G. 1978. *Teaching language as communication*. Oxford: OUP.

Widdowson, Henry G. 1991. "The description and prescription of language". In Alatis, James (ed.) *Linguistics and language pedagogy: The state of the art.* Washington, DC: Georgetown University Press, 11-24.

Widdowson, Henry G. 1998. "Communication and community: The pragmatics of ESP". *English for Specific Purposes* 17 (1), 3-14.

Widdowson, Henry G. 2000. "On the limitations of linguistics applied". *Applied Linguistics* 22 (4): 3-25.

Widdowson, Henry G. 2003. *Defining issues in English language teaching*. Oxford: Oxford University Press.

Wolff, Hanns P. "Arzt und Patient". In Sass, Hans-Martin (ed.). *Medizin und Ethik.* Philipp Reclam jun. GmbH & Co: Stuttgart, 184-211.

WordSmith Tools. Vers. 4. Oxford: Oxford University Press. 8 Jan. 2008 <http://www.lexically.net/downloads/version4/html/index.html>.

Material

American Academy of Family Physicians

"Breast cancer: steps of finding breast lumps early". American Academy of Family Physicians. 5 Dec. 2007 <http://familydoctor.org/online/famdocen/home/women/reproductive/breast/018.html>.

"Colorectal cancer screening". American Academy of Family Physicians. 5 Dec. 2007 <http://familydoctor.org/online/famdocen/home/common/cancer/risk/556.html>.

"Melanoma: a kind of skin cancer". American Academy of Family Physicians. 5 Dec. 2007 <http://familydoctor.org/online/famdocen/home/common/cancer/types/666.html>.

"Prostate cancer: what you need to know". American Academy of Family Physicians. 5 Dec. 2007 <http://familydoctor.org/online/famdocen/home/common/cancer/types/361.html>.

Better Health Channel (BHC) – Victorian Government

"Bowel cancer". Better Health Channel. 28 April 2008 <http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Bowel_cancer?open>.

"Breast cancer". Better Health Channel. 28 April 2008 <http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Breast_cancer?open>.

"Melanoma". Better Health Channel. 28 April 2008 <http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Melanoma?open>.

"Prostate cancer". Better Health Channel. 28. April 2008 <http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Prostate_cancer?open>.

Cancer Society of New Zealand

"Breast cancer". Cancer Society of New Zealand. 28 April 2008 <http://www.cancernz.org.nz/InformationSheets/CancerInformationSheets/>. (Link: Breast Cancer pdf).

"Prostate cancer". Cancer Society of New Zealand. 28 April 2008 <http://www.cancernz.org.nz/InformationSheets/CancerInformationSheets/>. (Link: Prostate Cancer Information Sheet pdf).

"Bowel cancer". Cancer Society of New Zealand. 28 April 2008 <http://www.cancernz.org.nz/InformationSheets/CancerScreening/>. (Link: Bowel Cancer).

"Melanoma". Cancer Society of New Zealand. 28 April 2008 <http://www.cancernz.org.nz/HealthPromotion/SkinCancerControl/InformationSheets/>. (Link: Melanoma pdf).

Health Canada

"Breast cancer". Health Canada. 5 Dec. 2007 <http://hc-sc.gc.ca/iyh-vsv/disase-maladies/breast-sein_e.html>.

"Preventing skin cancer". Health Canada. 5 Dec. 2007 <http://hc-sc.gc.ca/iyh-vsv/disase-maladies/cancer_e.html>.

"Prostate cancer". Health Canada. 5 Dec. 2007 <http://phac_aspc.gc.ca/ccdc-cpcmc/topics/cancer_prost_e.html>.

"Screening for colorectal cancer". Health Canada. 5 Dec. 2007 <http://hc-sc.gc.ca/iyh-vsv/disase-maladies/colorectal_e.html>.

National Health Service (NHS) Direct (CKS)

"Cancer of the breast female". NHS Direct (CKS). 5 Dec. 2007 <http://www.cks.library.nhs.uk/patient_information_leaflet/cancer_of_the_breast_female>.

"Cancer of the colon, rectum or bowel". NHS Direct (CKS). 5 Dec. 2007 <http://www.cks.library.nhs.uk/patient_information_leaflet/cancer_of_the_colon_rectum_or_bowel>.

"Cancer of the prostate". NHS Direct (CKS). 5 Dec. 2007 <http://www.cks.library.nhs.uk/patient_information_leaflet/cancer_of_the_prostate>.

"Cancer of the skin". NHS Direct (CKS). 5 Dec. 2007 <http://www.cks.library.nhs.uk/patient_information_leaflet/cancer_of_the_skin>.

Appendix

Summary (in German)

Englisch als Fachsprache erhält zunehmend mehr an Aufmerksamkeit von Seiten der Sprachwissenschaftler/Innen, Englischlehrenden, Arbeitnehmern und Arbeitgebern. Diese Entwicklung spiegelt sich in der Etablierung fachsprachlicher Zertifikate wider. Während Kurse und international anerkannte Zertifikate zu Geschäftsenglisch schon alltäglich sind, sind Sprachkurse zu Themen wie Medizin und Pflege noch eine Seltenheit, darüber hinaus wurde bis jetzt noch kein Zertifikat geschaffen.
Gerade im Bereich der Medizin herrscht Nachholbedarf wie auch eine geringe (aber ständig wachsende) Anzahl an englischen Arbeitsbüchern zeigt. In dieser Diplomarbeit wird argumentiert, dass medizinische Genre wie das der Patientenaufklärungsbögen in englischer Sprache in den Englischunterricht für (zukünftiges) medizinisches Personal miteinbezogen werden sollen und müssen da diese eine Reihe von interessanten Fachvokabular, Mehrwortlexemen und Sprachfunktionen wie definieren und beschreiben beinhalten.

Die vorliegende Arbeit basiert auf den so genannten *ESP Approach*, eine Schule der Genreanalyse. Als Methode wurde die auf einem Korpus basierende Genreanalyse gewählt. Der Korpus dieser Studie umfasst 20 Patientenaufklärungsbögen, die nach speziellen Kriterien ausgesucht wurden. Zu den Kriterien der Korpuszusammenstellung zählten unter anderem die Aktualität der Texte und die Themenwahl. So wurden nur Patientenaufklärungsbögen zu den Themen Brustkrebs, Prostatakrebs, Darmkrebs und Hautkrebs berücksichtigt. Ziel war es diese Bögen hinsichtlich Vokabular und Struktur zu analysieren. Für die Vokabelanalyse wurden die Computerprogramme RANGE und FREQUENCY verwendet, während eine manuelle Analyse für die der so genannten *move structure* Analyse notwendig war.

Die Analyse des Vokabulars zeigt, dass Vokabular der Allgemeinsprache wie auch der Akademischen und Technischen Sprache in Patientenaufklärungsbögen zu finden ist. Obwohl Wörter der Allgemeinsprache das Vokabular rein rechnerisch dominieren, stellen Wörter der akademischen und fachspezifischen Sprache einen großen Teil des Vokabulars. Überraschenderweise implizieren die Resultate der

Analyse, dass Patientenaufklärungsbögen akademischer und technischer sind als akademische Texte. Die Vielzahl und auch Bandbreite an akademischem und medizinischem Vokabular in derartigen Bögen spricht für etwaige Verständnisprobleme von Seiten der Leser. Im Rahmen der Vokabelanalyse wurden auch Mehrwortlexemen und Floskeln Aufmerksamkeit geschenkt. Die Resultate der Analyse zeigen, dass Mehrwortlexeme vor allem für die Benennung von medizinischen Konzepten herangezogen werden. Phrasen spielen in Patientenaufklärungsbögen auch eine wichtige Rolle, das sie eng mit Sprachfunktionen und daher mit kommunikativen Elementen des Genres verbunden sind.

Ein Kernbereich der vorliegenden Arbeit beschäftigte sich mit der Analyse der kommunikativen Bausteine des Genres. Eine Analyse dieser Bausteine ergab, dass Patientenaufklärungsbögen ähnliche Strukturen aufweisen. So werden Themenbereiche wie die Krankheit selber, Symptome dieser, Arten der Diagnose und Behandlung abgedeckt. Weiters ist es üblich Krankheitsprävention anzusprechen. Im Großteil der Bögen wird darüber hinaus auf weiterführende Literatur verwiesen. Interessanterweise, konnte festgestellt werden, dass Bildmaterial nur in wenigen Bögen inkludiert wurde. Dieses Ergebnis war dahingehend überraschend als dass eine größere Anzahl von visuellem Material erwartet wurde.
Im Rahmen der Analysen wurden auch Sprachfunktionen berücksichtigt. Die Resultate ergaben, dass Sprachfunktionen wie Definieren, Benennen, Beschreiben, Erklären und Beraten wichtige Rollen spielen.

Im letzten Teil der vorliegenden Arbeit wurde die Integration von Genre und Genreanalyse in den Sprachunterricht, das Curriculum- und Kursbuchdesign besprochen. Hierzu wurden verschiedene Methoden vorgestellt. Im Großen und Ganzen konnte festgestellt werden, dass Genre und Daten von Genreanalysen für die Planung und das Design von Unterrichtsstunden und Kursmaterial herangezogen werden sollten.

Curriculum Vitae

Name:	**Ulrike Gruböck**
Nationalität:	Österreich
Geburtsdatum:	23.12.1981

AUSBILDUNG

Datum:	seit Wintersemester 2002
Bildungseinrichtung:	Universität Wien
Studienrichtung:	**Lehramtsstudium** Englisch, Psychologie und Philosophie
Zusatzqualifikation:	CerTESP – Certificate in Teaching English for Specific Purposes (ESP)
Schwerpunkte:	ESP, Applied Linguistics, Ethik, Österr. Gebärdensprache

Datum:	2001-2002
Bildungseinrichtung:	Twin Tower College Finchley, London
Schwerpunkt:	Englisch

Datum:	1996-2001
Bildungseinrichtung:	HBLA für wirtschaftliche Berufe, Krems
Abschluss:	Matura mit gutem Erfolg
Schwerpunkte:	Gesundheit und Soziales
Zusatzqualifikation:	Abschlussprüfung in Servierkunde und Kochen

BERUFSERFAHRUNG

Datum:	2002-2007 (Sommermonate)
Funktion:	Ferialarbeiterin
Zuständigkeit:	Administration, technische Übersetzungen, Geschäftskorrespondenz, Reisebuchungen
Name des Arbeitgebers:	Moeller Gebäudeautomation, Wien

Datum:	Juli 2005 und Juli 2006
Funktion:	Betreuerin/ Lehrerin
Zuständigkeit:	Englischunterricht, Freizeitprogramm
Name des Arbeitgebers:	Kinderfreunde NÖ

Datum:	2001-2002
Funktion:	Au-pair
Zuständigkeit:	Kinderbetreuung
Name des Arbeitgebers:	privater Haushalt in London

Datum:	Juni 1999-August 1999
Funktion:	Ferialpraktikantin in der Großküche eines Pflegeheims
Zuständigkeit:	Essenszubereitung
Name des Arbeitgebers:	Dr.-Thorwesten-Heim, Krems

SPRACHKENNTNISSE: Deutsch (Muttersprache), Englisch (C2), Französisch (A2), Österreichische Gebärdensprache (fortgeschritten)