UNIVERSIDAD DE SALAMANCA

Facultad de Filología
Departamento de Filología Inglesa

TESIS DOCTORAL

Attribution and thematization patterns in science popularization articles of *The Guardian* newspaper

Blanca García Riaza

2012

UNIVERSIDAD DE SALAMANCA

Facultad de Filología
Departamento de Filología Inglesa

TESIS DOCTORAL

Trabajo para optar al grado de doctor presentado por
Blanca García Riaza bajo la dirección de la Dra. Izaskun
Elorza

2012

Vº Bº Directora   Autora

Dra. Izaskun Elorza   Blanca García Riaza

To my parents, María and Enrique

I found myself in a dirty, sooty city.

It was night, and winter, and dark, and raining.

I was in Liverpool

C. Jung

Foremost, I want to specially thank Dra. Izaskun Elorza, supervisor of this doctoral dissertation, for her invaluable contribution to the satisfactory development of this work; for her advices, patience and lessons, which have been crucial in this hard process. Thanks for guiding me in every step I have taken in my academic life and, above all, for always relying on me.

I am indebted to my many of my colleagues who have always encouraged me to struggle and get the better of myself and have always offered their help and knowledge along the elaboration of this doctoral dissertation. My special gratitude to Reyes, Juanpe, Ovidi, Nely and Anna, great colleagues who have always been there.

I owe my deepest gratitude to the teachers from the University of Liverpool: Michaela Mahlberg, Geoff Thompson, Mike and Nell Scott, who made me feel at home while I stayed there in 2009 and provided me with a stimulating environment to develop this work. Thanks also to the reviewers, Mike Scott and Tom Bartlett, and the proofreader Mehr. I., who have enormously contributed to make this work possible.

A very special gratitude has to be expressed to my family, without whom I would not have been able to accomplish this work, and who have taught me all that cannot be found in books. Thanks to my parents, who have always supported and encouraged me to do my best in all matters of life, to my brother and sister, who have always relied on me and encouraged me to achieve whatever I set myself, thanks for their love and support, for their patience and comprehension. I also want to specially thank Sofía, Arturo, Marta and Carlos, who, with the gift of their company, made my days more enjoyable and worth living.

Thanks to my friends Alicia, Miri, Laura C., Lucía, Luis, Isaac, Julio, Carlos and Reme. To Miriam, Víctor, Mariví, David and Vero. Thanks also to Héctor, Linda, Diego, Rodrigo, Carlos G. G., Carlos A., for helping me get through the difficult times, and for all the emotional support, comraderie, entertainment, and caring they provided.

Thanks

Abstract

This doctoral dissertation places its focus on the discourse of news media in English, which has traditionally been an attractive and fruitful source of information about language for both discourse analysts and corpus linguists. While many of the analyses have focused on scientific discourse, the alternative approach to science popularization proposed focuses on its description as a newspaper genre, placing our emphasis on the description of textual features typical of science popularization discourse. The popularization of science is normally carried out by a number of journalists who write in scientific sections and act as mediators between scientists and lay people, and who elaborate discourse through a series of linguistic and discursive choices by which they obtain credibility for the facts under comment. One of those resources is that of explicit attribution which presents discourse as deriving from someone different from the journalist, and introduces reported language events in the text.

In this context, this study was created with the aim of providing the phenomenon of attribution in newspaper discourse with deeper insights that help improve the knowledge about language event's reports in science popularization articles. For this purpose, a corpus of electronic versions of science popularization articles from the British newspaper The Guardian has been compiled and analysed to get data on the presence of attribution in science popularization articles and on the behaviour of language events reported, both in first paragraphs and in the rest of

the text, taking into account thematisation patterns in the texts. The model of language that has been taken as reference in this study is Systemic Functional Linguistics, which focuses on how language works and conceives language as a social semiotic, also taking Discourse Analysis as the perspective adopted for the analysis of language.

We have obtained relevant results on the phenomena of explicit attribution, simple and complex processes of language events' reporting and quoting, as well as on the use of the epistemic stance according to and reporting verbs as reporting signals in science popularization articles. We conclude that a close look at the data reveals that the traditional approach doesn't account for the phenomenon of attribution in science popularization articles appropriately.

Resumen

En esta tesis doctoral, el foco de atención se centra en el discurso de los medios informativos en inglés, una atractiva y productiva fuente de información sobre el lenguaje tanto para analistas del discurso como para lingüistas de corpus. Mientras la mayoría de los análisis sobre el discurso sobre ciencia en los medios de comunicación se han concentrado en el discurso científico, el enfoque alternativo que proponemos estudia los artículos de divulgación científica, tomándolos como un tipo de género periodístico y concentrándose en la descripción de elementos textuales típicos del discurso de divulgación científica.la divulgación de la ciencia normalmente se lleva a cabo por una serie de periodistas que firman artículos en secciones científicas, y que actúan como mediadores entre los científicos y el público general, no especializado. Estos periodistas son los encargados de elaborar el discurso de divulgación de la ciencia utilizando una serie de opciones lingüísticas y discursivas con las que consiguen obtener credibilidad a cerca de los datos proporcionados. Entre estos recursos disponibles, nos centramos en el fenómeno de la atribución explícita, que presenta fragmentos del discurso, atribuyéndolos a voces externas diferentes a la del periodista, e introduce el fenómeno del discurso referido en la narración de los divulgadores.

Este es el contexto en el que hemos desarrollado este estudio, que tiene como objetivo ahondar en el conocimiento de la atribución explícita en el discurso de los medios de comunicación, haciendo un análisis detallado sobre el discurso referido en los artículos de divulgación científica. Para desarrollar este análisis, hemos compilado y analizado un corpus de versiones electrónicas de artículos de divulgación científica del

periódico británico The Guardian *para obtener datos sobre la presencia de atribución explícita en los artículos, así como pautas sobre el comportamiento del discurso referido en los mismos, tanto en primeros párrafos como en el resto del texto, y teniendo en cuenta las pautas de tematización encontradas en los textos. El modelo de lenguaje que se ha tomado como referencia en este estudio es el de la lingüística Sistémico-Funcional, que se centra en el funcionamiento del lenguaje, al que concibe como semiótica social. La perspectiva adoptada para el análisis del lenguaje es la del Análisis del Discurso.*

Del análisis llevado a cabo a estos textos, hemos obtenido resultados relevantes sobre el fenómeno de la atribución explícita, el discurso referido, tanto directo e indirecto, como simple y complejo, así como sobre el uso del adverbio de modalidad epistémica according to *y los* verba dicendi *como marcas de discurso referido en artículos de divulgación científica. Los datos obtenidos han revelado que, tras un análisis exhaustivo de las pautas encontradas en los textos, las descripciones tradicionales sobre el discurso referido no dan cuenta del fenómeno de la atribución explícita de una forma adecuada.*

Indexes

_____

_____

Chapter 3. The encoding of attributed voices in the text

Chapter 4. Information expliciting: Theme and Rheme in discourse

_____

PART B. EMPIRICAL STUDY

Chapter 5. Study description

Chapter 6. Analysis and results

Chapter 7. Conclusions

Annex

Table Index

_____

_____

Figure Index

_____

_____

Chapter 1

Introduction

Chapter 1

Introduction

1.1.General Introduction

1.2. Structure of the study

_____

> *"The mere formulation of a problem is far more essential than its solution,*
> *which may be merely a matter of mathematical or experimental skills.*
> *To raise new questions, new possibilities, to regard old problems from a new angle*
> *requires creative imagination and marks real advances in science."*

Albert Einstein

1.1. General introduction

There has been a growing concern in linguistics about how words behave and what their roles are in specific contexts. Immersed in this area of enquiry, studies about language using corpora have played a major role. In opposition to the study of language as an abstract or static entity, corpora have gained a relevant position as linguistic repositories which supply researchers with real instances of language in context, valuable data which allow us to get a deeper insight in how language works.

_____

This process has been fostered by the development of information and communication technologies, which have favoured the creation of specific software to assist linguists in the analysis of language. Moreover, the generation of computing tools has meant a great advance in corpus linguistics studies, not only because it has made the analytical process more powerful, but also because it has triggered new discoveries which have shaken the foundations of linguistics, such as the importance of collocational patterns in language use.

Corpus linguistics offers not only methodological guidelines to analyse language, but also a theoretical framework which is getting more and more consolidated thanks to the contribution of linguists and resources which have made knowledge into this field advance. For this reason, corpus linguistics is has been chosen as the research methodology followed in this study.

The model of language that is taken as reference in this study is the one initiated by Firth and Halliday, which focuses on how language works and conceives language as a social semiotic.

In broader terms, Discourse Analysis is the perspective that we have adopted for the analysis of language. It examines "how humans use language to communicate and, in particular, how addressees construct linguistic messages for addressees and how addressees work on linguistic messages in order to interpret them" (Brown and Yule, 1983: IX).

In this context, this study was created with the aim of providing the phenomenon of attribution in newspaper discourse with deeper insights that help improve the knowledge about language event's reports in science popularization articles. Patterns observed in this study can be useful both for linguists and

journalists who deal with texts popularizing scientific findings for the non-specialised reader.

There are a great number of studies which have focused on the discourse of mass media, a relevant source of texts, which we benefit from in this study of the British newspaper *The Guardian*. Within this newspaper, we focus our attention on science popularization articles, and, more specifically, on the attribution processes that appear in the narration of language events reported in the texts. Traditionally, the description of language reports has been based mainly on the structural dichotomy between direct and indirect speech reports (Calaresu 1998: 870). However, as it will be shown, a close look at the data reveals that this traditional approach doesn't account for the phenomenon of attribution appropriately.

1.2. Structure of the study

The study that we present here is divided into two main sections. In the first section (Chapters 2, 3 and 4), a theoretical revision is present, and a broad vision is offered on the most relevant aspects of newspaper discourse and thematisation, whereas in the second section, (Chapters 5 and 6) we deal with the empirical part of the research, followed by a final chapter with the discussion and conclusions of the results (Chapter 7).

In Chapter 2, we examine the different characteristics of newspaper and science popularization discourse by comparing them to media and scientific discourse, with which they share some common features. Taking a step further,

we outline the specific features that characterize science popularization discourse and its popularity for text analysis in the last decade, also comparing them with different types of discourse that share some features but do also present distinctive features.

In this chapter, Discourse Analysis is set as the analytical framework that considers the kind of language used to communicate something and is recognized by receivers as coherent. However, the role of Information and Communication Technologies in the broadcasting process is not ignored.

Chapter 3 expands into one of the most frequent phenomena in science popularizations, that of evaluation, and more specifically, in attribution and reporting processes. Here, the notion of evaluation is explored as the umbrella term for the writer or speaker's attitude towards the linguistic events reported in a text. It considers the different models that have been created for the description of the phenomena of evaluation and their implications for the analysis of written newspaper articles. Moreover, we focus on two specific ways of expressing evaluation: reporting and stance adverbials; two different options available for the science popularization writer to manifest his/her position towards the narrated events.

Elements dealt with in Chapter 3 lay the foundation, for Chapter 4, in which the relevance of the information structure in written texts is considered. They also create the basis for the analysis of attribution patterns present in language event's reports from a corpus of science popularization articles that will be carried out later in this study.

Chapter 4 is guided by the idea that the election of what comes first in a sentence can appear as something casual that speakers and writers choose without purpose. However, elements in initial position have an important role in the clause, as well as in defining the way information reaches the audience. This is the departure point for a chapter that expands into the notions of theme and rheme as elements which organize the way the message is expressed and how the information reaches its receiver. In this way, Chapter 4 reviews the most relevant thematic structures at the same time that it deals with aspects such as delimitation of thematic structures or multiple themes. Furthermore, we describe the concepts of thematic progression and informativity, to provide readers with different perspectives on the study of theme. As a final part of this chapter, we narrow our focus and describe how themes behave in speech reporting clauses and its logico-semantic character as a projection structure.

To sum up, in the theoretical part of the study (Chapters 2 to 4) we review the most relevant literature on the three notions that will be taken as basis for the analysis that is going to be developed in subsequent chapters of this thesis. First, we examine the peculiarities of newspaper and science popularization discourse. We focus on the special characteristics of science popularization articles, a text type with scientific content but printed in a publication aimed for a general, non-specialized audience where the journalist acts as a mediator who brings the knowledge of the scientists to the general readers of the newspaper. Secondly, and as a second basis for this study, we explore in further detail one of the most common phenomena in science popularization discourse: the attribution of messages conveyed to voices different than that of the journalist. Reports and quotes are studied through the different classifications made on the issue, from a

systemic functional perspective as the guiding perspective to deal with this form of epistemic modality. Thirdly, we bring to light the importance of theme structure, reviewing in depth the roles of theme and rheme and how they contribute to the thematic progression of texts and the expliciting of information. Bearing all these three elements in mind, in the second part of the study, the empirical part, we present an analysis on the thematic choices made by journalists of science popularization articles from the British newspaper *The Guardian* in the quotes and reports introduced in their texts to attribute scientific messages to external voices.

The empirical part of this study starts with an overview in Chapter 5 on recent studies which have as well set their focus of attention on the language of science popularization articles. Most of the studies described have also benefited from corpus methodology to reflect the textual features that are conventional for a community of speakers, and are used as the basis for the study that has been carried out.

In Chapter 5 we formulate the hypotheses of the study, which can be divided into different parts, according to the different areas of enquiry which will be dealt with later. To achieve the objectives set forth in this study and test the aforementioned hypotheses, a corpus of science popularization articles was collected. With regards to the methodology, we have followed a corpus-based approach. To assist us with the computerized analysis carried out, we have used the processing software *WordSmith Tools*, that enables linguists to make statements about the typical features that can be found in the pieces of language studied.

Chapter 6 is devoted to the description of the data obtained in the analysis of the corpus, the presence and relevance of attribution in the corpus, and the significance and frequency of this phenomenon in the corpus. Once a general overview has been provided on the relevance and presence of the phenomenon of attribution, we first focus on simple reports and their relevance and presence in the corpus, paying special attention to their behaviour in first paragraphs. Secondly, we concentrate on complex reports, which link several occurrences of attribution into a single string. Double, triple, quadruple and quintuple cases are described both in terms of presence in the corpus and of patterns that most frequently appear in science popularization articles from the British newspaper.

Subsequently, the focus shifts to adverbial reporting. Data on the presence of the epistemic stance adverbial *according to* on the corpus are presented, so that the relevance of this adverbial is established, especially in initial sections of the articles compiled for the corpus. A further analysis is carried out to find out how this adverbial relates to transitivity system (Halliday 1985: 101), as well as to obtain information on the elements that most frequently appear on the right of the adverbial. Relevant outcomes have been achieved in this part of the analysis, which is followed by an inspection of the different positions that *according to* usually occupies in the corpus. Data on thematic, rhematic and mid-sentence occurrences of the adverbial are presented together with the relevance of the different locations in clause structure.

Chapter 6 deals with one of the most relevant reporting signals, namely reporting verbs, which are analysed for the purpose of gaining a deeper insight into the use and patterns of this reporting signal in science popularization articles

of the British press. Tense in reporting verbs encountered in the corpus has also been analysed in search of repeated patterns in order to identify the different verbs that most frequently appear in the corpus and classify them according to their function.

Finally, in Chapter 7 we gather the most important outcomes acquired throughout the study and also consider the degree of consecution of the objectives and the relevance of this research for the field in which it is framed. Conclusions on the results are also revealed in the final part of this study, together with some of the most relevant pointers for future research that have come to the front in this study.

Part A: Theoretical Framework

Chapter 2

Newspaper discourse and science popularizations

_____

Chapter 2

Science popularizations within newspaper discourse

2.1. Introduction

2.2. Mass media

    2.2.1. Why study mass media?

2.3. Newspaper discourse, a mirror for reality

    2.3.1. News text types and newspaper genres

    2.3.2. News: values, structure, and creation

    2.3.3. News texts

2.4. Science popularizations as news

    2.4.1. The anatomy of science popularization articles: Characteristics

    2.4.2. Science popularization sections in newspapers

    2.4.3. Science popularizations as objects of linguistic study

    2.4.4. A hybrid scientific communication mode

       2.4.4.1. Scientific discourse versus science popularization discourse

2.5. Corpus linguistics and the study of texts

Science is a sun:

everybody must move closer to it

for warmth and enlightenment.

Louis Figuier

2.1. Introduction

The fact that advances in communication and information technologies developed in the 20th and 21st centuries have changed the world surrounding us is not news anymore. Access to information through the internet has become the most popular means for people to stay updated on the latest international events, and newspapers have profited from a growing readership that demands current and reliable information every minute. In this chapter, we will deal with the role of mass media as the most powerful information dealer, and, more specifically, with how mass media constructs its own news discourse. We will also outline the

distinguishing features of science popularizations, a hybrid genre that is the textual object of the research carried out for this doctoral dissertation. We will devote this chapter to identifying the characteristics of mass media discourse in general and newspaper discourse on the one hand, and science popularization and scientific discourse on the other. We will establish comparisons among these characteristics highlighting the similarities and differences of the four types of discourse.

2.2. Mass media

In today's society, media are not only producers of information, but also important institutions that present culture, politics and social life, together with the way in which they are linguistically built. Media discourse does not only reflect the textual features that are conventional practice in a given community (Kim and Thompson 2010: 54), but also how values are encoded in society (Garret and Bell 1998: 64-65). We will then refer to mass media as encompassing all broadcasting organizations, which either through printed, internet-based, visual or oral channel, have become the dominating presenters of language in society (Bell 1991: 1).

Bell makes an interesting distinction between 'discourse' and 'text'. We have referred to language in media as 'media discourse' and will refer to news within media as 'texts', but what is the difference between text and discourse?

In this respect, Garret and Bell (1998: 2-3) point out that there has been a traditional lack of agreement in the definition of both terms, albeit they propose to

_____

consider text as the "outward manifestation of a communication event" (Garret and Bell 1998: 3) and discourse as not being only concerned with language but also with the context of communication and factors such as who, with whom, why, what kind of society or through what medium the communication is being carried out. We propose to take into account Cook's (1989: 6-7) definition of discourse as the kind of language used to communicate something recognized by receivers as coherent.

Modern media, and specifically media based on Information and Communication Technologies (hereafter ICTs) cannot be seen anymore in terms of the traditional dichotomy of text and discourse. The notion of 'media texts' has evolved alongside the technology for producing and broadcasting such texts, therefore media texts are no longer considered as "words printed in ink on pieces of paper" but have broadened in scope to include speech, music, image…(Bell and Garret 1998: 3).

A further terminological distinction to be considered at this point of the study is the difference between two dimensions in the media: 'outlets' and 'outputs'. Bell refers to outlets as the "publications, television channels or radio stations which carry the content", while outputs are defined as the products generated by the different outlets and the period of time that is to be covered (Bell 1991: 12). World media runs on a cycle of 24 hours; newspapers have daily editions and broadcast media, radio and television, transmit bulletins which make recapitulations of the past 24 hours' news (Bell 1995b: 307), but something

_____

different happens with internet-based broadcasting, where webpages are constantly being updated with information that changes several times a day.

Regarding the nature of media discourse, another remarkable fact is that outputs, as communicative events, are highly conditioned by the audience to which they are addressed. Media communication is obviously different from face-to-face communication, but the most important feature that differentiates these two types of communication is the fact that mass communication entails place and time asynchrony, and, what is even more relevant, a disconnection between communicator and audience.

The concept of audience has inevitably varied in the last years due to advances in ICTs. Recent advancements have created media outlets dependent on technology and new internet-based communication channels that have taken over radio and television, not to mention printed media. As Bell (1991: 85) points out, the audience dimension is one of the main defining characteristics of mass media, and he articulates this relationship between audience and media outputs in the following six different features:

- The audience is larger than in other communicative events

- Mass media content is publicly more accessible in relation to other contents

- The audience is heterogeneous

- There exists simultaneous contact with diverse individuals

- The flow of communication is unidirectional and impersonal

- The mass audience can be considered a creation of modern society

2.2.1. Why study mass media?

Mass media studies have traditionally been described in books and introductions as 'the media' by making reference both to the institutions and the messages they create. The media, including press, cinema, radio, television, as well as new communications via ICTs, have been considered as conveyors of social, political and moral messages produced in response to temporal events in a specific society that help readers form

> definitions of reality or perceptions of the world […] socially constructed,
> internalized, and then to a large degree seen as natural or common sense. Thus, the
> media often help naturalize particular modes of perceiving the world and offer
> various forms of pleasure to go along with these perceptions. (Davis 1993: 165)

As Bell states (1991: 2), news are carriers of the images and stories of our days, as well as a linguistic reflection of the society's values, which are conveyed in a way that influences the audience's reception of the content. He also points out that mass media communication has a series of characteristics that make it different from other types of oral or written communication, namely that it has multiple originators, a massive and fragmented simultaneous audience and a growing accessibility to different kinds of public all over the world.

The ways in which news stories are told show interesting uses of language by themselves because language is both an instrument and an expression of messages through media. General public, researchers and communicators themselves are interested in the content of media communication as well as in the way language carrying that content is constructed (Bell 1991: 3). This situation

_____

has enabled Bell to list some of the main reasons that may have made language research be fraught with media language examinations that we present as a summary in Figure 1 below.

MAIN REASONS TO STUDY MEDIA LANGUAGE

- Accessibility of media as a source of data for the language feature/s observed

- Interest in some aspect of media language in its own right, such as headlines

- Interest in the way media language use some specific feature also found in casual, ordinary speech

- Taking advantage of how media communication manipulates language in a revealing way (news' editing)

- Interest in media's role in affecting language in society

- Interest in what language reveals about media's structure and values

- Interest in what media language reveals as a mirror of society and culture

- Interest in how media affects attitudes and opinions in society through the way it presents people and issues

Figure 1. Main reasons for the study of media language (Bell 1991: 4)

The motivation for developing the numerous media language studies carried out in recent years differ depending on the availability of texts and the informativity of media language on different societies' values (Bell 1991: 3, Bell 1995a: 23, Garret and Bell 1998: 3-4).

On the one hand, much of the recent research carried out in the field of media discourse falls into the umbrella of 'Critical Discourse Analysis' (hereafter CDA), a framework that proposes an interdisciplinary approach to the study of language and whose main representatives are Teun van Dijk (1985, 1988, 1993,

_____

1998, 2009) and Norman Fairclough (1989, 1992, 1995). CDA focuses on the manipulative power of language (Van Dijk 1985, 2009) and takes into account, as well as the information conveyed in texts, text-external factors such as the interpretation of the message or the power that mass media can have on people (2009:13). In the same framework, we find the work carried out by Fowler (1991), who applies the Critical Linguistics framework to the analysis of news, and, according to Bell (1995a: 28), takes a political point of view and specifically concentrates on lexis rather than on linguistic features or discourse structure.

On the other hand, Fairclough (1989, 1992, 1995) analyzes media texts from the perspective of Halliday's Systemic Fuctional Grammar[1], and presents three different dimensions to discourse analysis (Fowler 1991, Bell 1995a: 29). Firstly, 'text analysis', which is concerned with the language contained in texts and its characteristics, drawing as well its attention to the interpersonal elements in the text. Secondly, 'analysis of discourse', which takes into account factors such as text reception and production, as well as to the features of discourse practices in the media. And thirdly, 'analysis of social practice', which concentrates on the role of discourse in society, especially in its relations to ideology and power, an idea which is based on CDA approach.

Among alternative analyses and descriptions of media discourse and the language of mass media, it is pertinent to mention Bell's (1991, 1995a, 1995b) illustrative and detailed description of news' structure, production and audience. The fact that Bell was first a journalist and then a linguist provides him with a

---

[1]We will deal with Systemic Functional Grammar in greater detail in Chapter 3.

_____

unique perspective allowing him to present the parts of the news production process which had previously remained more opaque. Bell emphasizes the idea that news is not a product of just one hand, but the result of a group effort drawing from many external sources of information to create a 'story', a basic notion in news writing which differs from that of 'articles' in that the former possess a specific structure, values and viewpoint which the latter cannot be said to have (Bell 1995a: 26).

2.3. Newspaper discourse, a mirror for reality

2.3.1. News text types and newspaper genres

Among the many mass media that we have available nowadays, newspaper discourse has received special attention from discourse analysts. In Ljung's opinion, this is due to the fact that news discourse has a more genre-like status than other types of press discourse:

> The very fact that the news category has received such massive attention, while this has been denied to the other newspaper categories, seems to indicate that news texts have a much stronger categorical status than the others and that news is, somehow, more genre-like than other kinds of newspaper writing. (Ljung 2000:132)

An additional question that comes to the forefront when dealing with the special attention given to newspaper discourse concerns the fact that news discourse has sometimes been referred to as a kind of 'genre' (Bell 1991: 12-13), while others have been alluded to just as a specific discourse type. Any research which involves a particular kind of texts in relation to others, demands a

typological classification of texts with respect to the possible types of discourse dealt with (Biber 1988: 206). The concept of genre is not a straightforward one, although it seems obvious that groupings of texts which display some kind of similarity, either linguistic or functional, can be considered to belong to a common category, that of genre (Ljung 2000: 132). Genre has been used to refer to the categorizations which depend on external criteria which relate to the speaker's or writer's purpose, and which can be portrayed out of the comparison with 'text types', "classes of texts that are grouped on the basis of similarities in linguistic form, irrespective of their genre classifications" (Biber 1988: 206). It was Swales (1990) who conducted a more in-depth study the concept of genre , concluding that exemplars of a genre display various patterns of resemblance in terms of structure, style, content and target audience. When all high-probability expectations are met, the exemplar will be viewed as a prototype of a particular genre (Ljung 2000: 140). As Swales asserts:

> Genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. (Swales 1990: 58)

In a more media oriented sense, Bell defines genre in two fashions. First as 'the particular kind of media content in which you are interested – news, classified advertising, game shows, weather forecasts, and so forth' (1991: 12) and second as 'largely a matter of defining what kind of copy does and does not qualify as what you are looking for' (1991: 17).

_____

Among the many classes of press news that can be found in different newspapers, Bell (1991: 14) distinguishes different classes of news: 'hard news', 'feature articles', 'special-topic news' and news elements: 'headlines', 'bylines', 'photos', etc. Hard news are those time-bound to immediacy, in contrast to 'soft news' which do not have such a temporal dependency, and which report accidents, crimes, etc. On the other hand, Bell (1991: 14-15) considers special-topic news to be those that appear in specific sections of newspapers and are written by several journalists who are explicitly working on a subject such as business or sports, under the direct control of a specific editor.

2.3.2. News: values, structure, and creation

Independent of the section in which news are published, there are common features that enhance the newsworthiness of the information conveyed and, therefore, justify the inclusion of a certain piece of news in a newspaper edition. Bell (1991: 156-158) has proposed one classification of three types of news: content of news, news process and quality of the news text. The first type put forward establishes twelve characteristics that concentrate on news discourse itself and leave apart factors relating to properties and actors who create the content in the news. These categories help us clarify what is considered a piece of news in newspaper discourse, and can be seen in Table 1 below.

News aspect	Description
NEGATIVITY	Negative events make the basic stuff of 'spot' news
RECENCY	The best news is something which has just happened
PROXIMITY	Geographical closeness can enhance news value
CONSONANCE	Its compatibility with preconceptions about the social group or nation from which the news actors come
UNAMBIGUITY	The more clearcut a story is, the more it is favoured
UNEXPECTEDNESS	The unpredictable or rare is more newsworthy than the routine
SUPERLATIVENESS	The biggest the event is, the more it gets covered
RELEVANCE	The effect on the audience's own lives or closeness to their experience
PERSONALIZATION	Indicates that something which can be pictured in personal terms is more newsworthy than a concept, a process, the generalized or the mass.
ELITENESS	Reference to elite persons as actors can make news out of something that would be ignored by ordinary people.
ATTRIBUTION	The eliteness of a story's sources. High valued news sources need to be elite in some dimension, particularly socially validated authority.
FACTICITY	The degree to which a story contains the kind of facts and figures on which hard news thrives: locations, names...

Table 1. Values in news actors and events. (Bell 1991: 156-158)

_____

Taking into account Bell's classification, which is comparable to the work of authors such as Hernando y Hernando (2006: 94), totally or partially used for analyses on newspaper discourse (Elorza 2010), we can state that these categories are not mutually exclusive, but act in a cumulative way identifying stories as more newsworthy if they satisfy more than one factor. Moreover, Bell (1991: 160) asserts that "there are also factors in the quality or style of the news text which affect news value: clarity, brevity and colour". The relevance of a news piece is not based only on the content and message, but also on how it is conveyed to the public, factors that condition both the quality of the message and its reception. What Bell is implying here is that both meaning and form are key factors in the successful development of news process: production and meaning (re)construction.

In relation to news production, it is necessary to highlight, as Bell (1991: 36) asserts, that the production process is a cyclical chain developed in a newsroom where staff members constantly interact making it difficult to identify who has produced which language forms. The journalist in charge of writing the article keeps his/her role as author, but there are many people who handle the text produced before it is broadcast in the media: chief reporter, subeditor, editor, principal and animator. This fact leads us to the idea that news cannot be based just on the facts reported, but on the editorial work of a large team of people who influence, in greater or lesser degree, the final product presented to the audience. Despite being those who create the content of the articles and sign with his/her name, journalists are not as original as it may be thought. As Bell (1991: 41) asserts, news is, in fact, what somebody tells a reporter, and conversely much of

_____

what a journalist writes in an article are paraphrased or quoted excerpts from someone else's speech. The different ways in which the journalist inserts those excerpts into the existing news discourse is one of the basic features of news media and will be discussed in greater detail in Chapter 3.

2.3.3. News texts

News' structure is not a simple or personal decision of the journalist, but a reflection of the complex model of information structuring followed by authors as convention. This model divides the text structure into three different types of content: an 'abstract', the 'attribution' and the 'story itself' (Bell 1991: 169). The abstract consists of a headline and a lead which includes the main event of the story, normally also giving information about the actors and setting concerned in the event. Attribution, in this case, refers to where the story came from, agency, journalist's name, place and time, although it may not always be explicitly stated. The story thus begins with the first paragraph of the text, and is subsequently organized in 'episodes' that are then divided into 'events'. The number of episodes and events within a story is not restricted, but dependent on the length of the news piece and thus on the quantity of information that the author has available on a specific subject. As Bell (1991: 172) states, the order of events is crucial in personal narratives as well as in newspaper discourse, but chronology means nothing in news discourse. It is one of the most striking features of newspaper discourse that events do not follow a chronological order, and what is more, events introduced in the first stages of news discourse are sometimes

_____

returned to in later stages of the discourse. In Figure 2 below, we can find the

complete model structure of news texts designed by Bell.

(see Appendix 1).

Figure 2. Model structure of news text (Bell 1991: 171).

As we previously noted, Bell's concept of attribution refers to the

circumstances surrounding the composition of the article, such as the time and

place, when and where it was elaborated, and the source of the information

conveyed. In this study, the concept of attribution will be used with a different

meaning. 'Attribution' has been considered as a process by which external voices

included in the text are explicitly marked as coming from a source different from

the author, and thus, the information provided in the text is grounded on a voice

not corresponding to the journalist. In this same respect, and as Thompson states,

_____

> The main expectation that a reader brings to a newspaper report on current affairs is that it will give a reasonably objective 'true' picture of the event and of the evidence, reactions, and opinions of the people involved. (Thompson 1994: 151)

Moreover, Thompson (1994: 151) divides people whose information and opinions are presented into newspaper discourse into two different groups. On the one hand, he distinguishes 'important' people, such as politicians and experts, and on the other, 'ordinary' people who have been involved in the news piece reported. These two groups normally appear in the articles as participants, witnesses or commentators of the event reported, although not only people can be considered participants, but also institutions, which are 'voiced' and presented as saying or thinking. In the reporting of events, a journalist can be present at the event making him/her a first-hand witness, but the most common situation is to rely on the account of someone else and report the information using reporting structures that emphasize the first-hand ("*The child fell*") or second-hand ("*The neighbours say the child fell*") character of the information. This broad conception of the notion of attribution will be developed in greater detail in Chapter 3 of this study.

2.4. Science popularizations as news

2.4.1. The anatomy of science popularization articles

Science popularizations, also named popular science or 'popular science writings', are defined as science writing for the general public that provide

_____

information on advances in scientific breakthroughs, as Kim and Thompson (2010: 54) state, but also about discoveries and inventions. The presentation of scientific knowledge has been conceived in the last few years as "scientists having regular interaction – direct or mediated – with the general public, […] apparently seeking to fill the traditional gap that exists between the scientific community and people in general", as Calsamiglia and López Ferrero (2003: 147) state. That means that popularization discourse is typically formulated in a way that constructs lay versions of specialized knowledge and integrates them into the non-expert reader's previous knowledge. As a consequence, several strategies of explanation (definition, example, metaphor…) are used as semantic vehicles enabling readers to associate new information with previous knowledge (Calsamiglia and Van Dijk 2004: 370) and the discourse is structured according to two components: journalistic and didactic (Hernando and Hernando 1994: 63). An additional characteristic of science popularization discourse is the recontextualization of knowledge, which is first shaped in a specialized context and then reconstructed in a different communicative context, addressed to the lay audience. It is the communicative situation that conditions the construction of scientific discourse (Calsamiglia 1997: 11), which is also highly dependent upon the channel through which it will be broadcasted: newspaper, scientific magazine and so on. According to Cassany, López and Martí (2000), there are three relevant processes carried out by the journalist in charge of popularizing a scientific event: denomination, re-elaboration and textualization. However, Calsamiglia and Van Dijk (2004) only recognize two of them (re-elaboration and textualization) while Williams (2009) considers that the linguistic transformation of the language of

science into science popularization discourse is achieved by a dual process of recontextualization and transformation:

> Popularization of scientific texts can be broadly defined as a variety of genres, spoken and written, which involve on the one hand a recontextualization of scientific discourse, and on the other hand a transformation of specialized knowledge into 'everyday' or 'lay' knowledge in the sphere of the mass media or other institutions. (Williams 2009: 467)

Along the lines of Calsamiglia and Van Dijk, Hernando and Hernando (1994: 62) consider that popularization consists of three different movements: expansion, reduction and variation. Expansion refers to the incorporation of elements that are not present in the original source text. Reduction implies the suppression of information that is not relevant, necessary or convenient in the popularized version, or the summary or synthesis of specific contents. The third movement, variation, entails the changes made in how the information is presented in the popularized version with respect to the source text, mainly regarding lexis, but also modifications of other linguistic aspects.

As Ciapuscio notes, the reconstruction of scientific information into science popularization articles implies an intellectual challenge as it entails "*reelaborar contenido científico con fidelidad conceptual y modal para un lector no experto, a quien además de informar hay que interesar en los temas de la ciencia*" (1997: 28). The fact that the world of the specialist and the world of the lay people come together in science popularization articles has made them become, in Moirand's (1997) words, a 'meeting point' for scientists and the average citizen. Building a bridge between specialists and non-specialists entails a series of conditions that

represent science and scientists outside their traditional area of influence, research centers and universities, where science had been closed off from the needs and interests of the general public. The journalist needs to communicate with press readers to convey information regarding a science and/or technology event but, moreover, the scientist has to persuade readers that the information in the article is both important and useful to them (Hernando and Hernando 2006: 61). As Calsamiglia (2003: 140) states, streams of democratization and globalization of knowledge have progressively compelled these two worlds to unite filling the gap that traditionally existed between these two separate audiences. That gap does not only refer to the different characteristics of both "discursive communities" (Swales 1990: 24), experts versus common people without specialized background knowledge, but also to the fact that the two communities perceive scientific "objects" (Swales 1990: 24) in quite different ways. Scientists focus on the immanent values of scientific advances within their specialized contexts while lay people value discoveries based on their application to ordinary life, their utility and consequences for life.

Finally, we would like to present some of the main principles which rule the nature of science popularization defined by Calsamiglia and Van Dijk (2004: 371). As the object of consideration in this study, the following principles emphasize the text structures that play a role in presenting knowledge to non-specialized readers:

1.- Popularization is a social process consisting of a large class of 'discursive-semiotic' practices, involving many types of mass media, books, the Internet, exhibitions and other genres of communicative events that aim

to communicate lay versions of scientific knowledge, as well as opinions and ideologies of scholars, among the general public.

2.- Popularization in general, but especially the press, is not primarily characterized by specific textual structures, but rather by the properties of the communicative context: participants and participant roles such as scientific sources, specialized journalists, lay public, their respective purposes, beliefs and knowledge, as well as the relevance of such knowledge in the everyday lives of citizens.

3.- Given the concept of context in tenet number two, these context properties of popularization discourse are relevant for the linguistic analysis of the 'textual' (verbal) structures of such discourse.

4.- Popularization involves not only a reformulation, but also a recontextualization of scientific knowledge and discourse originally produced in specialized contexts to which the lay public has limited access. This means that popularization discourse must always adapt to the appropriateness conditions and other constraints of the media and communicative events.

5.- The mass media are not passive mediators of scientific knowledge, but actively contribute in the production of new, common knowledge and opinions about science and scientists – including information and views not originating from scientific sources. That is, despite their dependence on other institutions and organizations for most of their information and

_____

broadcasting, media managers and journalists ultimately decide what and how to publish (or not to publish) about science, scientists and scientific knowledge, as part of complex processes of news production.

6.- The role of new knowledge production by the mass media needs to be further contextualized in relation to other media functions especially entertainment. Mass media's role of producing new knowledge needs to be further contextualized in relation to its other functions such as entertainment.

2.4.2. Science popularization sections in newspapers

Specific newspaper sections dealing with particular areas of knowledge have recently been created to fulfill the newspaper's social function of providing information to the audience. This is the case of *Science* sections in newspapers all around the world, which have notably increased the number of science articles published in the last few years (Hyland 2010) and have reached a mass readership but have not yet reached the front pages of newspapers (Hernando and Hernando 2006: 23).

British newspapers hold a well-established tradition on popularizing science through the press that can be clearly perceived by the existence of clear-cut sections in national newspapers such as *The Guardian.* However, this is not the case in Italy or Spain, where the creation of specific sections dedicated to providing readers with specific news on international discoveries and research is

_____

only just a recent accomplishment in the evolution of written broadcasting. The Italian newspaper *La Reppublica* recently included the section *Scienze* among its daily contents and only in the last few years has the Spanish newspaper *El País* included *Ciencia* among the sub-sections found within the broader section *Sociedad* (Elorza 2010: 107). The number of newspapers with specialized science sections, as well as the number of science popularization articles published is notably increasing (Hyland 2010: 118) to such a degree that newspaper readers are now used to finding a science section previously not present in publications for general readership:

> While many popular science books are written by scientists for an elite educated audience, the public gets most of its information about science from specialized magazines like *New Scientist* and *Scientific American*. Most daily newspapers now have specialized science sections and the number of science articles in the press has been increasing. (Hyland 2010:3)

This new trend has created a section in which the journalist tries to fill the traditional gap existing between the scientific community and lay people, who constitute the newspaper's audience. The journalist closes this gap by reporting the latest national and international scientific advances and discoveries that could result in relevant changes in our daily conception of life (Calsamiglia and López Ferrero 2003: 174). As Burkett (1973: 39) asserts there are four main reasons science has to be recontextualised for the lay reader. The first reason is simply because science plays an important role in the general knowledge of human beings. Second, it is convenient that current research is explicitly and directly explained to the public. The third reason is that the work of both scientists and journalists share the common objective of providing information. The fourth and final reason is the effort to reconcile the humanities and the sciences, disciplines

_____

traditionally far from each other. Calsamiglia (2003: 140) adds the democratization of knowledge to Burkett's list of reasons for reconstructing of science popularization discourse. This fifth reason stems from the fact that news media are responsible for informing society on all subject areas. Therefore, as Kenward points out, science should not be hidden "behind an impenetrable screen of jargon and arrogance" (Kenward 1988: 31).

2.4.3. Science popularizations as objects of linguistic study

Science popularizations have recently turned into frequent objects of study for linguists. As García Riaza and Elorza (in press) state, they "started to draw the attention of linguists because of its relationship to scientific discourse and their potential to compare different recontextualizations of scientific knowledge" as well as by the fact that:

> Scientific/Academic use of language has been recognized as a specific register with its own norms, patterns and style, affecting not only terminology but ways of presentation and reasoning through particular discourse genres and procedures. No wonder, then, that the way a unit or a piece of knowledge is selected and transformed to be presented and explained to non-experts can be a very rich topic of research, since it demands rigorous recontextualization conveyed through discourse procedures affecting both global and local levels of text. (Calsamiglia and López Ferrero 2003: 148)

It is worth mentioning that the possibility for analysts to build their own corpora (O'Keefe 2002: 260) together with the availability of already compiled collections of newspaper texts (Scott 1997) have also acted as determinant factors

for the increasing number of studies carried out in the field (García Riaza and Elorza in press).

As Gil Salom (2000-2001: 442) asserts, the association of new technologies with the necessity to popularize science in the 20th century has reached such a level that readers are unable to stay updated with new discoveries. These science popularization texts create an authority that legitimizes journalists' words (Caldas-Coulthard 1994: 303) and is expressed through a series of linguistic and discursive choices that transform real facts in text (McCabe and Heilman 2007: 139).

The accessibility and quantity of texts published daily in the aforementioned newspaper sections has identified science popularization discourse as a fruitful and reliable source of information about language, but this does not exclude interferences with other text types and genres. As Myers (2003: 265) asserts, only scientific texts that are not addressed to specialists can be considered popularizations, bearing in mind that there are other different genres of scientific discourse whose texts are written by and addressed to specialists. Different studies have presented the model of a cyclical process in the transference of scientific knowledge. Moirand (1997: 189) defends the idea that knowledge conveyed through science popularization texts settles on a common background of knowledge possessed by the audience, who is no longer unaware of scientific advances. Newspaper's scientific sections have built a cultivated audience who knows about the areas of science and technology that concern them (Myers 2003: 268 quoting Irwin and Wynne 1996) and can relate scientific abstractions to their

_____

experiential knowledge. In this cyclical model, the journalist is conceived as a 'mediator' (see Figure 3 below), and scientists as 'social agents', people with important knowledge to reveal to society.
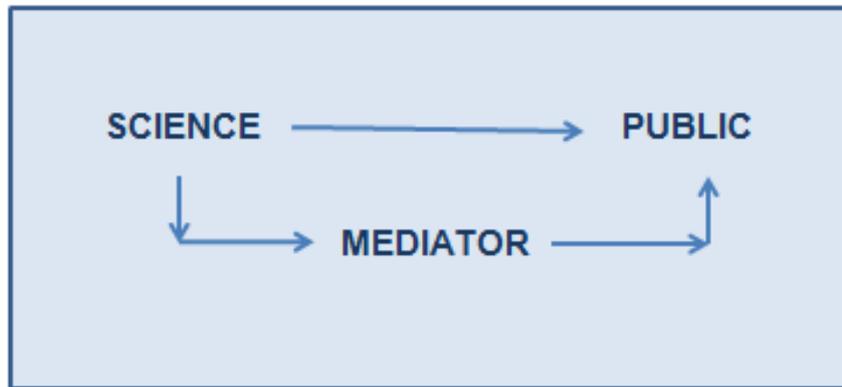
Figure 3. The role of the journalist as a mediator in science popularization articles
(Moirand 1997: 176)

Other contrastive analysis between scientific texts and science popularizations have also been carried out in the last years to characterize both text types (Myers 1990), while some others have concentrated on the analysis of a single feature, like William's analysis of metaphorical expressions of cancer (Williams 2009).

2.4.4. A hybrid scientific communication mode

The relation between neighbouring genres has always been a topic of research as well as a topic of disagreement for linguists and discourse analysts. But, as Myers (2003: 270) significantly posits, the discourse of science and

_____

science popularization can be conceived as a continuum, where we can just differentiate between two distinctive registers dependent upon the audience. In this way, the continuum in scientific communication is based on different modes of speaking for different rhetorical purposes. In the lines of Myers, Calaresu and Colti add that science popularization belongs to what they call the 'secondary' scientific communication, and is carried out at a 'popular' level

> where the scientific knowledge is broadcasted to a wider lay public trough the mass media such as television, or, in the case of written transmission, through the daily or weekly press, or through publications specifically designed for mass scientific popularization. (Calaresu and Colti 2007: 529)

Science popularizations stand in the middle of a cline that goes from media to academic discourse, passing through newspaper discourse and scientific discourse respectively, to arrive at the central position where science popularization discourse receives the influence from both parts of the cline as seen in Figure 4 below, where Elorza (2011) illustrates the process she calls 'the intertextual popularization process'.
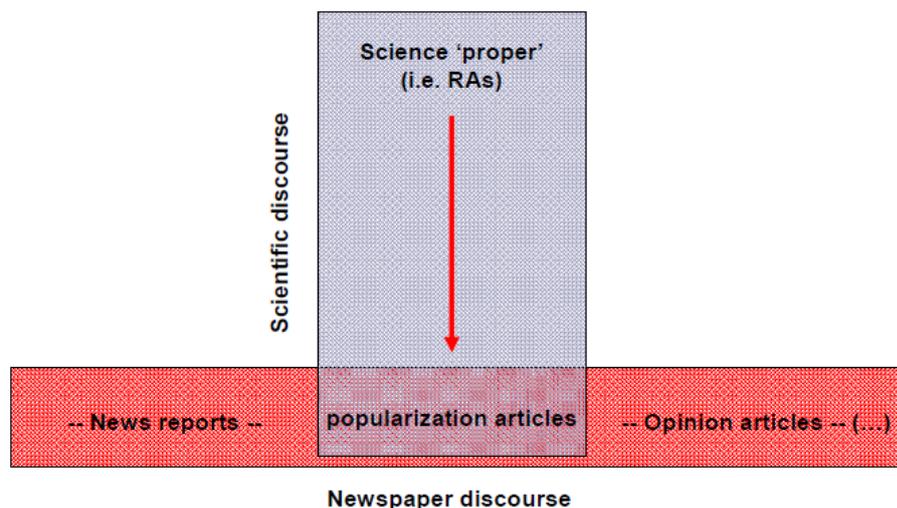
Figure 4. Interrelated character of scientific and newspaper discourse (Elorza 2011)

Conversely, Myers (2003: 266) indicates that the traditional view on this phenomenon is that of dominance, that is to say, that there are two different discourses, separate from each other and belonging to different communities: the experts and the lay people. This traditional 'dominant view' also states that the science popularization discourse is a 'translation' from the discourse of science. Scientists are considered the 'translators' in their condition of producers of what is known as scientific knowledge. According to this approach, communities of lay people are realms of ignorance when it comes to science and scientists remedy the situation by sharing knowledge unidirectionally. This dominant view of scientific discourse prevalence over that of science popularization brings about a series of textual transformations that are also conceived as qualitative downgrading. The translation of one discourse into another means adapting scientific knowledge to the general public. This does not only entail changes in form, but also changes in content of a discourse which ends up being distorted, simplified and dumbed down. In other words, science popularization discourse has traditionally been considered an adapted version of scientific discourse for ignorant readers who would otherwise not understand the highly specialized science discourse. From this perspective, the journalist writing science popularization articles is a mediator who serves, not only as a vehicle of content, sent from scientists to the lay people, but also as a kind of 'translator' who, by adapting scientists' words, is also creating a specific register with its own norms and conventions.

_____

2.4.4.1. Scientific discourse versus Science popularization

discourse

As we have previously seen, newspaper discourse is a direct inheritor of the characteristics of media discourse, which constraints the ways in which newspaper news, either printed or broadcasted through internet, is constructed. Much the same happens with academic and scientific discourse, or better said, academic scientific discourse, which aims to present facts by providing readers with evidence of the writer's assertions.

When we identify a text as scientific discourse, we are drawing our attention to a series of characteristics that identify it and, at the same time, distinguish its language from that of other kinds of discourse. As Halliday (2004: 143) states, it is the combined effect of the distinctive features of scientific discourse together with the relationships that such features establish throughout the text that drives us to think that what we are reading is indeed the discourse of science. Scientific discourse creates a very different view of the world than that of spontaneous conversation. In contrast to science popularization discourse, scientific discourse reshapes experiences into scientific theories addressed to a specialized readership through 'technical language', which "compacts" and "changes" the nature of everyday words (Halliday and Martin 1993: 172). The language of science is thus fraught with technical words that make it a discourse too complex to be understood by a common readership. As Halliday (2004: 131) asserts, readers have to be experts in order to understand scientific discourse, it is not enough to

_____

be an educated person to grasp the meaning of the language of science, which has been written by experts for their peers. Moreover, scientific language is impregnated with the authority that the 'possession of knowledge' gives experts, who exercise the power of possessing privileged information. Scientists convey this privileged information via the so-called "scientific metaphor" (Halliday 2004), which is not just a construal of meaning, but a re-construal by which information structure is challenged and re-aligned:

> Every scientific text, however specialized and technical, contains a mixture of levels of wording, from most congruent to most metaphorical, right up to the end. All scientific registers, likewise however specialized and technical, construe the full metaphoric range of semantic space opened up by their own histories, right up to the present (Halliday 2004: 121).

As Adam Smith highlights, another distinguishing feature of science popularization discourse is that unlike academic research articles, the topic takes second place to editorial treatment in popularizations, frequently adding a new context in order to frame the information from a different perspective. In a contrastive study carried out by Myers (1990: 1994) he argues that facts are endowed with an authority in science popularization that texts do not always present in scientific articles. He concludes that this is so because the narrative style of popularizations emphasizes "the immediate encounter of the scientist with nature" (Myers 1994: 179). Myers calls this narrative style a 'narrative of nature'. In contrast, he recognizes that the narrative style of most scientific research reports emphasizes "the concepts and techniques through which the scientist conceives of and delimits nature", naming this narrative style a 'narrative of science':

_____

> The professional articles create what I shall call a narrative of science, they arrange time into parallel series of simultaneous events all supporting their claim, and emphasize in their syntax and vocabulary the conceptual structure of the discipline. The popularizing articles, on the other hand, present a sequential narrative of nature in which the plant or animal, not the scientific activity, is the subject, and in their syntax and vocabulary they emphasize the externality of nature to scientific practice. (Myers 1994: 179)

From a qualitative perspective, science popularizations do not present most of the details on the materials, methods and technical procedures which are found in scientific research articles, due to the fact that they are written in a less formal register in order to be more appealing to the newspaper readers. Popularizations also show patterns that include the presentation of findings as a breakthrough, typically by constructing a rhetorical contrast between past and present facts, and the exemplification of a Problem-Solution pattern which is sometimes made explicit in the headline (Adams Smith 1987: 635). Additionally, from a quantitative perspective, Adams Smith points out that popularizations tend to be shorter than their corresponding scientific source articles.

Fagan and Martín Martín (2004) have contrastively analyzed psychology and chemistry research articles, finding that the chemistry texts have a tendency to be much longer (3,000 and 1,100 words per article respectively). According to the data, the length of research articles varies too much between disciplines to establish an average length. Furthermore, Ozturk (2007) has studied the length of research articles in two different journals within the same discipline (applied linguistics), and has acquired data suggesting that both variability is still too high within the same discipline (the average number of words per article is of 1,220 words in one journal but of 757 words in the other) and also that the variability

within each journal is very high as well, with lengths ranging from 426 to 2,429 words in *the Journal of Studies in Second Language Acquisition* and from 145 to 1,771 words in the *Journal of Second Language Writing* (García Riaza and Elorza in press).

In this same respect, and as García Negroni (2008: 6) asserts, scientific writing has been characterized by the properties of objectivity, neutrality, impersonality and accuracy, constructing a space for expository dialogue where the author sets his/her position with respect to the scientific community and seeks for inclusion in it by the presentation of the results attained in a specific area of knowledge (García Negroni 2008: 16). Scientific discourse is a collaborative construction of knowledge, where reference to previous research and researchers in the field is a common practice that aims to confront previous studies and statements. At the same time, and as Maingueneau (1999) states, the author becomes subject of the enunciation and projects an image of him/herself, an appropriate 'discursive ethos' into the scientific community that makes him/her stand in the respectable position of somebody who possesses a rigorous and grounded knowledge in the field. One way to construct such 'discursive ethos' is by citation. Citation is a common practice that, as Hyland expounds,

> is central to the social context of persuasion as it can provide justification for arguments and demonstrate the novelty of one's position […]. By acknowledging a debt of precedent, a writer is also able to display an allegiance to a particular community or orientation, create a rhetorical gap for his or her research, and establish a credible writer ethos (Hyland 2000: 20).

An additional characteristic of scientific texts to highlight is that, in contrast to science popularization articles, research articles in scientific journals follow a

quite rigid format when presenting scientific information. The typical structure is formed by an introduction, a description of materials and methods, a list of the outcomes obtained in the study and, finally, a discussion of the results to conclude the article. As Gil Salom (2000-2001: 430) asserts, such a rigid structure would reduce the creativity of science popularization articles, which evolve in a particular way depending on the scientific findings reported or the controversy around them.

A summary of the characteristics of newspaper discourse, scientific discourse and science popularization discourse previously discussed can be seen in Table 2 below:

		Discourse type		
		NEWSPAPER DISCOURSE	SCIENCE POPULARIZATION DISCOURSE	SCIENTIFIC DISCOURSE
Characteristics	AUDIENCE	General public	Lay people, non-specialised readers who want to get information on the latest scientific findings	Scientific community, specialised readers
	AUTHOR	Journalist	Specialized journalist (*The Guardian*) vs. Scientist as author (*Nature*) (Calsamiglia and López Ferrero 2003: 149)	Scientist, expert

_____

ACCESIBILITY	Accessible in daily newspapers: written and online versions	Accesible in specialized sections created in newspapers	Available in specialized journals and websites with restricted access
PURPOSE	Contribute to provide readers with general information	Keep average citizen informed of scientific advances	Generate knowledge in a specific research field

Table 2. Comparative chart of characteristics of newspaper discourse, science popularization discourse and scientific discourse.

Taking into account Hyland's description of how citation is used (2000: 20), it is remarkable to state at this point that, as Thetela (1997: 103) points out: "If we take into account that the major goal of scientific research is to establish facts, evidence can thus be seen as a positive evaluation because of its contribution to this goal". She takes a step further and differentiates between 'topic-oriented evaluation' and 'research-oriented evaluation' in academic research articles. Topic-oriented evaluation (TOE) refers to a more restricted observation of the world, while research-oriented evaluation (ROE) can be defined as the position of the writer observing the research reported. This latter type of evaluation engages both reporter and reader in a communicative event where the writer refers to the scientific community by reporting or commenting on the results of a research process (Thetela 1997: 105). Furthermore, Thetela divides the focus of this text type into three different issues: What the writer observes, the methods employed by the researcher to study the topic, and the findings and results of the research process (Thetela 1997: 106), three issues that can be seen in Table 3.

Evaluation	Research-oriented Evaluation (ROE)	Evaluation focused on the research process carried out by the attributee	Methods → PROCESS	Usefulness
				Control
			Findings → PRODUCT	Significance
				Certainty
	Topic-oriented Evaluation (TOE)	Evaluation focused on the research topic	Not considered in Thetela's study	

Table 3. Evaluation system of research articles based on product/process values, according to Thetela (1997)

It is significant that Thetela leaves aside the description of TOE in her study. She does not consider the speaker's/writer's evaluation of the actual topic of the scientific facts being reported with the same detail as she does with ROE.

2.5. Corpus linguistics and the study of texts

The high degree of informativity of press discourse has enabled the study of large collections of newspaper texts through 'corpus linguistics', described as "a way of investigating language by observing large amounts of naturally-occurring, electronically-stored discourse, using software which selects, sorts, matches, counts, and calculates" (Hunston and Francis 2000: 14-15). This has brought about a qualitative change to the understanding of language at the first stages of

_____

linguistic research (Tognini-Bonelli 2001: 1). The scientific analysis of spoken

and written discourse has provided analysts with large amounts of data for studies

on different language varieties. These studies also provide analysts with a much

more objective perspective of language making it possible to dismiss subjective

assertions and invented examples (McEnery and Wilson 2001: 103) while

benefiting from real data extracted from large collections of texts that make

language "look rather different when you look a lot of it at once" (Sinclair 1991:

100). Here Sinclair refers to the new perspective that results from examining

language with computerized tools. This is the beginning of computerized corpus

studies, which has heightened interest and increased the number of studies on how

language is used in a certain way in different types of texts and genres.

The discourse of news media is now widely studied in computerized

research and considered to be a productive source of information on language

both for discourse analysts and corpus linguists, some of whom have set their

focus of attention in the representation of scientific knowledge in science

popularization texts. Some contrastive analyses between scientific texts and

science popularizations have been carried out to characterize both text types

(Myers 1990) while others have concentrated on the analysis of a single feature,

like William's analysis of metaphorical expressions of cancer (Williams 2009).

Studies like Calsamiglia and López Ferrero (2003) and Myers (2003) center

their attention on what they call 'science for the general public' or 'popular

science' to set the object of analysis on the popularization of science, normally

carried out by a series of journalists who write in scientific sections and act as

mediators between scientists and lay people (Gil Salom 2000-2001: 443). In this

_____

text type, scientists have a mediated relationship with the audience, who reads about their research advances and discoveries through the words of the journalist reporting them.

In this chapter, we have examined the different characteristics of newspaper and science popularization discourse by comparing them to media and scientific discourse, with which they share some common features. Taking a step further, we outlined the specific features that characterize science popularization discourse and its popularity for text analysis in the last decade. More specifically, we identified one of the most common features of newspaper discourse in general, and of science popularization discourse in particular. This feature will be discussed in more detail in Chapter 3: the attribution of information to voices different from the journalist.

Chapter 3

The encoding of attributed voices in the text

_____

Chapter 3

The encoding of attributed voices in the text

3.1. Introduction

3.2. Text as an interactive representation of language

3.3. Evaluation as an inherent feature of discourse

  3.3.1. Averral and attribution: two sides of the same coin

  3.3.2. Processes to encode attribution in the text

    3.3.2.1. Reporting

    3.3.2.2. Epistemic Stance adverbials

"We all have a need to explain the world,

both to ourselves and to other people..."

David Straker, Changing Minds

3.1. Introduction

This chapter explores the notion of evaluation as the umbrella term for the writer or speaker's behaviour towards linguistic events. We will examine various models that have been built to describe the phenomena of evaluation and their different implications for the analysis of written newspaper texts, our main focus of concern. We will also focus on two specific ways of expressing evaluation: reporting and stance adverbials, two different options available for the science popularization writer to manifest his/her position towards the narrated events.

3.2. Text as an interactive representation of language

As Sinclair (2004: 53) states, language models experience a process of continuous internalization of reality within two different planes. On one hand, participants in language exchanges interact among each other, therefore constituting the 'interactive plane'. On the other hand, language creates a record of experience which belongs to the individual, and thus to the 'autonomous plane' (Sinclair 1981). This is a dynamic view of language where experience is recorded and internalized through discourse allowing the writer to build the text by meaning negotiation and its evaluation. As a counterpart to Sinclair's model, Halliday establishes a Systemic Functional model (Halliday 1985, Halliday and Mathiessen 2004), which considers a set of three different dimensions to interpret the different ways in which language constructs meaning, respectively called Ideational, Interpersonal and Textual. The ideational dimension refers to the construction of experience, and how logical relations are established throughout the communicative event: when, how, what, where and why things happen. The interpersonal dimension deals with the negotiation of meaning and how people interact and show their feelings, opinions or attitudes, and the textual dimension refers to devices used to keep the information flow of the controlled text.

In this study, we will mainly focus on the interpersonal dimension of language, concentrating on the resources available for the writer and reader to position him/herself towards the information being conveyed in the text. Grammatical systems of Mood (establishing role relationships between

writer/speaker and reader/hearer) and Modality (as carrier of the writer/speaker's assessment on the certainty of his/her message) (Berry 1975: 66) will be considered part of the phenomenon analysed in this study but will not be analysed due to the limitations of its scope:

> [i]s concerned with the social, expressive and conative functions of language, with expressing the speaker's 'angle': his attitudes and judgements, his encoding of the role relationships in the situation, and his motive in saying anything at all. (Halliday and Hasan 1976: 26-27)

The interactive aspects of written text can be approached from two complementary perspectives, 'information-oriented' and 'function-oriented' methods of language analysis. The information-oriented approach, focuses on "the ways in which writers take the (imagined) reader's expectations, knowledge and interests into account in construing their text and in signalling the relationships between parts of the text" (Thomspson and Thetela 1995: 104). Conversely, the function-oriented approach concentrates "on the ways in which writers more or less overtly conduct interaction with their readers, particularly by assuming for themselves and assigning to the readers roles in the interaction […], and by intruding in the message to comment on and evaluate it" (Thomspson and Thetela 1995: 104). In this study, we follow the function oriented approach to the analysis of written discourse put forward by Thomspson and Thetela (1995). We specifically focus on evaluation which is understood to be a consequence of interaction between people, which is verbally expressed in texts in a way that shows the sharing of experience, and not only that of information (Sinclair 2004: 52). As we described in Chapter 2, Thetela (2007) leaves aside the description of Topic-Oriented Evaluation in her study. She does not consider the

speaker/writer´s evaluation of the scientific facts being reported in as much detail as she does with Research-Oriented Evaluation, which concentrates on the research process carried out by the attributee. As it has already been noted, one of the characteristics of science popularization discourse is that of introducing into the texts, the voices of scientists and experts that provide justification for arguments and demonstrate the novelty of one's position (Hyland 2000: 20). As Myers has shown, this is a form of evaluation that concentrates mostly on the scientific finding rather than on the process. The main focus of this dissertation involves a complementary vision into the study of evaluation carried out by Thetela (2007), which focuses on the ways in which the writer introduces experts' voices into the text.

Evaluation can be defined as 'the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about' (Hunston and Thompson 1999: 5). However, evaluation is a very broad term which encompasses a great variety of phenomena. Channell (1999: 39) has defined the evaluative function in texts as 'whatever carries the expression of the speaker or writer's attitude or emotional reaction,' while Conrad and Biber (1999: 57) state that evaluation refers to the different ways in which "speakers and writers convey their personal feelings and assessments in addition to propositional content". All these definitions make reference to one specific phenomenon that has been called 'internal evaluation' (Labov 1972: 370-5), that is to say, a functional element present at different points of the text which marks, in a grammatical, semantic or prosodical way, elements

in the text through which the author has chosen to convey his/her point of view. In Cortazzi and Jin's words, 'internal evaluation' is,

> […] a rhetorical underlying that marks off the evaluated element grammatically, semantically or prosodically. The speaker can mark almost any element to make it stand out from the rest of the text in this way. (Cortazzi and Jin 1999: 107)

'Internal evaluation' has to be differentiated from a section found in the 'Problem-Solution' pattern of texts, a piece of discourse where the author sets the point of view in the text to proceed towards its conclusion, and which forms "a secondary structure" that concentrates on the evaluation section of the text (Cortazzi and Jin 1999: 107). In this study, we are focusing on the internal evaluation in which the author explicitly assigns a mark to the element carrying the evaluative meaning. Evaluation is thus a marked option within discourse, and everything which is not explicitly marked can be considered as non-evaluative. Evaluation is a phenomenon which lies completely on the writer´s desire to emphasize the presence of a subjective element and its manipulative power to influence the reader's construction of meaning while reading the text.

The relevance of evaluation lies in the fact that a single phenomenon performs a wide range of different functions which are able to operate simultaneously in a text, and have no reason to be mutually exclusive. The multiple functionality aspect of evaluation has been emphasized through different studies developed in the field, among which we would like to highlight Hunston and Thompson's model (1999), that portrays evaluation as having a triple functionality in the text.

Firstly, evaluation expresses the writer's opinion and feelings about something, thus reflecting the value system of his/her community and transferring it into the text, which conveys the writer´s ideology. Secondly, evaluation creates a relationship between the writer and the reader, that is maintained throughout the text. Through this relations, the writer exploits the resources of evaluation in a different way to establish a variety of interactions with the reader that vary from manipulation to hedging or politeness. Persuasion and hedging are key terms when dealing with the relation established between writer and reader in a text. Persuasion refers to the writer´s manipulation of the reader to perceive things in a particular way, while hedging refers to the way in which the certainty of a statement is expressed and adjusted through evaluation. Finally, evaluation serves as an element to organize discourse, a monitoring device to establish a mutual awareness about the flow of narratives. As we can note, the three different functions of evaluation pointed out by Hunston and Thompson (1999: 6-10) can easily be related to the three language metafunctions that Halliday put forward (Halliday 1985, Halliday and Mathiessen 2004) in his model: 'Ideational', 'Interpersonal' and 'Textual'. This model, as presented in Figure 5 below, refers to the capacities of evaluation to organize discourse, build a relations between the writer and reader and express the writer's opinion.

Functions of evaluation (Hunston and Thompson 1999)	Metafunctions in SFG (Halliday 1985, Halliday and Mathiessen 2004)
Express the writer's opinions about something	IDEATIONAL
Build relations between writer and reader	INTERPERSONAL
Organize discourse	TEXTUAL

Figure 5. Functions of evaluation and metafunctions of language.

3.3. Evaluation as an inherent feature of discourse

As we previously mentioned, the presence of evaluation in texts is a pervasive fact of discourse. However, this presence does not imply that evaluation is always visible or explicit to the eye of the reader. Evaluation can be conveyed to the writer through a wide range of options in order to make it more or less overt. A fact that is worth noting here is that there is no correlation between explicitness and degree of manipulation, that is to say, an overt evaluation can be less influential than an implicit one and vice versa.

Evaluation can be recognized in texts by identifying three different patterns that signal the presence of an overt mark made by the writer, who wants the reader to be aware of his/her own positioning regarding a statement. According to Hunston and Thompson (1999: 21), evaluation entails the comparison of one textual element with another by using a range of comparators (adjectives, adverbs

of degree, comparator adverbs) which make explicit that the value of the compared element is being contrasted to another. Evaluation comes from a single source, and that makes it a subjective device by means of modals and markers of uncertainty, as well as markers of value to convey a specific evaluative meaning, to the reader.

The disclosing of the writer's attitude towards the meaning constructed in texts has not yet reached a common denomination. Studies on this field have focused their attention on several different aspects of the same phenomenon, as seen in Table 4 below. While Hunston (1994, 1999) talks about the phenomenon of 'evaluation', Martin and White (2005) and Martin and Rose (2003) prefer to call it 'appraisal', Chafe and Nichols (1986) 'evidentiality'and Conrad and Biber (1999) call it 'stance'. Additionally, Labov (1984) talks about one specific aspect of evaluation, 'affect', Hyland (1996) deals with 'hedging', and Bednarek (2006) refers to the aspect of 'sourcing', as 'attribution', as 'epistemological positioning' and as 'epistemic modality' (Stubbs (1996).

Approaches to evaluation	'entity focused'	'proposition focused'
Chafe & Nichols 1986		evidentiality
Ochs & Schiefflen 1989	affect specifiers	affect intensifiers
Biber & Finnegan 1989	affect	evidentiality
Wierzbicka 1990	emotion	
Bybee and Fleischman 1995	evaluation	modality
Niemeier & Dirven 1997	emotion	
Conrad & Biber 2000	attitudinal stance	epistemic stance
Hunston & Thompson 1999	Opinions about entities	opinions about propositions
Hunston 2000	'status' and 'value' on the 'autonomous plane'	'status' and 'value' on the 'interactive plane'

Table 4. Approaches to evaluation in Martin and White 2005: 39.

Evaluation can convey several specific meanings when used in texts. Some of the most common features of evaluation are those of expressing goodness (how good or bad the writer considers the propositions of the text), importance, expectedness or certainty (the expression of how sure the writer is on the truth-value of the statement claimed). As Conrad and Biber point out (1999: 57), the term 'stance', or evaluation, is used to cover the assessment and expression of personal feelings mainly in three domains:

- Epistemic stance refers to the certainty-doubt of the proposition, its reliability and limitations, including the writer´s comments on the source of the information provided, a domain to which we will devote a deeper insight in this study.

- Attitudinal stance covers the expression of the speaker's attitudes, feelings and judgements on the value of the knowledge conveyed.

- Style stance focuses on the manner in which information is presented to the reader.

One of the main difficulties in depicting evaluation is being able to refer to different planes of discourse. As Hunston (1999: 182, quoting Thetela 1997) points out, not all evaluation in a text is focused on the same element. It is possible for multiple things to be evaluated in a single text. They differentiate between 'discourse-entities', (the evaluation made on pieces of discourse themselves) and 'world-entities' (the evaluation of elements external to discourse). This distinction can be related to the differentiation made by Sinclair (1981) between interactive and autonomous planes of discourse. Whenever a

discourse act within the discourse itself is evaluated, we can say that evaluation is happening at the interactive plane, while the evaluation of an external entity should be placed in the autonomous plane of discourse. In Hunston's words:

> The ideological space of discourse is constructed both by the way the world is labelled (evaluation on the autonomous plane) and by the way the argument is constructed (evaluation on the interactive plane). What counts as knowledge or as a valid argument (interactive status) is as important as what the world is seen as made up of (autonomous status). (Hunston 1999: 205)

As a counterpart, Martin prefers the term 'appraisal' to refer to the "semantic resources used to negotiate emotions, judgments, and valuations, alongside resources for amplifying and engaging with these evaluations", (1999: 145). This is a narrower term that refers to the notion of 'lexicogrammar' put forward by Systemic Functional Linguistics, to meanings conveyed by the writer. This concept has also been pointed out by authors such as Cortazzi and Jin (1999: 107). Martin and White (2005) establish a complex taxonomy to classify the different variants of appraisal found in texts. They have set their model of appraisal within Systemic Functional Linguistics by stating that appraisal can be classified among the resources in the interpersonal dimension of language meaning.

Apart from making reference to this classification, it is relevant to highlight that Martin and White (2005: 9) distinguish three language strata in which language is realized, comprising three different levels of abstraction: phonology and graphology, grammar and lexis, and discourse semantics. The notion of appraisal must be taken into account when focusing on the third level, which concentrates on the meaning beyond the clause at the textual level. Discourse semantics is concerned with certain aspects of discourse organization such as how

people, places or things are introduced and kept track of within the text, a notion which is similar to the concept of evaluation we are dealing with in this chapter. In their model of appraisal, Martin and White (2005) also provide the following reasons to justify the inclusion of appraisal into discourse semantics:

- The realization of the writer's attitudes usually scatters over discourse, disregarding grammatical boundaries.

- A particular attitude can be expressed through a range of different grammatical realizations.

- The existence of a grammatical metaphor, which, according to Halliday (1985: 321), is a verbal transference; a variation in the expression of meanings which involves a non-literal use of a word, and implies that for any given semantic configuration there is one congruent realization in the lexicogrammar. In particular, metaphor is an irregularity of content that consists of the use of a word with a different meaning than its proper definition. Finding a relation to the word´s proper definition in terms of similarity entails a process of tension between wording and meaning where attribution and grading opinion processes are present.

Taking into account the three aforementioned reasons for including appraisal in the interpersonal dimension and discourse semantics, it can be said that appraisal is one of the major discourse semantic resources construing interpersonal meaning. The notion of appraisal is further developed in Martin and White's model as built by three different domains: 'attitude', 'engagement' and 'graduation'. We will briefly review 'attitude' and 'graduation' and later focus on

_____

the domain of 'engagement', which sets one of the frameworks of reference used in this study.

'Attitude' deals with the feelings, emotional reactions, judgement of behaviour and evaluation of entities. The domain of 'attitude' is further divided into three different regions of feeling: 'affect', 'judgment' and 'appreciation'. 'Graduation' refers to the up and down-scaling quality of engagement, whose locutions can be graded with greater or lesser degrees of positivity or negativity. Graduation acts, accordingly, both in scaling intensity or amount, and in measuring prototypicality and preciseness, according to category boundaries.

As one of the main areas of focus of this study, the 'Engagement' dimension of appraisal has been said to deal with

> […] the ways in which resources such as projection, modality, polarity, concession and various comment adverbials position the speaker/writer with respect to the value position being advanced and with respect to potential responses to that value position – by quoting or reporting, acknowledging a possibility, denying, countering, affirming and so on. (Martin and White 2005: 36)

As we can see in Martin and White's model (2005), 'engagement' is a very complex dimension of appraisal that provides the writer with an ample set of possibilities to reveal his/her position towards the meaning in the text. Among the many resources offered by 'engagement', we will explore explicit projection accomplished through quoting and reporting in science popularization texts, together with the presence of epistemic stance adverbials.

'Engagement' is a pervasive characteristic of texts, where all utterances may be considered attitudinal in some way or another. This is because it allows the speaker or writer to manifest their position or opinion, towards the different

statements made explicit in the text. Two relevant concepts to point out at this stage are 'heteroglossia' and 'monoglossia'. Bakhtin (1981) has defined monoglossic events as those where the communicative environment is built with a single voice that does not consider references to other viewpoints. This monoglossic mode is thus in Bakhtin's words (as quoted in Martin and White 2005: 99), 'undialogised,' whereas, heteroglossia is a 'dialogised' mode, in which different voices are invoked in the text, and the writer/speaker presents propositions as one with alternatives that can be engaged in the communicative context. Martin and White go a step further in Bakhtin's analysis and make a deeper classification of heteroglossia, sub-dividing it into 'dialogically expansive' and 'dialogically contractive'. When heteroglossia allows other voices to be considered as alternatives to the writer/speaker's voice, it is said to have an expansive meaning, while heteroglossia functions as a restrictive device when it acts as a challenge or deflects the validity of the locutions encoded in the text, thus contracting the scope of plausible options available to the reader/listener (Martin and White 2005: 102).

As a result, engagement deals with the ways in which position is achieved linguistically, by standing against, for, or neutral with respect to other voices and alternative viewpoints in the text. By acknowledging other voices, the author engages him/herself with alternative positions and locutions which play a role in the communicative event, as in: *I'm kind of upset by what you said* or *He's a true friend* (Martin and White 2005: 94).

Within the system of engagement, Martin and White differentiate four resources to mark the aforementioned inter-subjective stance. The first of those

_____

devices is 'disclaim', which positions the writer/speaker as rejecting a viewpoint that s/he considers contrary to his/her own. Conversely, we find 'proclaim', or the explicit support of the text's author towards the different voices that have been encoded in the text. By proclaiming a point of view or statement, the speaker/writer is personally warranting the validity or plausibility of the proposition, suppressing the possibility of further alternatives and positively influencing the listener/reader towards the proclaimed voice. The third device used to present the author's positioning is 'entertain', a resource by which the proposition taken into account in the text is presented as one of the many possible options, and thus considered more as a subjective opinion than as a well-grounded statement. Lastly, the fourth of the devices considered in this approach is 'attribution', a process by which the inclusion of external voices in the text is explicitly marked as coming from a source other than the author. As a result, the information provided is grounded on an external voice that may not bear relation to the writer/speaker's opinion. Martin and White´s system of 'engagement' model (2005), used to mark the inter-subjective stance in texts, can be seen in further detail below in Figure 6:
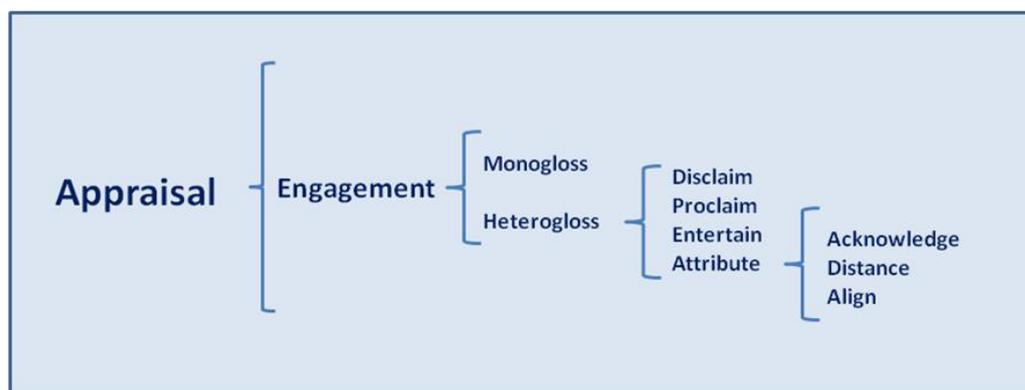
Figure 6. The system of engagement (Martin and White, 2005).

Having described Martin and White´s model (2005), we shift our focus to the heteroglossic dimension of attribution, which can be defined as the phenomenon by which the text's internal authorial voice is dissociated from the propositions conveyed by attributing them to an external source (Martin and White 2005: 111). It is important to note that this greatly overlaps with the concept of 'entertain', which presents the internal voice of the writer/speaker as a source using verbs which express mental processes (*believe, think...*) and circumstantials (*in X's view...*) (Martin and White 2005: 111). The overlapping of Martin and White's concept of 'entertain' with the phenomenon of evaluation seems to be more closely related to the concept of 'averral' than to that of attribution. Sinclair (1988) stated that texts aver anything which is not explicitly attributed to another person and, therefore, 'entertain' can indeed be considered a form of mental averral, in which the writer/speaker makes his or her own opinion and feelings explicit.

Within the system of attribution developed by Martin and White (2005), three additional categories can be distinguished. An attributed locution in a text can be said to be 'acknowledged' when there is no explicit indication of the attitude that the source has with respect to the proposition. The main feature of this type of attribution is the presence of reporting verbs. These verbs act either as associations, or also dissociations of the responsibility of the attributee on what is being reported, as in Example 1 below taken from *Sci_TG Corpus* compiled for this study:

(1) *The US said it was in Bali to be "constructive" and wanted the meeting to agree a roadmap to a new agreement on climate change which would be concluded by 2009* (Sci_TG0706).

Acknowledgements are obviously heteroglossic, as they establish an overt relationship between voices in a dialogic communicative event. Attributions however, can also fulfil the function of 'distancing' the attributee's voice from the statement reported. Several textual elements explicitly mark the distancing of the voice, such as the verb *claim*, which, in Caldas-Coulthard's words "detaches him/herself for the responsibility of what is being reported" (1994: 295). Evaluative elements are not rare devices in discourse. On the contrary; discourse is full of evaluative elements, some of which help to attain the goals of a text and are in essence, positive, while others negatively evaluate entities in a text (Hunston 1989: 204).

When dealing with the management of interaction in written discourse, it is also relevant to point out the concept of the 'reader-in-the-text' (Thompson 2001, Thompson and Thetela, 1995), a notion which takes into account the inclusion of the reader in the unfolding of the text's message, and his/her role as participant in a relationship with the writer. This relationship can be achieved through both 'interactive' and 'interactional' resources in the text. Interactive resources are those which help the reader manage the flow of information throughout the text, while interactional resources encourage the reader to collaborate in the development of the text; a collaboration achieved through processes such as the voice attribution included in the text, or through the use of modality and evaluation. As Tadros (1985) points out, the reader´s voice, categorized as having

interactional purposes, also serves an interactive purpose by predicting the next steps in the text. It is important to bear in mind that texts are not only built to satisfy the writer's expectations, but also to let the readers take part in the process by recognizing their role in a dialogic use of language where questions, whether real or imaginary, are constantly being asked and answered (Bakhtin 1986). As Thompson asserts:

> Not all hypothetical members, by any means, are attributable to the reader-in-the-text; but it is an inherent part of the pattern that they need a source - someone who is projected as (potentially or actually) entertaining the belief which the Real member then contradicts, qualifies or confirms. This someone may be the writer, the reader, other people, or any combination of the three (2001: 64).

In his view, attribution is seen as inherent to the dialogic relation between reader and writer, characteristic as well of academic text production and narratives in general, where writers need to express who the narrative belongs to and who performs the evaluation by the use of three different overt options: commands, questions and statements. These three different options are present in the text because the writer needs to justify his/her decisions against possible criticisms (Cortazzi and Jin 1999: 109). These types of interactive signals constitute a form of meta-discourse as they enclose writers' reflections on pieces of text, on which she/he makes comments by using modality and attitude markers closely related to hedging and evaluation (Thompson 2001: 73).

Another aspect of the author´s position towards the expression and evaluation of knowledge, is 'epistemological positioning', which has been characterized as the linguistic expression of assessment on knowledge (Bednarek 2006: 635). It deals with the identification of the source of knowledge, its basis

and validity, and is very similar to the notion of 'evidentiality', which Bednarek defines as "the linguistic marking of evidence" (Bednarek 2006: 635).

Anderson (1986: 273) clarifies the concept of evidentiality by asserting that not all that seems to have an evidential meaning is in the end a sign of evaluation, but that evidentiality has to be considered as a special phenomenon which shows the justification for a claim that has been made. Evidentiality is not encoded as the main predication of the clause, but as a specification added to a claim which has an indication of evidence that is not an inferred function, but a primary meaning. Evidentials include "all linguistic forms that mark the writer/speaker's basis of knowledge as something seen, heard, inferred or told" (Bednarek 2006: 636). However, as Mushin (2001) asserts, it covers much more than evidence marking. It also involves the writer´s attitude towards meaning in the text and the information conveyed, which can be expressed as certain, doubtful, authoritative, valid or reported, among others. As some authors (Bednarek 2006: 338; Biber, Johnasson, Leech, Conrad and Finegan 1999; Chafe 1986a) have put forward, epistemological positioning involves a series of elements covering the wide range of questions that the reader asks himself/herself when reading the text: Evidentiality answers questions concerning the basis of the knowledge expressed through the text, while 'epistemic modality' deals with the writer´s degree of certainty regarding the content expressed. 'Mirativity' deals with the degree of deviation that the knowledge has from the writer's expectations, and, finally, 'extent' is concerned with the possible limitations of the knowledge.

In regards to the terminology applied to the reliability of knowledge reported in texts, Cornillie (2009: 45) states that evidentiality can be easily

_____

confused with epistemic modality due to the functional similarities of both phenomena. He stresses that evidentiality, alludes to the "perceptual and/or epistemological basis for making a speech act", while epistemic modality refers to the "evaluation of the chances that a certain hypothetical state of affairs under consideration (or some aspect of it) will occur, is occurring or has occurred in a possible world" (Cornillie 2009: 46, quoting Nuyts 2001: 21). In the case of evidentiality, we can distinguish between 'direct evidentials', where the speaker/writer has directly witnessed the action reported; and 'indirect evidentials', in which the speaker/writer has not directly witnessed the action, but has heard about it through other voices, namely 'hearsay makers', 'reportives' or 'quotatives.' Unlike evidentiality, epistemic modality evaluates the likelihood of the reported action. This is an evaluation that ranges from absolute certainty on the validity of the events, to the absolute certainty on their falseness (Cornillie 2009: 46).

Moving beyond the multiple interpretations of evidentiality, we will now focus on attribution. According to Hunston (1999: 181), attribution is only a small part of a larger phenomenon, called 'the ascription of information or opinion in a text to sources, which may be animate or inanimate'. These different sources have also been called 'evidentials'. The concept of 'evidentiality', or 'the kinds of evidence a person has for making factual claims' (Anderson 1986: 273), is also taken into account as a special grammatical phenomenon that is not exclusively based on those elements considered to have an evidential function. Evidentials are competing choices which justify a claim made by a person who had the information contained in the claim, available to them. Evidentials are not

_____

generally embodied in texts as predication of clauses, but as explanations of the claims made, containing their specification as the primary meaning of clitics or free syntactic elements.

(2) *The author is* allegedly *a member of a comedy troupe and presumably was trying to be witty* (Cornillie 2009:46)

(3) *It* seems *to be a good movie* (Cornillie 2009:46)

The justification of a claim through evidentials can be expressed as direct evidence plus observation, evidence plus inference, inference, or expectation, reasoned from logic or facts. However, it can also be omitted when the fact that is claimed is observable by both speaker and listener, a case where its presence would be regarded as redundant (Anderson 1986: 274, 277).

In the same respect, Hunston and Thompson (1999: 3) describe how the writer conveys his/her opinion on the probability of events happening by using the traditional term 'modality'. This concept overlaps with the concept of 'evidentiality' (Chafe 1986a), but has been modified by Martin and White (2005: 2) to accompany the term 'epistemic modality' referring to 'how the textual voice positions itself with respect to other voices and other positions'. It is worth noting that epistemological positioning is concerned with both attribution and averral. As Bednarek (2006: 646) states, it is necessary to distinguish between sourcing, which is the identification of the Sayer, Senser or Experiencer (according to the Systemic Functional Grammar), and evidentiality, which identifies the source of knowledge, as we shall later see.
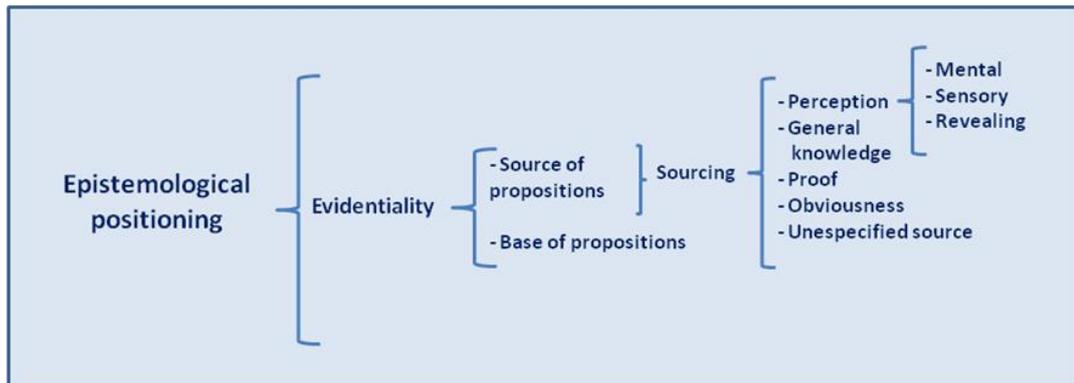
Figure 7. Epistemological positioning system (Bednarek, 2006)

A further distinction which I want to highlight is the one Bednarek (2006) establishes between 'source of propositions' and 'basis of propositions'. With the first term, she refers to the person responsible for the information conveyed. In the case of popularizations, she directly points to the scientist who provided the writer with the information, and therefore, to whom the speech event is directly attributed. ´Sourcing´, or knowing who is established by the author as the source of information in a piece of news, is closely linked to the source of propositions. As we will see in the taxonomy presented by Thompson (1996a), reproduced in Section 3.3.2.1, sourcing is normally directed to the writer him/herself or to others, specifying the source in different degrees. Conversely, the latter refers to the evidence of the source of knowledge where or how the evidentiality was retrieved by the writer. After a newspaper corpus analysis, Bednarek (2006: 640) found that there are five possible bases for the information conveyed through sourcing, namely perception, general knowledge, proof, obviousness or unspecified source. Perception refers to the knowledge grasped through three different means: a mental, a sensory, or a revealing process.

Mental perceptions are commonly expressed by verbs such as *seem, appear, look,* and even verbs such as *infer.* Sensorial perceptions refer to the acknowledgement of information that a human being sees or hears through his/her senses, whereas revealing perceptions deal with processes through which the information is revealed to the writer in one way or another. Some of the verbs conveying this meaning are *reveal, show, betray and confirm.* The next category, general knowledge, is believed to belong to a general background shared by a community of speakers or at least, by both writer and reader. Proof, however, is based upon some kind of test found to be true and relevant, and obviousness and perception, closely related to each other, refer to the obviousness of the validity of the information, either because it is self-evident or because it does not need further evidence to be tested. Finally, there is the existence of unspecified bases of information, which are those through which information reaches the writer by following a specific, but not specified process, such as in the verb *emerge.*

3.3.1. Averral and attribution: two sides of the same coin

The difference between averral and attribution is something that we cannot ignore in this study. According to Sinclair (1988), texts aver (attribute to the writer/speaker) everything which is not explicitly attributed to another person. Therefore, any attribution is a marked option in the text where the author chooses to present the source of the attributed discourse. By attributing information, whether spoken, written or thought, to somebody else, the author is presenting it as having been derived from someone different, an external source which comes

into the text and carries the responsibility of the proposition. Taking into account the framework of evaluation set by Hunston (1999: 178), attribution and averral conform the interplay between sourcing and evidentiality of language events, either written or orally constructed, which are presented as deriving from a source different from the writer or are attributed to the voice of an external source. On the contrary, stretches of language which are not attributed to a different source are said to be averred to the writer's voice. When an utterance is averred, the writer is fully committed to the content of the averral (Sinclair 1986), which is supposed to be reliable and relevant enough to be included in the text.

As Hunston and Sinclair note, "the distinction between averral and attribution […] can be used to position the reader to attach more or less credence to the various pieces of information", because "in general, a writer assumes responsibility for what is averred, but delegates responsibility for what is attributed to the attributee" (Hunston 1999: 178 and Sinclair 1986). If the writer of a text decides not to aver a stretch of language, s/he signals it by attributing it to someone else, and therefore, "it is a report from the start and the attributed proposition will be couched in a sub-report or quote" (Sinclair 2004: 57).

It is therefore crucial for the study of evaluation and the development of this study to differentiate between these two concepts, as they help identify the reader's involvement with the reported message. On the one hand, we can easily identify the reporter's voice in texts, conveying his/her own ideas and world view through the propositions present in the text. Readers normally accept the information as valid, and rely on the message that the author is conveying. On the other hand, the validity of the content which is not averred is conveyed through

the attribution of statements to respectable sources that reinforce the validity of the information in the context of the text. As Hunston (1999: 178) states, "by manipulating attribution and averral, the writer evaluates the story highly, yet evinces a scientific caution and deference to authority which her readers in turn are likely to evaluate highly". The writer is ultimately the evaluator of the reported message, which is also averred. Averrals made in a text are also attributed. Although the author makes the source of attribution the responsibility of the message reported, it is the writer who, in the end, attributes that message. As Bednarek states on this topic, "In a strict sense, the attributed proposition is thus both part of an averral and at the same time an attribution" (2006: 642). In this same respect, Thompson (1996a: 211) puts forward the term 'self-projection' to refer to a writer/speaker´s marked intention to make the proposition clear to the speaker.

Depending on the degree of alignment between knowledge and the world, a series of different ranks can be assigned to the information conveyed by the text´s author through attribution and averral. Hunston (1999: 186) presents a distinction between 'averred fact', 'attributed opinion' and 'averred hypotheses or between what is claimed as true, and what is asserted as possibly being true. The contrast is established between statements that create a hypothetical world and those which convey claims about the real world. In journalistic writing, the persuasive power of the text makes the aforementioned distinction much more obvious. Writers typically encode evaluation as a phenomenon with an evident truth-value that can be either accepted or challenged by the reader, or as an opinion that is not subject to truth-value.

As Sinclair notes (2004), an added factor in the relationship between averral and attribution is that every attribution is also averred, implying that all attributed stretches of language are, at the same time, averred by the fact that they are included in the text. It is the author who has deliberately chosen to include the attributed propositions in his/her text and therefore, he/she is ultimately responsible for their presence in the text. The relationship here established between attribution and averral is not unidirectional, but rather bidirectional. Taking into account Sinclair's assertion of the dependence of attribution upon averral, we could state that averral is also dependent on attribution. As shown in Example 4 below, every averral present in a text is a form of self-attribution or marked option through which the author conveys a meaning and attributes its credence to him/herself:

> (4) *For Dr Eggan, a young scientist of 32 who is itching to get on with research, the result has been agony and frustration: "I've spent the last three years of my life trying to get this sorted. At least a third of my time is still spent keeping the accounts and equipment separate."* (Sci_TG0601)

However, as Coulthard (1994: 6) affirms, as shown in Example 5 the author has the responsibility of remaining the evaluator of the truth-value, even when quoting.

> (5) *"My left foot is slightly larger than my right foot" [in a shoe shop]* (Sinclair 1986: 44)

A difference between attribution and based averral should be made clear at this point. Attribution refers to a person´s thoughts or speech being different than

those of the reporter (e.g. "He" in Example 6 below). Based averrals are those in which the reporter interprets the linguistic evidence him/herself in order to report it, on his/her own basis (e.g. the reporter's evaluation of the meaning of the behaviour of the poll of activists in Example 7 below) or in light of specific evidence (e.g. "Tests" in example 8 below), which stands in contrast to non-based averral, where the author plainly reports his perception on the facts reported (e.g. Example 9 below)

 (6) *He said that Mr. Duncan Smith was wrong*

 (7) *Yesterday's poll of activists showed that Mr. Duncan Smith was wrong*

 (8) *Tests found that Mr. Duncan Smith was wrong*

 (9) *Mr. Duncan Smith was wrong*

 (Bednarek 2006: 647)

An additional difference between attribution and averral is the degree of objectivity that the two phenomena display. As previously mentioned, with attribution, the writer detaches himself from the words transcribed in the text, which belong to an external voice responsible for the meaning of his/her utterances. This characteristic makes attribution much more objective than averral, which depends on the subjective interpretation of the speaker/writer. Both in the cases of based and non-based averrals, the speaker/writer turns into referent for the truth value of the meaning expressed. In the case of the based averrals, the speaker/writer is responsible for the interpretation of data. In the case of non-based averrals, the truthfulness of the information given rests entirely on the speaker/writer´s subjectivity. A visual representation of this difference in degree of subjectivity between attribution and averral can be seen in Figure 8 below.
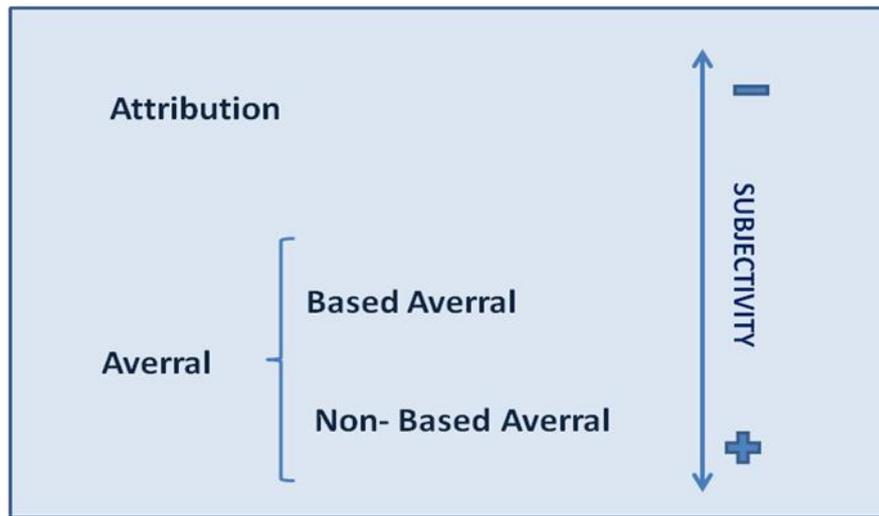
Figure 8. Cline of subjectivity of attribution, based and non-based averrals, adapted from Bednarek (2006: 649)

3.3.2. Processes to encode attribution in the text

In reference to the presence of attribution in newspaper discourse, Bell (1991: 52) states that it is typical for news discourse to mark many language events on the basis of hearsay or mindsay due to the fact that newspaper discourse consists of "embedded talk." As Calsamiglia and López Ferrero point out:

> Prescriptive guidelines for journalists stress the use of citation to give authority and legitimacy to what is said through the strict (literal) quotation of words spoken by others. It is one of the procedures by which to obtain credibility for the facts under comment as well as one of the means for certifying knowledge of what is being said. Literal quotation is intended to be objective knowledge, far removed from the subjectivity of the journalist. (Calsamiglia and López Ferrero 2003: 153)

By hearsay and mindsay, Bell refers to what the writer reports as coming from another voice, and thus it is not the writer's knowledge which is reported, but that of another specified or unspecified person or group of people. In the case of mindsay, what the writer is reporting are not actual words, but thoughts. The distinction is further developed to take into account that there may also be a

source of a 'perceiver' (also named 'experiencer'), or the person who is attributed

as having witnessed the reported facts (Bednarek 2006: 643-644). Conversely, the

Systemic Functional Grammar makes a distinction between the author of a

linguistically expressed event and that of a thought. Halliday and Mathiessen refer

to hearsay as, 'quoted speech', and its source as 'Sayer', while mindsay is

categorized as 'quoted mental experience' and its source as 'Senser' (2004: 252,

201).

As seen in Hallidayan SFG, Figure 9 below, the different types of encoding

illustrating the view of the speaker/writer towards the information conveyed can

be associated with the three metafunctions:

Metafunctions in SFG (Halliday 1985, Halliday and Mathiessen 2004)	Encoding of author's evaluation in the text
TEXTUAL	Modality
INTERPERSONAL	Attribution Engagement Epistemological positioninig
IDEATIONAL	Affect, judgement, Evidentiality, Stance

Figure 9. Related SFG metafunctions and the encoding of author's evaluation in texts

To strengthen the relationship between Halliday's metafunctions and the

phenomena which includes the evaluation of the author within the text, we have to

deal with the concept of 'register.' This concept according to Halliday and Hasan (1989: 22) is "the set of meanings, the configuration of semantic patterns that are typically drawn upon under the specified conditions, along with the words and structures that are used in the realizations of these meanings". The context of the situation which characterises genre encompasses elements such as the participants in the situation, their verbal and non-verbal actions and the surrounding objects and events that influence the verbal action. The context of situation described by Halliday and Hasan (1989: 23) can be divided into three main components: field, tenor and mode.

In this respect, as Martin (1992: 494) asserts, there is a systematic correlation between register variables and macro-functions. Therefore, field mainly influences ideational choices, and tenor influences interpersonal choices.

3.3.2.1. Reporting

Language in use is both a continuous negotiation between participants and a constant record of human experience. The negotiation carried out in discourse emphasizes the interactive character of language, and therefore, the intense activity at the 'interactive plane' (Sinclair 2004, 52). One of the means through which this negotiation can be achieved is that of reporting what has been said or thought by a person other than the writer/speaker, and then creating "another voice in the story" (Martin and Rose 2003: 23). The reported material, which may appear as direct quotes, indirect quotes or averred information in the text, is supposed to be verbatim. However, much of it seems to be indirect quotes in

_____

paraphrase, and a lot of unattributed information draws heavily on what a newsmaker has stated (Bell 1991: 61).

As Maldonado González (1999: 3551) presents, the differentiation between the original production and the reconstruction of a language event is a universal of language:

> *La posibilidad de reproducir un discurso es un universal del lenguaje. En todas las lenguas, la reproducción de un discurso es un fenómeno lingüístico distinto a la producción original del mismo; los hablantes tienen siempre la posibilidad de citar palabras – propias o ajenas – y no sólo de hacer referencia a ellas.*
> (Maldonado González 1999: 3551)

Maldonado González also covers various possibilities for carrying out the report: from the mention of a speech act, without naming the speaker, to the interior monologue, or the general description or summary of a language event, or the literal transcription of the words uttered. In this same respect, Caldas-Coulthard (1994: 295) asserts that the reporting of what other people say is a major feature of different text types, including news in press and narratives. In news reports, 'news' is what is conveyed in the text, while voices encoded in a news report are considered 'accessed voices' (Hartley 1982), accessed by the writer who then acts as controller of the text's flux. In contrast, 'unaccessed voices' are those which are neglected in the text and on which the author is making some type of discrimination: sexual, racial, of content, source, etc.

The writer in charge of deciding what is going to be said and how it is organized in a report can attribute voices to people either explicitly in a direct way or implicitly. The study focuses on explicit attribution, a term that corresponds to what Calsamiglia and López Ferrero call 'inserted citation'. This can be defined as

the presentation of discourse as deriving from someone other than the journalist (Hunston 1999: 178), or as an explicit mark that identifies the source of information (Conrad and Biber 1999: 67).

In SFG, reporting is normally carried out through 'projection'. This has been defined as the relationship between two clauses where the secondary clause is projected through the primary clause, which acts as a locution or as an idea, or as "the logico-semantical relationship whereby a clause comes to function not as a direct representation of (non-linguistic) experience but as a representation of a (linguistic) representation" (Halliday and Mathiessen 2004: 377, 441). With regard to this definition, Thompson (1996a: 209) expounds that the relationship between clauses in projection differs greatly from that of 'expansion' – where the secondary clause expands the primary clause by extending, elaborating or enhancing the meaning of the first (Halliday and Mathiessen 2004: 377). This is because projected clauses need the meaning of the first clause to make sense, while 'expansion' clauses would not change without the relationship with another clause.
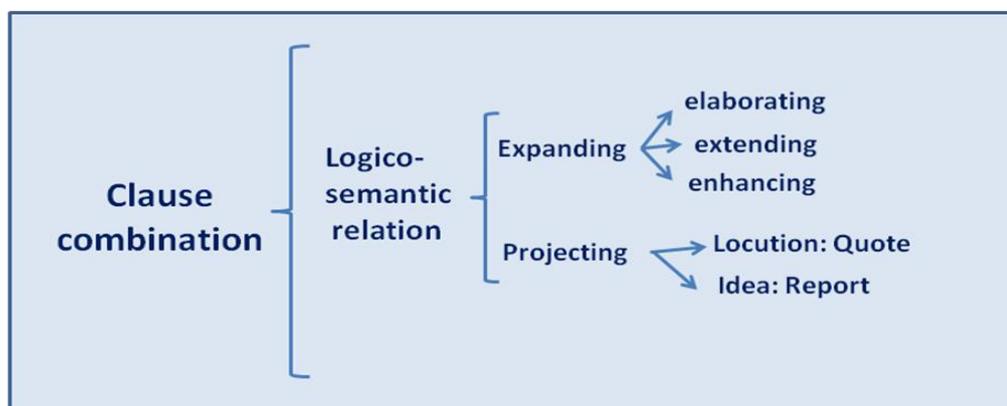
Figure 10. Logico semantic relations adapted from Thompson (1996a: 214) and Halliday and Mathiessen (2004: 373).

_____

As Halliday and Mathiessen (2004: 443) point out, some of the most common uses of projection include: attributing voices in news reporting, presenting views in scientific discourse, constructing dialogue in narrative and framing questions in conversational exchanges. Reporting structures depend on communicative purposes and the functions they aim to fulfil. Therefore they are not always encoded in the same way, as can be seen by the extensive literature available on the use of citations in academic discourse (Hoey 1993, Minelli de Oliveira and Pagano 2006, Tadros 1993, Thompson and Ye 1991, Shaw 1992, Swales 1990).

Reporting what other people say is a common feature in newspaper discourse, where the report stands for an original communicative event that has been reduced to just one informative move[1], in contrast to the three moves present in reporting fictional events: initiation, response and follow-up (Caldas-Coulthard 1994: 297). As Sinclair (2004) states, reporting is a category that relates the 'interactive plane' of discourse to the 'autonomous plane'.

One of the linguistic choices that journalists make when writing a science popularization article is that of reporting scientists' voices. Science popularization articles are characterized by presenting multiple voices that are usually conveyed through direct and indirect forms of reported speech which fulfil the function of giving authority to the words reported. As Calsamiglia and López Ferrero assert, "[Reporting scientific voices] is one of the procedures by which to obtain

_____

[1]   For further information on the moves that constitute the exchange structure, see Sinclair, J. and M. Coulthard. (1975) *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* Oxford: Oxford University Press.

credibility for the facts under comment as well as a means of certifying knowledge of what is being said" (2003: 153).

As McCabe and Heilman (2007: 139) state, events that take place in the world are expressed through a series of linguistic choices that journalists make to construct their narrations. In this sense, Thompson (1996b: 502) expresses that a guiding principle we should bear in mind in this framework is that speakers/writers have a wide variety of modes at their disposal by which they can insert language reports into text. This guiding principle should underlie every linguistic analysis and will also be considered in this study.

Nevertheless, we may still question what is actually considered reported language. Language reports have been defined as 'signalled voices in the text' as well as "any stretch of language where the speaker or writer signals in some way that another voice is entering the text, however muffled or ambiguous in fashion" (Thompson 1996b: 506). However, a broad classification of reporting signals is necessary to identify the choices made by the writer and to better understand the nature of reporting.

As Thompson (1994: 1) states, reports usually consist of two different parts: the 'reporting clause' and the 'message'. The reporting clause shows that another person's words are being reported, that the preceding speech is not the reporter's, but somebody else's. The message is the part where the words said or written are reported. In structural terms, reports can be classified into two different types: 'direct quote structures' and 'indirect report structures'. Direct quote structures convey the message in such a way that the original words (or thoughts) of the speaker or writer are used without any change at all. Direct quote structures

consist of a double clause structure. Reporting clauses contain the reporting signal and a quote containing the actual message communicated. The quote can come either after or before the reporting clause:

(10) *'I'll be back in a minute'* [quote], *she said* [reporting signal].

(11) *He asked Jenny* [reporting signal]:*'Could you please close the window?'* [quote]

The reporting clause can also come in the middle of the quote, but, in these cases, reporting clauses very often come after one complete element of the clause has been elicited (Thompson 1994: 3):

(12) *'What I still don't know', he said 'is how to get all pieces together'* [quote]

On the other hand, reports can also be classified as 'indirect report structures', a typology which often implies the presence of a 'reporting verb'. This structure also consists of two parts: the reporting clause in which the verb is included and the reported clause that conveys the message. This double-clause structure is also known as indirect speech report, in contrast to the direct speech reports mentioned earlier, and which conform the traditional notions of reported speech (Calaresu 1998: 269-270). In this study we will also consider 'reporting adjuncts', an alternative in indirect report structures used instead of the reporting clause that can be adverbs, prepositional phrases, non-finite clauses or subordinate finite clauses. 'Reporting adjuncts' are used in direct and indirect reports and are normally marked off from the rest of the sentence by commas, as it can be seen in examples below:

(13) *Apparently, his fears were not without foundation*

(14) *According to Simon, they spent an interesting evening looking at photos*

(15) *To quote McCullough: 'The ocean breezes at Long Branch did not suffice, so off he is taken to Saratoga'*

(16) *'The world is not hostile, nor yet it is friendly,' in the words of J.H. Holmes. 'It is simply different'*

(Thompson 1994: 22)

At this stage, and before concentrating on the reporting uses and structures in newspaper discourse, it is relevant to have a general look at the following three systems ruling projection: level of projection, mode of projection and speech function (Halliday and Mathiessen 2004: 443).

First, projection may be realized on two different levels, either as 'idea' or as 'locution'. Through projection, the content of a clause is projected via another clause, in which the former may be a representation of ideas or locutions. The reporting of thoughts is called 'idea', whereas the reporting of speech acts has been labelled 'locution'. As Thompson states in this respect (1996a: 210), language can not only be used to talk about things that happen in the world, but also about the language itself. When we are reporting stretches of language, or what Halliday and Mathiessen (2004: 441) call 'metaphenomena', the projection is giving an account of a linguistic representation which is normally signalled as not being our own but as coming directly from its original source. Thompson adds (1996a: 210) that quotes imply a more or less literal re-using of the wording of a different language event, that has no need to coincide in mood or register with the hosting projecting clause. Searle (1969: 29) has defined reports as "the mechanism by which linguistically expressed content is related to speech acts

performed." Conversely, reports exclusively project meanings and not words, which fit into a projecting message with which they need to coincide in some parameters. Caldas-Coulthard (1994: 296) also introduces the relationship between the pairs reporting/quoting and those of idea/locution. Reporting is directly related to idea, while quoting stands in connection with locution, the former being the form of expression of the meaning in the latter of the term pairs (see Figure 11 below).

Second, projection is also dependent on the same logico-semantic relations that generally apply to clause combination, namely hypotaxis and parataxis, and the constituency relation of embedding. When projections are represented hypotactically, that is to say, in the form of a dependent clause that cannot stand by itself without the meaning of the projecting clause, we are typically constructing 'reports'. On the other hand, when projections are represented paratactically – by means of clauses which are not mutually dependent - we are usually constructing 'quotes'. As Sinclair (1994) asserts, quoting means introducing one text into another, where the quoted text is assumed to maintain its integrity, and thus the wording of the language event reported is highlighted and protected from the requirements of the surrounding lexicogrammatical patterns. A quote can be perceived as a parallel alternative to reporting, but it is important to bear in mind the main difference between the two concepts. A quote cannot carry out the change from the interactive to the autonomous plane that reports can.

The third consideration in reference to projection modes is embedding, a relation in which projected clauses are 'rank-sifted' (Halliday and Mathiessen 2004: 467) to function as a Qualifier of a nominal group.

- Paratactical → quote → direct speech:

(17) *///"We can actually pinpoint particular genes that are responsible for providing adaptations for these organisms that are living in extreme environments," [α] /said Professor Grunden [β]///* (Sci_TG06453)

(18) *///The senior climate negotiator for the US, Harlan Watson, said: [α]/"We have problems with defining the numbers up front. In our view, that pre-judges the outcome of the negotiations over the next two years." [β]///* (Sci_TG0706)

- Hypotactical → report → indirect speech

(19) *///One microbiologist, Amy Grunden, at North Carolina State University, reported [α]/ that she had been working on a way to grow food in harsh conditions on other planets. [β]///* (Sci_TG06453)

(20) *///Chris Stringer, research leader in human origins at the Natural History Museum in London, said [α]/ that the new study was important because the Jebel Irhoud site was often neglected by palaeontologists. [β]///* (Sci_TG07216)

- Embedded clause:

(21) *///He said [α]/ the Earth was undergoing a rapid transition that could boost temperatures by 8C, making large parts of the surface uninhabitable and food production impossible. [β]///* (Sci_TG0630)

(22) *///In a study that will surprise few parents, it was found [α]/ that children preferred their cherished comfort blankets or favourite raggedy bear over duplicates apparently identical in every way. [β]///* (Sci_TG07219)

As the final system of projection, Halliday and Mathiessen (2004: 441-445) outline different speech functions of projection, emphasizing the prominence of projecting propositions and proposals over minor speech functions such as

greetings and exclamations. A complete outline of these three different projection

systems can be seen in Figure 11 below.
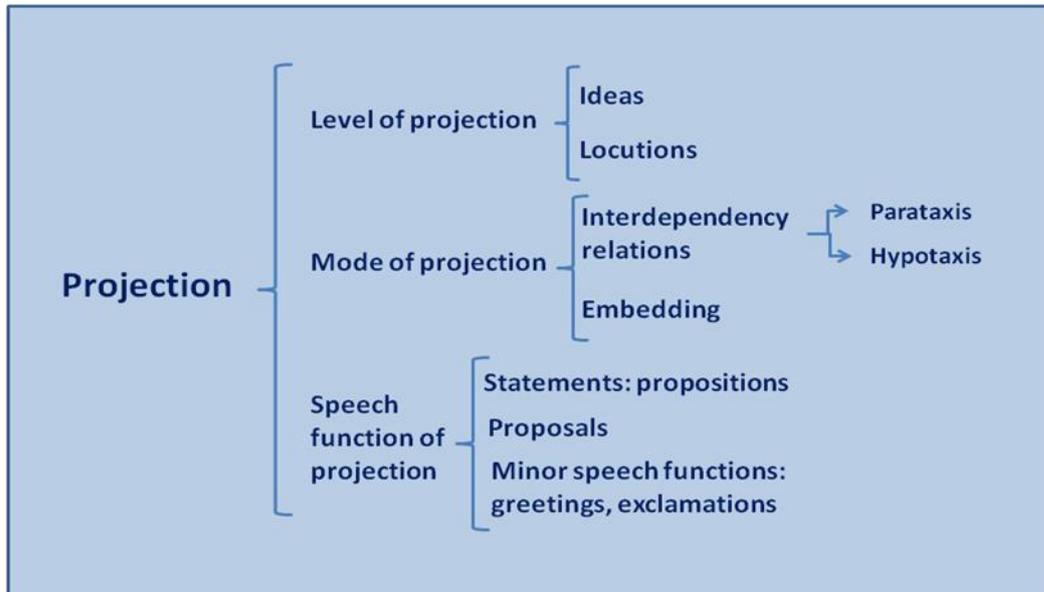
Figure11. Reports, Ideas and facts: three kinds of projection, adapted from Halliday and Mathiessen (2004: 441-445).

We will now focus on the discourse of newspapers, and as Thompson notes

(1994: 151), the reader's expectation that a newspaper report will provide a

reliable picture of the event coinciding with evidence provided. Evidence is

normally given through the presentation of information or opinions of people who

are considered important, and are therefore voiced and authorised by the author.

The frequency of reports in science popularization newspaper articles gives rise to

multiple ways of presenting scientists' discourse and integrating it into the

message conveyed. A triplet of 'citation formulae' put forward by Calsamiglia

and López Ferrero (2003: 155) creates a taxonomy of parameters to help identify

reporting cases in science popularization texts. On one hand, their classification is

based on traditional grammar making a distinction between direct and indirect

citation (used as a synonym for quotation). On the other hand, their classification is based on the way in which language events are inserted into the text (see Table 5 below).

Citation formulae	Characteristics
Direct citation	There is a fracture between the syntax of D1 and D2 because it entails the maintenance of two different deictic centres (affecting tense, space and time adverbs and person-reference words), as a result of the two different enunciations being put in relation one to another; the two segments are connected through juxtaposition and they are signalled by graphic markers such as (:)
Indirect citation	There is only one discourse, D1, with a single deictic centre, a subordinate clause introduced by a conjunction, and the correspondent agreement of tenses.
Integrated citation	It has the form of indirect citation but with segments – of greater or lesser extension – signaled as being cited directly/literally with clear graphic or typographic marking, mainly with quotation marks or marked fonts (boldface or italics). This type of citation allows mixing syntactic traits of direct and indirect style, a phenomenon rejected by grammarians but frequently used by journalists.
Inserted citation	Words of W2 are brought into the main discourse by means of markers such as 'según X' or 'para X', 'in the words of X', 'according to X' which have the function of assigning explicit words to a particular agent (literal or non-literal, depending on the use of graphic signs of quotation) without any communicative verb.

Table 5. 'Citation formulae' and its characteristics (Calsamiglia and López Ferrero 2003:155).

As seen in the table above, Calsamiglia and López Ferrero also establish a series of abbreviations that make reference to the different voices and pieces of texts that are brought together in the citation process. In their model, citation consists of a main text (D1) and a quoted text (D2); the former being the point of departure for the writer (W1) to introduce the cited voice (W2). Citation processes

are carried out within a specific framework, which is also designed by the writer (W1) with three elements (Calsamiglia and López Ferrero 2003: 156):

- The identity of the cited voice

- A communicative verb (*verbum dicendi)*

- Other elements: narratives, descriptions, modalities

These three elements of the attribution frame create a particular environment where the process of attribution is carried out. This specific environment helps both the writer and reader identify the phenomenon of citation. By including the three aforementioned elements, writers homogenize the procedures used to encode attribution into the text that provide a well-defined structure easily identified by readers.

Conversely, Thompson (1996b: 507) proposes to analyse the elements in the attribution framework along four parameters: voice, message, signal and attitude.

The term 'voice' refers to who or what is presented as the source of the language being presented. There is a set of different voices which can be encoded in the text, as the following examples taken from Thompson (1996b) illustrate:

The 'self', a voice in which the writer also holds the role of speaker:

(23) *I think he was a bit shorter than you are* (Thompson 1996b:507)

'Specified other(s)', which entails the inclusion of the voice of another person who is usually in a place and time different from the author's:

(24) *The two cricketers deserve better, as Graham Gooch admitted* (Thompson 1996b:508)

Parameter	Types		
VOICE	Self		
	Specified others		
	Unspecified others		
	Community		
	Unspeciable others		
MESSAGE	Quoted		
	Echoed		
	Paraphrased		
	Summarized		
	Omitted		
SIGNAL	Separate	Dominant	
		Equal	
		Subordinate	
	Fused		
ATTITUDE	Relevant	Negative	
		Neutral	
		Positive	
	Not relevant		

Table 6. Parameters in Language reports. (based on Thompson 1996b)

'Unspecified others', which relate to the presentation of a report without giving explicit identification of the source, even though it can be identified:

(25) *One of the women of the house allegedly flung boiling water on the crowd in the street* (Thompson 1996b:508)

This is sometimes the case with entities which have been previously revealed in discourse, and are referred to in subsequent parts of the text. (Thompson 1996b: 509)

'Community' makes reference to knowledge shared by a restricted or more ample group of speakers, either a restricted or a more ample group, that doesn't

need specification about the source, and therefore shows no explicit marking of reporting:

  (26) *The only rescuable items were a heavy rosewood desk, eastern, and a Wellington Chest whose top and side panels had split badly. Beggars can't be choosers.* (Thompson 1996b:509)

Finally, the last category is 'Unspecifiable others', where the introduction of another voice in the text impels the listener to ask for its relevance to appear co-operative. As with the community voice, the unspecified voice is general knowledge-dependent, with no explicit marks in discourse:

  (27) M*eanwhile, for all those Broke Shields fans, the ordeal is almost over. Her masterwork is coming out.* (Thompson 1996b:510)

The next language report element mentioned by Thompson (1996b: 507) is message. This element relates to the degree in which the function or content of the message presented matches the actual language event, conveyed either in the form of quotes, echoes, paraphrases or summaries. Quotes are usually recognized by the use of inverted commas or different explicit devices. They fulfil a triple function in the text. First, they show a direct relation to a different piece of discourse by exactly reproducing the words of another speaker. Second, the message conveyed through citation is more vividly presented to speakers, giving the impression of reproducing an original event. The third function of the quoted text is to detach the reporter from the content of the quote, words from which the writer distances him/herself. This procedure of detachment from the words uttered can be interpreted as follows: the writer's intention of disassociating his words

from those quoted, the writer's desire to show his/her humility in recognizing that the words quoted are the most appropriate ones to express the intended content, or simply that no other re-elaboration would have expressed his/her superiority as the author of the text. The words are quoted to explicitly mark that the writer himself would not have used those exact words. When no explicit reporting signals are found in a text that echoes the words of a source other than the writer, different functional, lexical or structural resources are used to recognize the string of discourse attributed to an external voice. Another means of conveying the message in reporting language events is paraphrasing. No trace of the original interaction is present in the paraphrased text, and the message is expressed according to the reporting context and the reporter's preferences, as in Example 28:

(28) '*According to scientists, whales in the Atlantic live longer than those in the Pacific*'

Moreover, summaries are built by a reporting word which introduces a noun or a prepositional phrase in which the message is encoded. However, there can also be cases in which the message itself is conveyed in the reporting words:

(29) *He apologized for the delay*

In the case of omission, no explicit data is given about the identity of the speaker, although a language event is actually encoded in the text. It is commonly done through the use of a "non-reporting language event verb" (Thompson 1996b: 518), or the association between a noun and a general verb:

(30) *He walked down the stairs, still muttering/ In December Pitt issued orders to his commanders in North America.*

The 'signal' parameter deals with the way in which the reporter indicates that it is a language report. We can contemplate the real relations between the structurally dependent signal, the message and the nature of the reporting signal that fits the message into the surrounding context and relates it to the rest of the elements in the text. 'Separate dominant' places the message in a following subordinate reported clause, while the 'Separate equal' gives the same importance to the signal and message in the text as in the case of quotes. In the case of the 'Separate dominant', signals are adjuncts with the function of marking the dominant message. In the case of 'Fused clauses', there is no explicit piece of language functioning as reporting signal since it is encoded in the message itself. Attitude is the third element of reporting and it alludes to the reporter's evaluation regarding the message of the original language event conveyed. His/her attitude towards the language event may not be overtly expressed or an 'evaluative space' (Thompson and Ye, 1991: 369) can be used by the reporter if considered relevant. If the reporter does not want to evaluate the content of the language event, he/she can leave it open, which is common practice in journalism.

As Martin (1999: 144) states, it is worth mentioning that SFL does not pay much attention to the semantics of evaluating text processes. Martin notes that SFL does not take into account 'how the interlocutors are feeling, the judgments they make, and the value they place on the various phenomena of their experience' (1999: 144). This leaves aside what Caldas-Coluthard considers to be the essential reason of representing speech in newspaper discourse: 'significance'. In her opinion, the reporter only reports those parts of the exchange that are

significant from his/her point of view. The exchanges are reduced to significant 'utterances' that tend to exclude the interpersonal and social features necessary to prime the informational content, the ideational meaning (Caldas-Coulthard 1994: 298-299). Conversely, as Pounds (2010: 114) states, attribution is not only relevant in terms of mediation of attitudinal and subjective expression, but also in terms of the level of mediation (direct/indirect reported speech) and the identity of the individuals reported/quoted. Attribution can be seen as a form of impersonalization when professionals are quoted or reported. However, when this phenomenon applies to the reporting of speech by non-expert participants, there seems to be a subjective element even though the source of that personal dimension is the reporting voice and not the expert's. The choice of the 'reporting process' is thus relevant and can be 'neutral' ("no indication of authorial position or authorial distancing from the attributed material") as in verbs such as *say, state, declare, note;* or 'infused' ("expressing authorial evaluation towards the events reported or the reporters") as in verbs such as *admit, confess* or *reveal* (Pounds 2010: 115). The reports presented in newspapers have the legitimizing force of the publication behind them, and therefore, as readers, we normally take for granted that what is being exposed is based on facts. The term 'facts', as opposed to 'fiction', refers to the events in the world that do not require verbalization, but can be verbally expressed through averral and stand in direct 'correspondence' to the real state of affairs in the world (Sinclair 1986: 49). Yet the essential difference between the two terms lies in the fact that in factual reports, averrals made by the writer depend on words produced elsewhere, what constitute a double, mutually dependent averral (Caldas-Coulthard 1994: 302). Reporting facts hence entails the presence of the author's position conveyed through averral, because "when

[authors] report or write about another speaker speaking, they can only choose to aver that the speaker said something and not to aver the factuality of what has been said, that is to say, they detach themselves from the propositions" (Caldas-Coulthard 1994: 299).
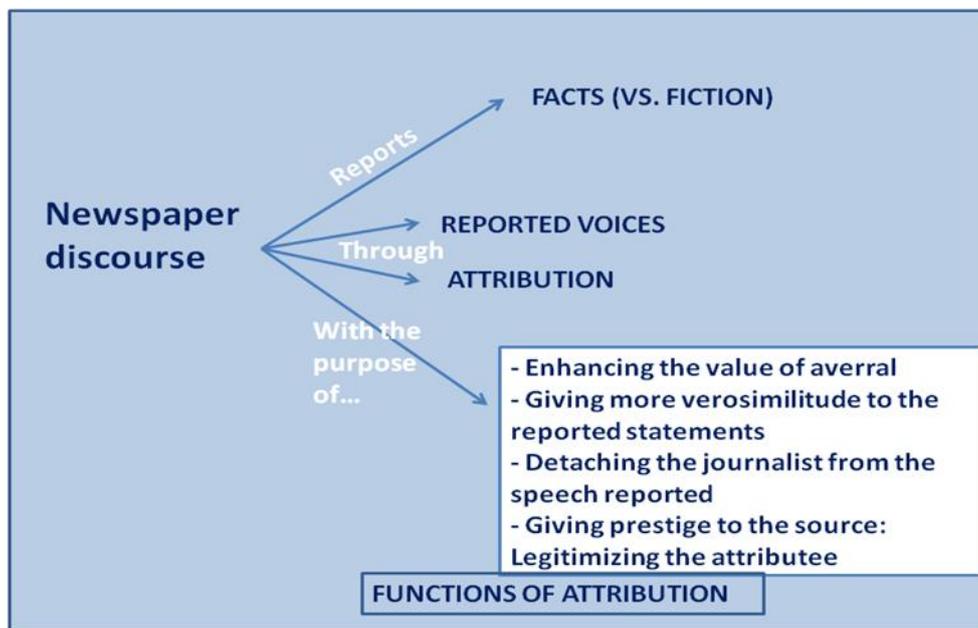
Figure 12. Reporting process in newspaper discourse (Adapted from Caldas-Coulthard 1994).

As previously seen, a core element in reporting processes is the verb used in the projecting clause for both quotes and reports. Caldas-Coulthard (1994: 306) classifies these 'verbs of saying' in a taxonomy according to their function in relation to the reported clause. Verbs are classified in three different categories: 'Speech-reporting verbs', 'Descriptive verbs', and 'Transcript verbs'. The first of these categories concentrates on verbs that introduce the report; while the second category presented by Caldas-Coulthard (1994) refers to descriptive verbs, which

mark the manner and attitude of the writer/speaker in relation to what is being said. The third category in this taxonomy is that of discourse-signaling verbs, which, as Caldas-Coulthard (1994: 306) states, do not normally report speech, but often accompany reporting structures.

Taking into account Caldas-Coulthard's taxonomy, we will focus on speech-reporting verbs (see Table 7 below) that can be further sub-divided into three categories. Neutral structuring verbs, also termed 'glossing', introduce a report but do not evaluate it. The intended meaning has to be extracted from the message itself. Conversely, 'Asertive,' 'Directive' and 'Expressive' metapropositional verbs make the writer/speaker's illocutionary force explicit by marking the stance towards the reported message. Finally, metalinguistic verbs categorize the type of linguistic event that is being reported in addition to making the speaker's illocutionary force explicit.

Speech-reporting verbs	Neutral structuring		say, tell, ask, enquire, reply, answer
	Metapropositional	Assertive	remark, explain, agree, assent, accept, correct, counter
		Directive	urge, instruct, order
		Expressive	accuse, grumble, lament, confess, complain, swear
	Metalinguistic		narrate, quote, recount

Table 7. Speech-reporting verbs in newspaper discourse (Caldas-Coulthard 1994).

Although the taxonomy put forward by Caldas-Coulthard (1994) is one the most comprehensive classifications in the field, Halliday and Mathiessen's (2004) model on reporting verbs (see Table 8 below), adds a complementary dimension to the descriptive parameters of a report. As this study focuses on reporting phenomena which occur in verbal clauses, we will briefly sketch their proposal on verbs functioning as 'process' in quotes. Process types are part of the 'transitivity' system through which the world and human experience are reconstructed as text. Each process relates to a specific domain of experience to construct meaning. Halliday and Mathiessen (2004: 172) identify the following six different process types: material, behavioral, mental, verbal, relational and existential.

		Proposition
General member		*say*
Verbs specific to speech function	a) Giving	a) Statements: *tell* (+ receiver), *remark, observe, point out, report, announce*
	b) Demanding	b) Questions: *ask, demand, enquire, query*
Verbs with additional circumstantial feature	a)	*Reply* ('say in response'), *explain* ('say in explanation'), *protest* ('say with reservation'), *continue* ('go on saying'), *add* ('say in addition'), *interrupt* ('say out of turn'), *warn* ('say: undesirable consequences')
	b) Manner specifying connotation	*Insist* ('say emphatically'), *complain* ('say irritably'), *cry, shout* ('say loudly'), *boast* ('say proudly'), *murmur* ('say *sotto voce*'), *stammer* ('say with embarrassment').

Table 8. Verbs functioning as process in quotes (Halliday and Mathiessen, 2004: 448).

In addition, within newspaper discourse analysis, Calsamiglia and López Ferrero (2003) present a classification of reporting verbs developed according to the results obtained by analysing their newspaper corpus. This classification is based upon the meaning each verb conveys in the text and distinguishes four different categories of reporting verbs: 'positive' (e.g. *assure*), 'popularizing acts' (e.g. *explain*), 'neutral' (e.g. *say, state*) and 'tentative' (e.g. *believe*).

From a functional perspective, reporting not only has the dimension of structural dependencies on the different elements of the report, but also the speaker's intentionality, discussed earlier in this chapter. The author's intentionality can be expressed by choices concerning the voices accessed in the text as well as the position where the explicit reporting signal is placed within the clause structure. The effects of placing the reporting signal before or after the message have not yet been considered in detail, but Thompson (1996b: 518) has pointed out that placing the reporting clause after the reported one is a marked option and a thematic choice at the same time, as seen in Examples 31 (rhematic attribution) and 32 (thematic attribution):

(31) *"The time has come to build on their experience and create a new generation of multilateral, WHO-aligned laboratories as a front-line defence against future pandemics," they write in Nature today.* (Sci_TG06381)

(32) *Professor Harry Lambright, a space policy expert at New York's Syracuse University, said: "It reaches the average citizen in the way a lot of space projects, including the space station, do not."* (Sci_TG0668)

The author defines how the report will fit the surrounding textual environment and chooses to emphasize that the projected clause is not overtly

_____

holding a hypotactical relation to the reporting clause. Here Thompson (1996b: 519) adds that "the reasons are broadly the same as, or very similar to, those which determine any choices concerning textual and ideational relationships between elements in a text"[2].

3.3.2.2. Epistemic Stance adverbials

Reporting in news discourse is not exclusively carried out through reporting clauses and reporting verbs as previously stated. There are different procedures that allow the author of the texts to encode the voice of a source other than him/herself. Epistemic stance adverbials are commonly found in newspaper discourse as explicit signals of the writer's detachment from the propositions expressed. As Conrad and Biber (1999: 64) state, epistemic stance adverbials are much more common than the other two semantic classes of stance adverbials (attitudinal stance and style stance) in conversation, academic and newspaper registers. As Thompson asserts (1994: 22), adverbs used as reporting adjuncts can be used to indicate that what you are saying is based on what someone else has told you rather than on first-hand information. It therefore signals a kind of report, but without identifying the original speaker or writer.

The variety of stance adverbials used to report other's voices is not very ample, just as studies on the subject are not yet fully developed. However, Conrad

_____

[2]   For further information on relationships established between elements in a text see Martin, J.R. (1992) *English text. System and Structure.* Philadelphia. John Benjamins. Pgs. 159-270.

and Biber have interestingly noted that the news reportage shows a consistent use

of certain stance adverbial to identify the source of information:

> […] the news reportage is the only register to make relatively common use of a stance adverbial marking the source of information: according to. This corresponds to the emphasis in news reportage on identifying the source of information. *According to* is used with sources that range from specifically named people and publications to sources identified only by their location. (Conrad and Biber 1999: 67) [Emphasis original]

Thompson (1996a) adds that stance adverbials constitute a device through

which logical coherence is achieved in texts making them a relevant way to signal

the type of connection being established between propositions. In this same

respect, Thomson and Zou (1999: 122) put forward the notion of 'evaluative

coherence', alluding to the writer's intention to convey a personal evaluation on

the topic of the text. Evaluative coherence is normally achieved through the use of

disjuncts (*unfortunately, obviously*), which have been defined as: "expressing the

writer's comment on the content or style of the sentence in which they appear"

(Thomson and Zou 1999: 123). They signal the absence of conjunction, which

focuses on local meaning, relations of time, comparison, cause and addition

realized by paratactic, hypotactic and cohesive conjunctions (Martin 1992: 27)[3].

In contrast to conjuncts, adjuncts, such as 'according to', have a reduced scope,

but add information and affect the sentence in which they appear, and have a

dependence on it.

'Comment adjuncts', a sub-category of modal adjuncts, are present in parts

of the text where they are relevant both for textual organization and transmission

of meaning. This is why Zou (1991) has referred to them as 'conjuncts with

_____

[3]   For a complete list of conjuncts see Halliday, M.A.K. and R. Hasan. 1976. *Cohesion in English*. London: Longman.

attitude'. Conversely, Biber, Conrad and Finegan describe 'according to' as an epistemic stance adverbial which signals the source of information:

> "the use of *according to* + NP in news is noteworthy, since it reflects the singular emphasis in news on stance adverbials that show the source of knowledge. *According to* is used with sources that range from specifically named people and publications to sources identified only by their location: "But drinkers will get off reasonably lightly, **according to reliable Westminster sources**" (Conrad, Biber and Finegan 1999: 871) [Emphasis original].

Hoey and O'Donnell (2008) have tried to move a step beyond simply describing the degree of presence of certain elements within a text type by relating frequencies of appearance with 'textual colligation' (Hoey, 2005: 115). They carried out an analysis on the use of words and group of words in a corpus which included 'according to'. They analysed the degree of 'keyness' (how outstanding these words are) (Scott, 1997) of 'according to' in a corpus of news from the British newspaper *The Guardian* and confirmed that this adverbial is prone to be present in newspaper discourse in text-initial sentences of the corpus.

In this chapter we have considered evaluation and its different forms, classifications and locations in order to specifically focus on the phenomenon of attribution, mainly accomplished through reporting and stance adverbials in newspaper discourse. Contrary to authors such as Piazza (2009: 170) who believe that attribution threatens the fidelity of the main text, this study concludes that attribution is not only a necessary process, but also an enriching device that widens journalist's perspectives on the information conveyed and provides the reader with a multiplicity of voices that have essential functions in the transmission of information. These elements will set the groundwork for the

analysis of attribution patterns present in a corpus of science popularization articles that we will carry out later in this study.

Among the models reviewed in Chapter 3, it is relevant to highlight the role of attribution within engagement system, and, more specifically, within epistemological positioning of the author in texts. As we have seen, attribution provides a source that identifies a voice out-of-the text as author of the information conveyed, a process that establishes either hypotactical or paratactical logico-semantic relations between two clauses in the text. Relations established between clauses in a report are not only relevant to classify them in relation to projection mode (quotes or reports), but also to provide the reader with more data that can be drawn by analysing the order of information in a text. This will be the basis for Chapter 4, where we deal with the organization of information in texts, and specifically with how the event reported by speakers/writers is primed to be located in specific parts of the text.

Chapter 4

Information expliciting and thematic progression

Chapter 4

Information disclosing: Theme and Rheme in discourse

4.1.- Introduction

4.2.- Thematic and information structures in systemic linguistics

    4.2.1.- Theme and rheme: clause as message

        4.2.1.1. – Side constructions of thematic structures: Equatives, predicated theme, thematized comment, passivization

        4.2.1.2. – Theme delimitation

        4.2.1.3. – Multiple themes

    4.2.2.- Given and new: clause as information

4.3.- Thematic progression and informativity

4.4.- Different perspectives on the study of theme

4.5.- Theme in reporting clauses

*The structure of language reflects in some way
the structure of experience, that is to say, the
structure of the world, including [...] the
perspective imposed on the world by the speaker*
Croft 1990:164

4.1. Introduction

The election of what comes first in a sentence can appear as something

casual that speakers and writers choose without purpose. However, this is just an

impression, and as elements in the initial position have an important role in the

clause especially in guiding the way information reaches the audience. As we

have seen in previous chapters, newspaper discourse and, more specifically,

science popularization discourse, is aimed at a lay audience that relies on the

_____

information that the journalist conveys in the article. Furthermore, this information is strongly supported by the voices of scientists, who validate and strengthen the scientific nature of the facts revealed. Scientists' voices are inserted in the journalist´s speech and contain the information on scientific findings that the journalist has chosen to report or quote. The way they are presented is also relevant, as the scientist being introduced first can have a special connotation, and a different nuance can be perceived by the reader if the information is presented prior to the source of information. This is one of the premises supporting the analysis being carried out later in this dissertation when we take a deeper look at how elements are thematized, and recurrently thematized, in the discourse of science popularizations from the British newspaper *The Guardian*.

When we look at language from the textual metafunction (see Chapter 3), we are dealing with the way in which speakers construct their messages so that they match effortlessly with the unfolding language event (Thompson 1996a: 141). It is relevant to bear in mind at this stage of the study that authors not only interact with their readers or speakers and represent human experiences through texts, but also organize the way the message is expressed, and how the information reaches its receiver. As we will see in this chapter, this is a fact that highly conditions the actual meaning of clauses. Moreover, as with thematized elements, the textual metafunction can help linguists to identify a given discourse, by consistently giving the elements a specific value. Speakers and writers thus turn into mediators between the meaning of words and the receivers by deciding how the message is structured to signal its relevance and the relationship established with other elements in the text. As Thompson asserts (1996a: 143),

one of the choices that most directly modifies meaning is that of theme. The fundamental aspect of theme is that its meaning depends on its existence as the result of a speaker/writer´s choice of which experiential element should form the point of departure for the message (Thompson and Thompson 2009: 58). We will now deal with this notion in its several aspects and contexts, particularly focusing on the behaviour of theme in reporting clauses.

4.2. Thematic and Information Structures in Systemic linguistics

Within the textual metafunction, we will discuss two complementing systems: Theme System and Information System, which assign two different pairs of elements, Theme-Rheme and Given-New, to the structural constructs, in order to describe the characteristics of meaning of utterances in context (Fries 2009: 12). As Fries (1983: 117) describes, there have traditionally been two approaches to the conception of the aforementioned two systems. One approach is called the 'combing approach', and equates theme with given. The second, called the 'separating approach', disentangles the two concepts, and separates information structure from thematic structure. This 'separating approach' presented by Halliday and supported by systemic schools in general, argues that, although theme typically coincides with given, and new tends to fall within the rheme, the opposite is also common in spoken English (Downing 2001: 26). As Fries states, "the information that is contained within the themes of the various sentences of a

_____

passage correlates with the method of development of the passage" (1983: 16), a statement which declared him as a follower of the 'separating approach'.

4.2.1. Theme and rheme: clause as message

Language is, in essence, communication, which has been conceived as a dynamic and cumulative process in which the speaker/writer and reader/hearer have a communicative need to meet in an area of convergence (Ndahiro 1998: 17). As Halliday and Mathiessen (2004:64) assert, as a result of this communicative need,

> we may assume that in all languages the clause has the character of a message: it has some form of organization whereby it fits in with, and contributes to, the flow of discourse. But there are many ways in which this may be achieved. In English, as in many other languages, the clause is organized as a message by having a different status assigned to the one part of it. One part of the clause is enunciated as the theme; this then combines with remainder so that the two parts together constitute a message. (Halliday and Mathiessen 2004: 64)

From their point of view, the fact that messages are organized as clauses lies on the different status given to some of its members. One part of the message is articulated as theme, while the rest of the clause is conflated with rheme so that both parts create a unity: the message. In this same respect, and as Thompson states, thematization refers to the way individual elements of the clause are included into the flow of texts, as well as a matter of co-operation with the reader. The reader perceives a clear starting point for the clause which makes it easier for him/her to integrate the meaning of the clause into earlier meanings, and helps "to

see immediately how the information that will come in the reminder of the clause is likely to fit in what has already been said" (1996a: 142). The signaling of relevance depends upon factors such as the language being written or spoken. Obviously, spoken language can rely on intonation patterns and written language cannot. However, in written language an item has thematic status by coming first. We will now concentrate on the characteristics and behavior of theme in written texts.

Going back to the eighties, Fries (1983: 117) referred to Mathesius' point of view as the 'combining' approach, which considers that:

> the thematic content or topic of a sentence is determined to a great degree by how the information expressed in that sentence relates to the information already available in the linguistic and non-verbal contexts. That is, the theme of a sentence must be known information and one can only tell what is or is not known information in a particular sentence by looking at how the information in that sentence relates to the information available in the context (Mathesius 1939 in Fries 1983: 116-117).

On the other hand, ascribing to the so-called 'separating' approach, we find the systemic linguistics point of view, which advocates the separation of different notions of information. First, they identify the information which is known, or obvious in the situation as 'given', as opposed to the notion of theme, considered as the information from which the speaker or writer proceeds in his/her discourse (Fries 1983: 117).

In the early 20th Century, the Prague School of Linguistics and authors such as Firbas (1964, 1986) or Mathesius (1939) developed their structuralist analysis, and came up with the label of theme to refer to "that which is known or at least obvious in a given situation, and from which the speaker proceeds in his

_____

discourse." Since then, many attempts have been made to achieve a consensus on the definition of theme. Some time later in the 20th Century, Halliday defined theme as "the element which serves as a point of departure for the message; it is that with which the clause is concerned" and as "the starting point for the message, it is what the clause is going to be about." He defined its counterpart, rheme, as "the reminder" of the clause (1985: 38-39). Theme is not only a formal label for the first element of the clause, but also a functional tag for an element which is set by the writer as a point of departure of the message; as a clause constituent which locates and orients the sentence within its context. In Halliday and Mathiessen's words "it is what sets the scene for the clause itself and positions in relation to the unfolding text" (2004: 66). We could say that theme is the leader of the message within the clausal elements, while rheme is found in the rest of the message conveyed in the clause, and develops the meaning which has been set in the theme. Clauses can be understood as a binary structure combining a theme and a rheme, as it can be seen in Figure 13 below.
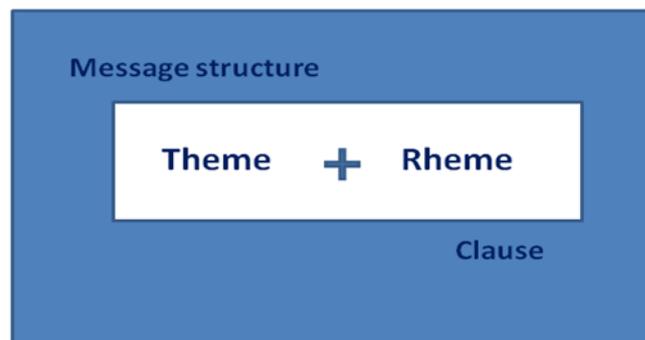
Figure 13. Message structure of a clause, based on Halliday (1985)

Many other authors have considered different variations on the concept of theme put forward by Halliday. For example, Huddleston (1991: 95) proposes that theme be separated and considered as expression or content. Berry considers that "theme has to do with the concerns of the speaker or writer, and that it is the cumulative force of the theme of the clauses of a text which indicates this concerns, rather than the theme of any clause individually", and suggests differentiating between theme as meaning and theme as form (Berry 1996: 13 quoted in Whangeng 1997: 44).

It is also relevant to note the contribution made by Firbas (1982) when dealing with the notion of theme. Apart from the verbal context of the clause in which theme is immersed, he stresses the presence of a situational context constituted by objects which are of immediate concern to producer and receiver of the message and by some other objects which are, irrespective of the time or place, always present. Both contexts, verbal and situational are, in Firbas' words:

> Embedded within the entire preceding context, equally consisting of a verbal part and a situational part accompanying it. In its turn, this sphere is embedded in a still wider sphere of common knowledge and experience shared by the producer and the receiver of the message, eventually, the entire contextual complex so far described within the wider context of human knowledge and experience. (Firbas 1992: 171)

As thematic structure is inherently thematic (Fries 2009: 12) together with the priming of theme for the initial part of (or beginning of) the sentence, it is easy to think of it as "what the clause is about" (Halliday 1985: 39). However, this concept is misleading because theme can easily be confused with subject. In this study, we will not adopt the idea that theme is what the clause is about but rather support the notion that words primed to appear in thematic position help readers

_____

to better and more quickly grasp what the text is about ('aboutness') (Wangheng 1997: 47) and its 'keyness' (items which show an unusual frequency in a text are identified and considered "key" words for the understanding of the text) (Scott & Tribble 2006: 57). As Wangheng asserts, there is a strong relationship between theme and 'keywords' in a text, as they can serve to highlight the uniqueness of a particular text in relation to others, or with the language in general (Wangheng 1997: 47). Keywords, which have either been defined as "words or phrases which in some sense characterize, represent typify or identify a text" (Collins and Scott, 1997: 3, quoted in Wangheng, 1997: 46) or as "words which occur with unusual frequency in a given text" (Scott 1997: 236), tend to concentrate in the initial section of clauses. Therefore, these keywords coincide with thematic position and work together to guide the reader through the meaning of the text. Conversely, Davies (1988, quoted in Gosden 1993: 60), proposes conceiving theme in a dual schema. On the one hand, we find an obligatory topic, realized by the grammatical subject (GS), and, on the other, an optional 'Contextual Frame' realized by elements which occur preceding the GS.

As it has been previously expounded, there is not a universal consensus on the definition of and considerations surrounding the concept of theme, nor have linguists come to share a common point of view on its identification. Identifying theme is not as straightforward of a procedure as it may seem and strongly depends on the mood of the clause it initiates. In the case of declarative sentences, the identification of theme is clearer, and, in the vast majority of cases, it coincides with the subject of the clause; a phenomenon which has been named 'conflation' (Thompson 1996a: 144). The fact that subject and theme are encoded

_____

in the same element is a common case of 'conflated theme' where the writer has chosen not to alter the natural tendency present for both theme and subject (frequently in the case of subject and consistently in the case of theme), as the first element of the clause. Unless there are reasons to alter this conflation, this is the normal election of theme, and therefore it is said to be the 'unmarked' option. This is in contrast to the 'marked' choice in which the writer voluntarily opts for fronting an element different from the subject. The element which is placed in the thematic position in this marked option can be any constituent, but it is common to find adverbial groups or prepositional phrases respectively functioning as adjuncts and complements in that position:

(30) *Little Bo-peep has lost her sheep* (Unmarked theme, conflation)

(31) *This responsibility I accept wholly* (Marked theme, Complement)

(32) *Without any warning, the computer shut down* (Marked theme, Adjunct) (Halliday and Mathiessen 2004: 73)

Complements are reasonably flexible constituents, but adjuncts cannot be moved as easily into theme position, and therefore constitute a marked point in the cline of markedness of thematic option in clauses (Thompson 1996a: 146). This can be seen in Figure 14 below, although they are considered as marked by Halliday and Mathiessen (2004: 73).

_____
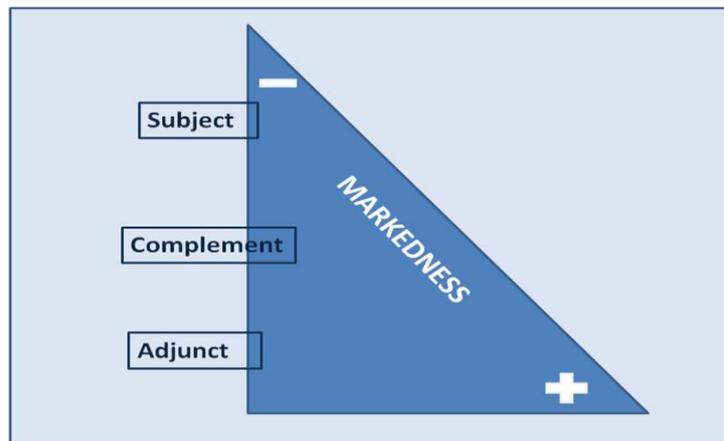
Figure 14. Cline of markedness in thematic options

In the case of non-declarative sentences, and more specifically, interrogative sentences, it is common in thematic position to find 'what one wants to know' (Halliday and Mathiessen 2004: 75) in thematic position. In the case of *wh-* interrogatives, *wh-* elements (*who, when, how, where, why…*) prevail in initial position, mainly due to the fact that they stand for the information that the hearer requires to be present. This is a basic characteristic of interrogative sentences (Thompson 1996a: 146). Conversely, in the case of yes/no questions, the hearer is not being asked for a complete response which fulfills the missing information, but only the confirmation of one of the two options of polarity: yes/no.

As Thompson (1996a: 146) asserts, in the case of imperative clauses, the marked option can be understood on account of the nature of the construction, which naturally emphasizes the role of the predicator as the most important element in the performing of the action by another person. The essential message in an imperative sentence is that of impelling the hearer to do some kind of action,

_____

but the option of placing the verb in the initial thematic position is not the only one. As Halliday and Mathiessen indicate (2004: 76), there is also a less marked option which not only conveys the meaning of "I want you to do something" but also that of "I want both you and me to do something" through the use of *let's,* as seen in Example 33.

    (33) *Let's* (theme)*/ go home now* (rheme)

As we have seen, it is the choice of theme that determines the content of the message (Thomas and Hawes 1997: 20) but, as Halliday points out, we have to bear in mind that the concept of 'subject' is not as straightforward as it seems and, as Halliday and Mathiessen (2004: 56-58) explain, there are three types of subjects which correspond to three distinct and separate functions. 'Psychological Subject' refers to the concern of the message, 'Grammatical Subject' refers to that of which something is being predicated, and 'Logical Subject' alludes to the doer of the action. In order to explain the functions accomplished by these three types of subjects, new labels were put forward by Halliday and Mathiessen (2004). Psychological Subject was re-named 'Theme', Grammatical Subject was re-named 'Subject', and logical Subject was termed 'Actor' (Halliday and Mathiessen 2004: 57):

    (34) *This teapot my aunt was given by the duke* (Halliday and Mathiessen 2004: 56)

This teapot -> Psychological subject

My aunt -> Grammatical subject

The duke -> Logical subject

The fact that subject and theme normally occur in a conflated way in declarative sentences can be a source of conflict or confusion between the two notions. As Halliday mentions, this is because "the typical, UNMARKED form, in an English declarative […] clause, is the one in which theme, Subject and Actor are conflated into a single element" (1985: 36) [Emphasis original].

It is worth mentioning at this stage that Halliday, in his consideration of theme and subject, points out that subject is a formal, grammatical category which determines other grammatical features, as well as "that of which something is being predicated" (1985: 33). This notion coincides with his concept of Actor as well as what nineteenth century linguists named Grammatical Subject (Thomas and Hawes 1997: 19). The notion of theme is thus reserved for what has been traditionally called Psychological Subject, or "that which is the concern of the message, what the speaker had in his mind to start with, when embarking on the production of the clause" (Thomas and Hawes 1997: 33). One of the most important advances led by Halliday is that of connecting the notions of 'psychological subject' and 'logical subject' (Theme and Actor) in what we now call theme, a concept which deserves further definition (Fries 2009: 16).

Taking into account Halliday's reflections, in the case of declarative sentences, we can perceive that a triple conflation can occur. as it has already been noted in this study, a conflation between subject and theme can exist, but this duet can also be conflated with Halliday's notion of Actor, as can be seen in Figure 15 below.
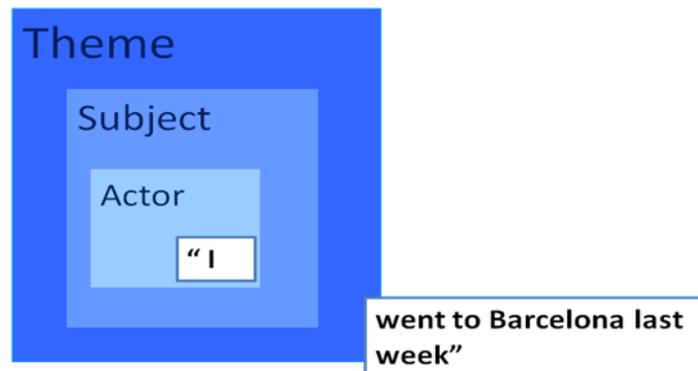
Figure 15. Conflation of theme, subject and actor in initial position of declarative

sentences

In the case of interrogative clauses, the subject is conflated with the actor, but not with the theme, which is usually a *wh-* element in the case of *wh-* clauses, or an auxiliary verb in the case of yes-no questions. A diagram on this can be seen in Figure 16 below.
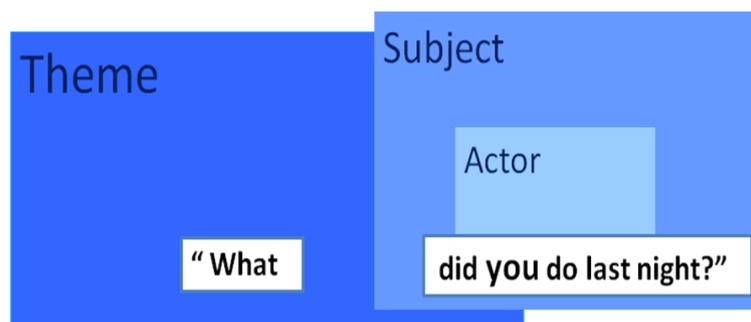
Figure 16. Conflation of subject and actor following theme in initial position of
declarative sentences

Finally, in the case of imperatives, there is a conflation between the theme and the actor of the clause, but not with the subject which is normally elliptical in these type of clauses (see Figure 17 below).

_____
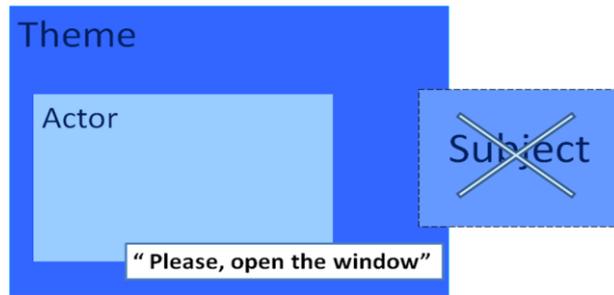
Figure 17. Conflation of theme and actor

4.2.1.1. Side constructions of thematic structures: Equatives, predicated theme, thematized comment, passivization

In addition to the fact that different representations of theme depend on the mood of the clause they initiate, there are some special thematic structures that speakers and writers use to establish a specific point of departure. One of these special constructions is that of 'thematic equatives', also referred to as 'pseudo-cleft,' which joins together a group of elements of a message, creating a single constituent which is fronted to thematic position (Halliday and Mathiessen 2004: 69, Thompson 1996a: 149). The expression 'thematic equative' responds to the parallelism that exists between the multi-element constituent, which occupies thematic position, and the rheme of the clause, which are equivalents expressed in the form of an equation. As seen in the examples below, the relation of equivalence is expressed through the presence of the predicator 'be' in the identifying clause (Halliday and Mathiessen 2004: 69):

(35) *What the duke did with that teapot was give it to my aunt*

(36) *What the duke gave to my aunt was a teapot*

(37) *The one who gave my aunt that teapot was the duke*

A second special construction that we will consider in this study is that of 'predicated theme', a thematizing structure which has also been named 'cleft sentence', and shares with thematic equatives, the characteristic of gathering several elements from the message in a single constituent. As Thompson expresses, the main function of the predicated theme is to emphasize the noteworthiness of the thematized constituent in a way, which often contrasts with other constituents in the clause or has been selected as the appropriate one among a series of alternatives (1996a: 152). Examples on predicated theme can be found below:

(38) *It's not the technology* (theme)/ *which is wrong* (rheme)

(39) *It is the second of these points* (theme)/ *that I shall be concentrating on in this talk* (rheme)

(40) *It wasn't until 1986* (theme)/ *that we finally came back to work in the UK* (rheme)

(Thompson 1996a: 151)

The third of the special constructions which we would like to emphasize in this section is that of 'thematized comment', a structure which enables writers or speakers to activate the message they want to convey by including a personal comment on the value or validity of what they are about to say. 'It' serves to establish the beginning of the message by introducing the comment of the author, a comment which allocates the attitude of the author before the beginning of the

actual message. As previously discussed in Chapter 3, the use of thematized comment in different types of discourse is closely related to the notions of appraisal and modality, which as Thompson (1996a: 152) observes, are modes of expression. As he also notes, the comment included in the 'it'-clause is not a meaningful component of the second clause. This is because it is not possible to rewrite it in the form of a single clause, as it is with predicated theme:

(41) *It is true* (theme)/ *that it took five years to do so* (rheme)

(42) *It is difficult* (theme)/ *to know exactly how to characterize what we have just noticed* (rheme)

(43) *It's interesting* (theme)/ *that you should say that* (rheme)

(Thompson 1996a: 152)

Examples of the special construction of thematized comment shown by Thompson include the whole first clause as theme, whereas authors such as Halliday and Mathiessen (2004: 97) have suggested that this structure is not a thematizing one, and its theme consists of only 'It'. In this study, we take Thompson's viewpoint and consider this structure as a thematizing device that takes place at key transition points of the text (1996a: 153).

A fourth thematizing structure which is predominantly present in spoken discourse is 'preposed theme'. Through the use of this device, speakers mark a strong separation between theme, which is constructed in a separate constituent, and rheme, where the theme is recalled through a pronoun. Examples of this structure can be found below:

(44) *Happiness,* (theme)*/ that's what life it's about* (rheme)

(45) *Your Mum*, (theme)*/ does she know you're here?* (rheme)

(Thompson 1996a: 153)

Finally, it is relevant to comment that a common device used to explicitly mark the theme in clauses is 'passivization'. The process of turning active sentences into passive ones entails the fronting of the 'sufferer' of the action, and normally requires the presence of a 'doer' introduced by the preposition 'by'. For example: *They were rescued by a soldier who spotted them both crying* (Thompson 1996a: 154).

4.2.1.2. Theme delimitation

Now that the notions of theme and rheme have been outlined, a second issue that must be considered in this study is theme delimitation. How far does theme extend in a sentence? We have already noted that there are discrepancies between authors, who analyze theme in different ways. As a general notion, Halliday and Mathiessen point out that the theme of a clause is the first group or phrases that have some function in the experiential structure of the clause (2004: 66). Thompson reinforces this idea stating that it is compulsory for theme to contain an element that plays a role in transitivity. In other words, every theme in a clause must contain a participant, a process or a circumstance (1996a: 159). This experiential element which initiates the meaning of the clause has been called by

_____

different names, depending on the author dealing with the notion. Halliday and Mathiessen have referred to it as 'topical theme' (Thompson 1996a: 79), but there are alternative considerations such as Thompson's which asserts that labeling a theme according to what is considered to be 'topic' in topic comment analysis can be misleading, and advocates the term 'experiential theme' which avoids the comment analysis context.

In this same respect, Ndahiro (1998: 101) states that there is a shared ground by the concepts of theme, subject and topic, three of the elements that are usually identified at the beginning of clauses and can create confusion in their interpretation. He defines subject according to a formal criterion, and identifies it with a particle that is not affected by the sequencing of the clause. Conversely, he describes theme in physical terms as containing everything up to the verb in the clause, and defines topic as having a discourse-semantic role of 'aboutness' within the clause. A diagram representing the three elements and their convergence area can be seen in Figure 18 below.
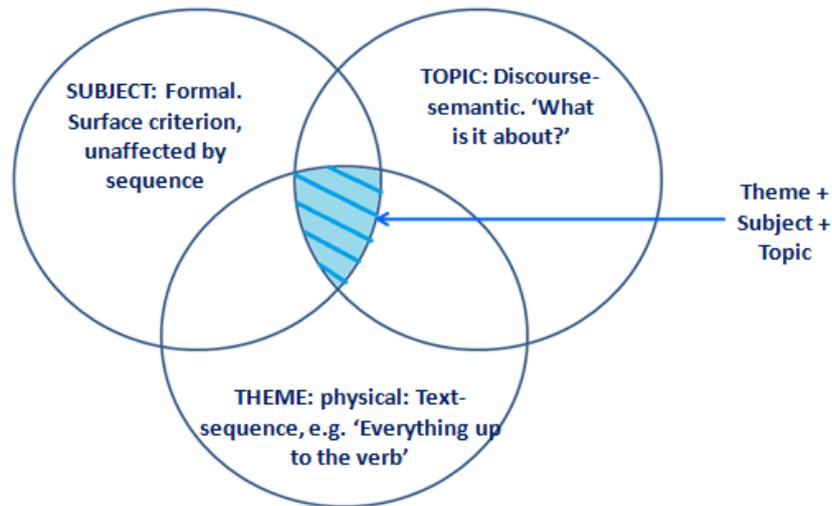
Figure 18. Subject, Topic and Theme convergence in initial position of clauses
(Ndahiro 1998)

An additional issue within this same topic is the inclusion or exclusion of verbal forms in theme. Wangheng (1997) considers that the main verb normally contains new information in the clause, and is thus more appropriately regarded as rheme (Wangheng, 1997: 50), but there are differing views on this issue. Based on Halliday's work on theme, Berry (1996) states that theme should not be defined by only the first ideational element in a clause. If only the first ideational element is analysed as theme, then an analysis of theme in text will not capture some of the co-referential elements. She claims that "the priority concerns, discoursal or causal, of a speaker or writer need not be ideational" (Berry, 1996: 19), and that the writer may choose to select a feature as theme because it relates to the surrounding text, or the concerns of the immediate clause, or something more closely related to the reader's concerns (Forey 2002: 54). Berry (1987, 1989) also asserts that it is not necessary to stress the relevance of subject over theme or the

other way round, because both of them are important for the notion of topic and the development of a text (Thomas and Hawes, 1997: 34).

In addition, Chafe (1976 as quoted in Thomas and Hawes 1997: 34-35) adds that subject is the item to which knowledge is attached, making it a more effective prompt for a passage about other elements that are not present in the initial position of the clause. Halliday and Mathiessen (2004: 66), on the other hand, consider that theme is only the first experiential element in the clause. Theme is generally unmarked, but, in the case of a marked theme, the following subject would also be included in theme structure (2004: 103).

We have already considered the various definitions and forms that theme can adopt, but nothing has been said so far about the constituents that form theme. Besides the constructions that have been reviewed in this chapter, theme can also be accompanied by elements that indicate that the clause belongs to a larger unit, the 'clause complex'. This refers to a structure consisting of more than one clause (Thompson 1996a: 154), or how the clause relates to other element in the clause complex. Conjunctions are elements which can frequently be found in thematic position, constituting a departure point. As Halliday and Mathiessen remark, this departure point does not absorb the full thematic potential of the clause (2004: 83) but helps the reader set the clause complex in the appropriate context.

In addition to conjunctions, adjuncts can also occupy thematic position. Circumstantial ones contribute to the experiential meaning of the clause, while conjunctives work to signal how the clause adjusts to the preceding text: *therefore, also, likewise, to sum up, on the other hand* (Halliday and Mathiessen

_____

2004: 82). Modal adjuncts may be perceived as a comment on the meaning of the message, due to the fact that they convey the speaker's opinion about the truth value of the message (Thompson 1996a: 157): *probably, usually, occasionally, obviously, in my opinion, honestly, evidently, broadly speaking, to my surprise…* (Halliday and Mathiessen 2004: 82).

Alternatively, Thomas and Hawes (1997: 99) have considered other thematic classifications apart from those concerning actor, theme and subject. They have noted that discourse participants are often present in media discourse as themes, and have different manifestations depending on the degree of visibility given to the self. Firstly, the writer can choose to adopt a visible position and identify him/herself through the use of pronouns (*We, our…*) or nominal groups associated with the publication in which they appear (*The Times, The Sun…*). Secondly, pronoun use can be identified as the cohesive potential of theme and their role as anaphoric references to something which has been stated before. Thirdly, they have signalled named and semi-named participants, which depend on the degree of explicitness of the given reference. Named participants' cases are those in which the writer provides the reader with enough information to easily identify the individual being discussed; conversely, semi-named participants are those from which a certain distance is implied in human terms. As Thomas and Hawes assert (1997: 102), to speak about 'David Cameron,' is to speak about the person himself, about the specific man, but, to refer to 'The prime Minister' alludes to the office, therefore represents a less human semantic component. Fourthly, they identify institution participants, which do not refer to specific people, but to establishments, academies or well known societies which have

_____

become news or have characterized a communicative event through one of their members.

4.2.1.3. Multiple themes

As it has been previously seen, discourse can be strongly or weakly thematized depending on the presence of marked/unmarked elements in thematic position (Perfetti and Goldman 1975: 258). So far, we have only dealt with themes which were relatively simple to identify; these themes could be named 'simple' as opposed to 'multiple' themes, which are formed by an additional element which complements the experiential theme in the clause (Halliday and Mathiessen 2004: 79). Therefore, multiple themes can have textual, interpersonal and experiential elements simultaneously, but the experiential element is its only compulsory characteristic (Montemayor-Borsinger 2003: 40). This means that in multiple themes, the thematic component does not end with the first experiential element in the clause, but extends further to include constituents varying from continuatives to conjunctions, vocatives or adjuncts. Continuatives (*yes, no, well, oh, now*) signal a move in the discourse, an advance in the flow of information. Conjunctions (Paratactic: *and, or, nor, either, neither, but, yet, so, then, for.* Hypotactic: *when, while, before, after, until, because, if, although, unless, since, that, whether, to, by, with, despite, as, even if, in case, supposing (that), assuming (that), given that provided (that), so that, so as to, in order to, in the event that, in spite of the fact that*) act as links or binders to the next sentence, depending on the

paratactic or hypotactic relation that exist between clauses. Conjunctive or discourse adjuncts, on the other hand, relate the clause to the preceding text and fulfill a textual function, whereas modal comment adjuncts (*occasionally, obviously, probably, usually, personally, honestly*) make the speaker/writer´s opinion explicit on the content of the message, constructing modality in the same sense that finite verbal operators do in unmarked theme of yes/no interrogatives. Vocatives are the last type of non-topical element in thematic position, mainly present in multiple themes to address somebody, as it can be seen in the example below:

(46) *well /but/ then /surely /Jean /wouldn't /the best idea* (theme)*// be to join in*
(rheme)

A further division can be carried out on the six types of non-experiential element in thematic position put forward by Halliday and Mathiessen (2004: 81). This concerns the force with which they are related to thematic position. On one hand, they establish an initial category where elements are 'inherently thematic'. That is to say, if they are present in the clause, they are primed to occur in thematic position (Hoey 2005: 129) (see Table 9 below). Within this category, we can include continuatives and conjunctions. The second category consists in 'characteristically thematic' elements, which are not as strongly tied to thematic position, but merely have a tendency to occur within thematic structure. In this second category, we could place conjunctive and modal comment adjuncts, vocatives and finite verbal operators construing modality.

_____

Strength of tie to thematic position	Type of non-experiential element
INHERENTLY THEMATIC	Continuatives
	Conjunctions
CHARACTERISTICALLY THEMATIC	Conjunctive adjunct
	Modal comment adjunct
	Vocative
	Finite verbal operator

Table 9. Relation between thematic elements and their ties to theme position

Based on this taxonomy and the distinction made between marked and non-marked themes, we could say that different initial sections of the clause, in the case of marked themes, show a growing degree of markedness that increases in the following order: no non-experiential theme, inherently thematic non-experiential theme and characteristically thematic non-experiential theme. At this stage of the study, it is relevant to present the conclusions that Halliday and Mathiessen (2004: 85) draw on the dichotomy of marked and non-marked theme. These assertions are significant for the analysis of thematic elements in newspaper discourse that are carried out in this work and will be presented in subsequent chapters.

1.- Initial position in the English clause is meaningful in the construction of the clause as a message; it has a specific thematic function.

2.- Certain textual elements that rhetorically and logically position the clause within the discourse, are inherently thematic[1].

_____

[1]   Stated as well by Hoey 2005: 129

3.- Certain other textual and interpersonal elements, that set up a semantic relation with what precedes, or expresses the speaker's angle or intended listener, are characteristically thematic. This includes finite operators which signal one type of question.

4.- Inherently thematic elements lie outside of the experiential structure of the clause, and therefore, have no status as participant, circumstance or process.

5.- Until one of the latter elements appears, the clause lacks an anchorage in the realm of experience; this is what completes the thematic grounding of the message.

4.2.2. Given and New: clause as information

The system of theme described above bears a close relationship to that of information, in which speakers and writers arrange the elements of discourse flow in order to communicate a specific meaning. Grammar, as it has been stated by Halliday and Mathiessen (2004: 89), construes structural units up to the rank of the clause complex. However, there are also non-structural elements, which help building texts with the help of semantics, and are able to reach as far as a novel. Non-structural resources for managing discourse flow create a series of semantic links across sentences which establish cohesive relations in a text.

_____

The information unit reflects the different status of what is already known by the speaker, or at least predictable, and what is new or non predictable. It is precisely the interaction between new and given information that creates meaning in discourse. Thus, we can say that the two essential elements necessary to achieve a correct disclossure of meaning are 'given' and 'new' (Halliday and Mathiessen 2004: 89). Chafe (1976: 30 quoted in Vande Kopple 1991: 313) defines given information as "that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance", and Halliday (1985) adds that for a part of the clause to be considered 'given' it has to be recoverable, which is to say, that its meaning and the meaning of the structure it occurs in are deduced on the basis of the previous information, relevance of the situation in which it is used, or readers' previous knowledge.

One of the main issues when dealing with the relationship of given and new in discourse is that of markedness. As we have seen about theme units, some elements tend to occur in initial position, while others, less prone to that location, are purposefully fronted by writers or speakers to emphasize their relevance.

As it has already been noted in this chapter, discourse needs an anchorage point which sets the initiation of meaning and creates a context for the reader to tie in new information. This anchorage point, in the case of information units, is a given element with which readers are familiar, and has the tendency to be presented in 'phoric' position (Halliday and Mathiessen 2004: 89). Given elements are already present in the context shared by the reader or listener of the text, thus belonging to a background of shared knowledge. This knowledge can be

verbal or non-verbal, and to a great extent, enables the comprehension of new information.

The most typical distribution positions a given element followed by a new element and, as Halliday and Mathiessen present (2004: 89), the pattern is marked by two circumstances. Elements in new position are marked by prominence, whereas given typically precedes new. Prominence given to new elements is accomplished through intonation in spoken discourse, where fronted elements hold tonic prominence. It is different with the case of written discourse, where the only explicit marking is that of position in the clause. In both cases, the element carrying this prominence is labelled as having 'information focus'. There is a series of elements in language which are inherently primed to occur at a given position. Because in order to be interpreted to their full potential, they need a previous element. A clear example of this dependency can be found in new elements containing anaphoric or deictic elements. They directly point out the meaning of an element which has been previously stated in the clause, typically in the given element (Halliday and Mathiessen 2004: 91).

Below, Ndahiro (1998: 120) represents the diverse characteristics of given-new structure taking into account the position of the verb as the center of the clause and presenting the different roles of elements found in pre and post verbal position:

_____

Pre-verbal	Verbal	Post-verbal
Common ground. Initial state	Transition-goal. Intermediate state	Target state
Given Information	New information: achievement of communicative goal	New information: achievement of communicative goal

Table 10. Pre verbal, verbal and post-verbal roles in clauses (Ndahiro 1998: 120)

Information and theme systems are not only coexisting, but parallel clause organizations which share their unmarked character. Neither of them breaks its standard structure unless necessary and explicitly marked by the author. we would, however, like to emphasize an interesting feature that Halliday and Mathiessen point out (2004: 93). They state that while Theme+Rheme structure is speaker-writer oriented, 'Given+New' is more listener-reader oriented. Although both are selected by the author of the text, theme is voluntarily chosen by the writer as his/her point of departure for the unveiling of the message, whereas given gets activated so that the new information is understood, as it is what the reader has accessible in his/her background knowledge

_____
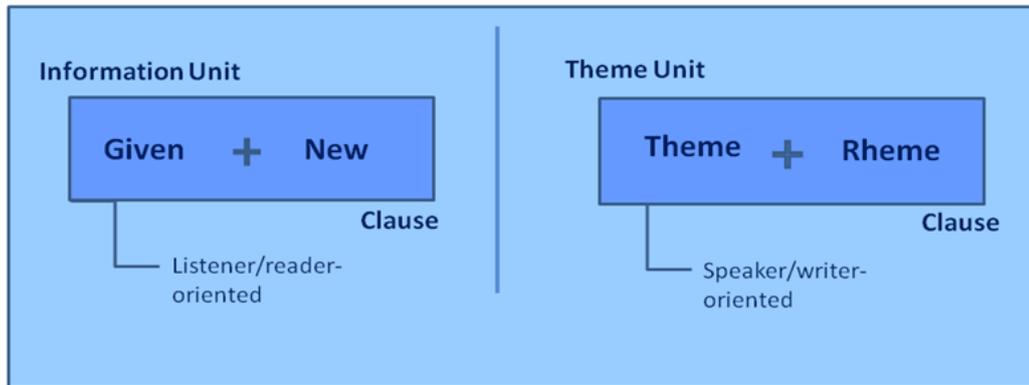
Figure 19. Information and Theme structures in English.

4.3. Thematic progression and informativity

If a text is to be appreciated as a text, it should have some kind of 'continuity', expressed through thematic organization, and be closely related to text connectivity and coherence. Theme choices are interesting by themselves, but they acquire an overall meaning when they are studied together along a text. They are understood as "any passage, spoken or written, of whatever length, that does form a unified whole" (Halliday and Hasan 1976: 1). Themes can also be taken as signals of the 'method of development' of the text (Fries 2009: 24) and its 'cohesion', the latter being understood as the property of language (Halliday and Hasan 1976: 2). The method of development of a text, taken as "the set of themes in a text" (Fries 2009: 23), is not yet a clearly defined concept, and has received criticism on being a simple phenomenon in most texts. The fact that texts have a simple method of development does not make them 'less good', but, as Fries

asserts (2009: 24), it is one of the potential textual meanings that still deserves more attention in the framework of meaning systems.

In this section of the chapter, we will consider a double approach to the study of thematic choices. On one hand, we will focus on the relations established among the themes present in a text, and how they create progression. On the other hand, we will concentrate on the different functions accomplished by themes when signaling the underlying coherence of the text (Thompson 1996a: 165). Thematic progression refers to the patterns created by the way themes and rhemes in sentences relate to each other. Fries (2009: 18) states that Thematic progression has the closest association to lexico-grammar. This is because it can be noted by observing the sequence of themes in successive clauses or sentences in a text, and examining how their cohesive connection to the information stated in previous sentences. In this same respect, Daneš (1970, 1974, quoted in Thomas and Hawes 1997: 9) has identified three different types of thematic progressions from theme to rheme between the clauses of a text. The first is called 'simple linear thematic progression', and basically involves a linear thematisation of rhemes in which the rheme of a sentence is selected as theme for the next one, and successively in the following ones. This can be seen in Figure 20 and example (47) below:
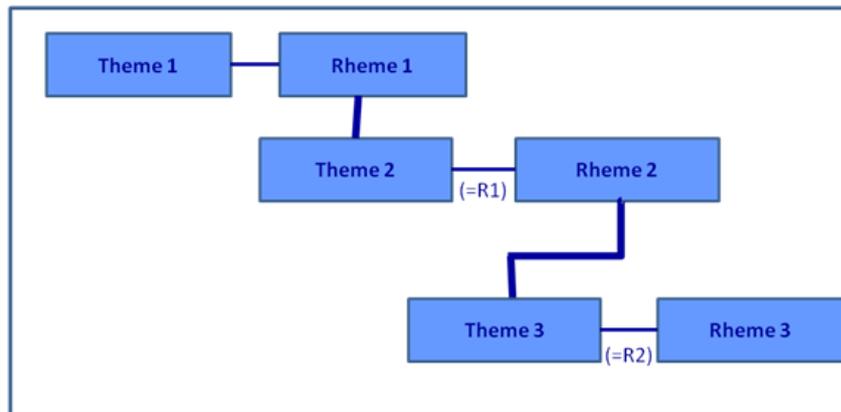
Figure 20. Simple linear thematic progression

(47)   a. *The first of the antibiotics was discovered by **Sir Alexander Flemming** in 1928.*

b. ***He** was busy at the time investigating a certain species of germ…* (Daneš in Thomas and Hawes 1997: 13) [Emphasis added]

In the second type of thematic progression put forward by Daneš, a single theme is recurrent in several different sentences, which are each linked up with different rhemes. This second model of thematic progression is called 'constant thematic progression', and an explanatory diagram on it can be seen in Figure 21 below, followed by an illustrative example:
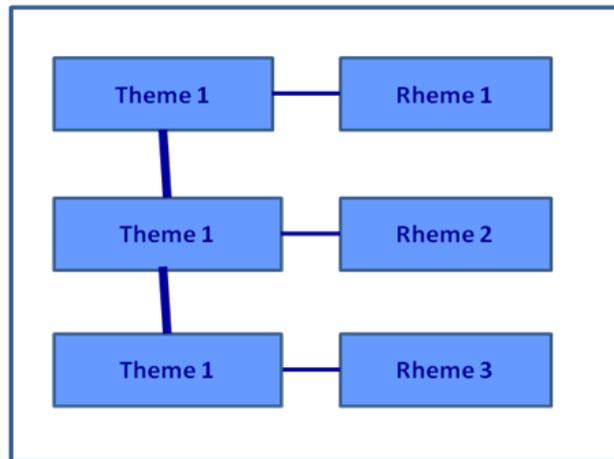
Figure 21. Constant theme progression

(48) a. ***The Rousseauist*** *especially feels an inner kinship with Prometheus*

b. ***He*** *is fascinated by any form of insurgency.*

c. ***He*** *must show an elementary energy…*

d. *Further **the Rousseauist** is ever ready to discover beauty of soul.*

(Daneš in Thomas and Hawes 1997: 14) [Emphasis added]

The third of the patterns suggested by Daneš is called 'derived theme progression', and proposes that the themes in several sentences are derived from a 'hypertheme', which may be explicitly stated or unclear in a particular sentence. A diagram on its structure can be seen in Figure 22 below, followed by an illustrative example:
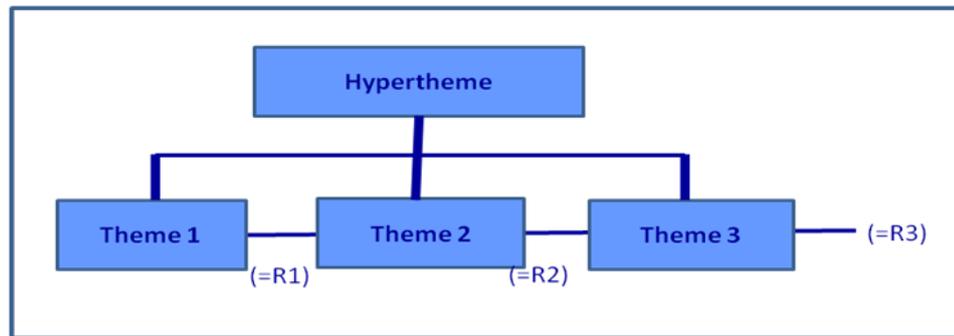
Figure 22. Derived theme progression

(49) a. ***New Jersey*** *is flat along the coast and southern portion; the northwest*
region is mountainous.

   b. ***The coastal climate*** *is mild…*

   c. ***Summers*** *are fairly hot.*

   d. ***The leading industrial production*** *includes chemicals…*

   (Daneš in Thomas and Hawes 1997: 15) [Emphasis added]

Thompson (1996a: 165) considers the different functions accomplished by themes, by pointing out that there are four main functions performed by theme choices in texts.

First, he points out that thematic choices serve to maintain the progression of the message conveyed in the text, and help readers keep the track of what the text is about. At the same time, make meaning progress towards the end of the message the writer intends to communicate. As Thompson asserts, this first function is especially well achieved through the choice of unmarked theme, which

often means maintaining the theme of the preceding clause or making progression evolve through the selection of an element from the preceding rheme.

Secondly, theme choices also mark a framework change necessary for the interpretation of the next sentence (Fries 1995). This second function is attained mainly by the selection of marked themes, either adjuncts or clauses, which include interpersonal and textual elements and function as contextual frames (hereafter CF) (Davies 1997: 55; Montemayor-Borsinger 2009: 115). CFs share the function of contextualizing the message in different ways, and different patterns of CFs, depending on their type and frequency, are characteristic of certain genres (Thompson and Thompson 2009: 57).

Thirdly, theme is selected to signal a boundary between two different sections in a text. This 'boundary-effect' is achieved through a series of theme changes that can be signalled by a summative element ('*All this…*') followed by themes which explicitly note the end of a section and the start of the next one. An example of this can be seen between the introduction and the development of ideas in newspaper texts.

Finally, themes also serve to signal the writer´s preference in choosing a specific point of departure for the message conveyed in the clause or clause complex. This explicit preference for a particular element, which the writer considers useful or important as the opening constituent of the clause, is usually marked by repeatedly choosing the same element in thematic position. The fact that the writer repeatedly chooses the same theme can also be related to Chafe's concept of 'working memory'. This concept takes into account the cognitive

capacity of the human being to maintain activated concepts, ideas and knowledge. According to Chafe (1986b), working memory is very limited and only keeps record of what is active and currently being used in our brain. This is why writers from time to time have to repeat the element chosen as theme.

In an attempt to relate the different types of thematic progression suggested by Daneš and pointed out by Thomas and Hawes (1997) with Thompson´s (1996a: 165) four functions accomplished by theme choices, we note that the first of the functions can be identified, in the case of maintenance preference, with the second of the diagrams presented by Thomas and Hawes (1997: 14), where the theme keeps repeating. Conversely if the option chosen is that of progression, this function can be identified with the simple linear thematic progression in which themes are subsequently taken from preceding clauses' rhemes (Thomas and Hawes 1997: 13).

In the case of the second and third of the functions identified by Thompson (1996a: 165), there is no straightforward correlation between any of the patterns presented by Thomas and Hawes. However, taking into account their model, corresponding diagrams have been designed to better illustrate each of the functions. On one hand, when theme is used to specify the framework being considered when reading the clause/s, either an adjunct or a full clause can be chosen as theme. This can be seen in the tentative diagram shown in Figure 23 below.
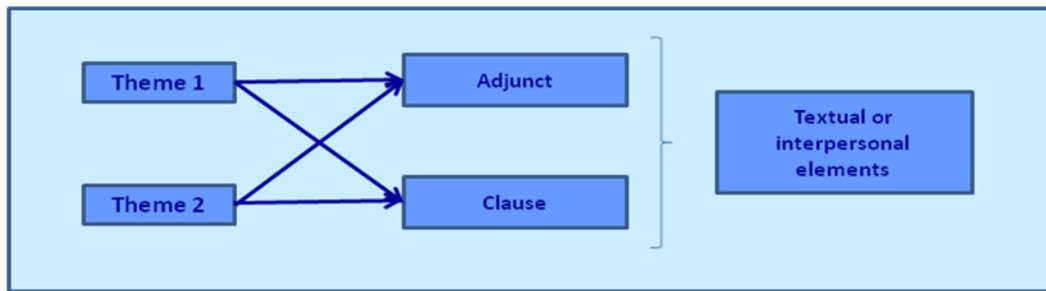
Figure 23. Theme choices of framework emphasis in thematic position

On the other hand, when theme is used as an element to mark the boundaries between sections in the text, we usually find a 'thematic triplet' (Thompson 1996a: 165). In a 'thematic triplet,' the first particle acts as a conclusion to the previous themes, followed by a particle which specifies the switch in theme that is to come in the third thematic option. This is an introductory particle of the new topic dealt with in the text. A proposal for a diagram on this function can be seen in Figure 24 below:
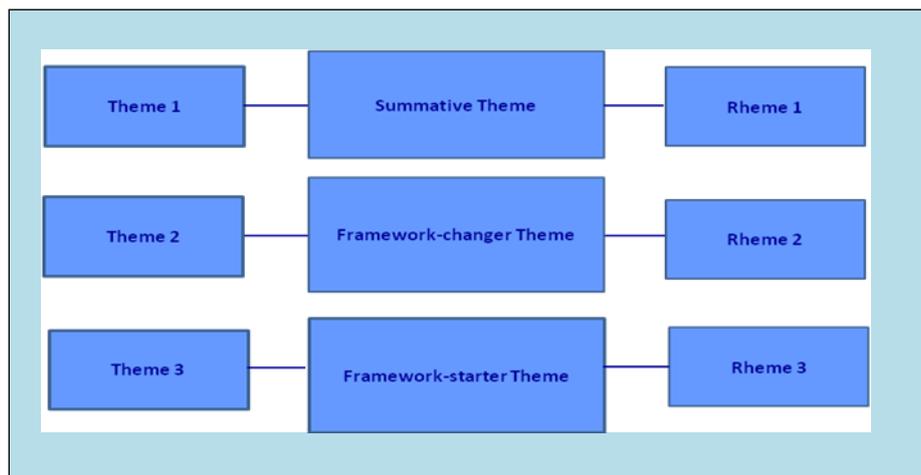
Figure 24. Theme choices of boundary emphasis in theme

The fourth function pointed out by Thompson (2004: 165), is the signalling of the departure point that the writer considers more relevant, and appropriate for repeated theme positions. This departure point can be clearly identified with Thomas and Hawes' notion of 'hypertheme' (1997: 15), or the repeated theme which consecutively occurs in a number of clauses of a text.

Similarly, in a study conducted in 1994 (quoted in Fries 2009: 25), Davies assigns each text unit (section of a text) to one of the functions: Interactive, organisational or topical, also pointing out that an examination of the themes in each of the units of a text showed that many of them (units) reveal a considerable amount of simple thematic progression.

To conclude this section, we will now extend our perception of clause as message outside theme boundaries, and briefly look at rheme. We point out that while theme sets the point of departure of the clause, rheme provides the content, which is the writer´s essential message. Not only does rheme carry the main weight of meaning, but it also sets the point of the text through consecutive rhemes that ground one idea on another (Thompson 2004: 165-166), and strongly depend on the writer´s thematic choices for the initial position of the clause.

4.4. Different perspectives on the study of theme

In the late twentieth century, during an epoch in which linguists focused on the structure of theme units, alternative studies and approaches were presented on the definition, status and behavior of theme.

One of the first definitions of theme was put forward by Mathesius. He expresses that theme can be considered "that which is known or at least obvious in the given situation and from which the speaker proceeds" (Firbas 1964: 268). From Mathesius' definition it can be deduced that all sentences have to start with an element which is identifiable as theme. However, Carter states that:

> If Mathesius is claiming that Theme items are known or at least obvious to the hearer as well as the speaker, then in what sense are these items known or at least obvious? Is this a purely 'textual' criterion, or does Mathesius include knowledge of the real world situation to which the text refers? If the former, then, as , for example, concedes, not all sentences can contain a Theme item. If the latter, the thematic status of an item can be divided only pragmatically, by reference to each individual hearer in a significant number of cases. (1986: 5-6)

As Carter states, Firbas considers that language is dependent, among some other principles, on what he calls 'communicative dynamism'(CD): 'A quality displayed by communication in its development of the information to be conveyed and consisting in advancing its development' (Firbas 1975: 317). In other words, the extent to which each element in a sentence contributes to the development of the communication. Theme is thus the element with the lowest degree of communicative dynamism, a measure which seems to increase as the clause evolves, making it stronger at the end (Thomas and Hawes 1997: 6; Vande Kopple 1991: 314).

Alternatively, as Daneš (1970, 1974) expounded, the organization of information in texts is ruled by principles of linear connectedness between sentences. For example, through linear progression, which does not necessarily require theme to take the initial position of the clause. A further point made by Daneš is that the comparison traditionally established between theme/rheme and given/new (Mathesius) is not automatically considered to be parallel. He insists that theme and given are not equivalent, and theme should not be identified with known. Conversely, Daneš considers that the relationship existing between rheme and new is identical. He states that it is necessary to extend the limits, set for discourse units at sentence level by the Prague School, in a way that theme becomes a relevant element in displaying text connexity (Thomas and Hawes 1997: 11).

In respect to the information unit, Daneš believes that a key concept is 'newness', characterized by not being mentioned in the preceding context. At the same time, it is related to what has been previously said in the same way in which rheme refers to theme. He asserts that although it is the rheme which represents the content of the utterance, it is theme that plays a key role in organizational terms by making the discourse move forward.

Regarding the application of the notion of theme to extended texts, Fries (1981: 1) claims that: "The information that is contained within the themes of the various sentences of a passage correlates with the method of development of the passage, and if the passage is outlinable, with the outline structure of the passage". His point is that the kind of information contained in the theme conditions the methods of development of the passage. This is due to the fact that thematic

_____

patterns correlate with levels of relevance of specific passages, or with the ideas that are coordinated o subordinated to others.

One of the most popular classifications of theme was carried out by Halliday (1985). He classified elements of clauses in terms of theme (or various themes) and rheme, establishing three different types of theme, dependent on the type of meaning that language conveys. First, he considers ideational themes which reflect "our experience of the world that lies about us, and also inside us, the world of imagination" (Halliday 1985: 53) and can be represented by processes, participants or circumstances around these processes. The ideational theme is the point of departure necessary for the correct interpretation of the clause. Subsequently, the rheme transmits the remainder of the message, usually leaving the most significant part at the end. As Vande Kopple (1991: 322) adds to this description: "Every clause has an ideational theme, whether it be marked or unmarked. But a clause may have two more themes, one from the interpersonal realm of meaning, and one from the textual realm of meaning". Interpersonal themes, on the other hand, are representatives of the interpersonal meaning described by Halliday as "meaning as form of action: the speaker or writer doing something to the listener or reader by means of language" (1985: 53). Interpersonal meaning is often realized through modal adjuncts and vocatives. Both interpersonal and ideational themes function as departure points in clauses. As Halliday states, if some words appear before the ideational theme in a clause, especially those bearing interpersonal meaning, they may be considered theme as well, but not if they appear after the ideational theme.

Thirdly, textual meaning enables writers to tie clauses together and create a cohesive and coherent message, as it denotes the relevance of the context, both for the preceding and following text, as well as the context of situation (Halliday 1985:53). The same happens with textual themes as with interpersonal ones; they only constitute part of the theme if preceding the ideational meaning, but not if they are following it (Vande Kopple 1991: 323).

4.5. Theme in reporting clauses

In this section of the chapter we will review the primary notion of reported speech, which we covered in detail in Chapter 3. In Thomas and Hawes' terms, speech reporting has been described as "the use of language in one situation to represent the way language was used in a previous situation" (1997: 28). This definition implies the existence of two different language events which are independent speech situations which are brought together in a clause or clause complex. As they assert, the first situation can be considered the original communicative event, or that in which a linguistic utterance was made, while at the same time, a second linguistic utterance has also been produced. Leech (1983) already named these two communicative events, 'primary' and 'secondary speech situations'. He noted that the speaker in the secondary situation, who is recounting the language event in the original situation, has two different options in carrying out that procedure: direct and indirect speech representation, otherwise called quotes and reports. In the case of quotes, the speaker in the secondary situation

repeats the same words that were uttered in the original situation setting formal and semantic links between the two utterances. Conversely, in the case of reports, what is conveyed in the reported clause is just the essence of the communicative event on the whole, carried out using the reporter's own words, while maintaining semantic links with the original utterance (Thomas and Hawes 1997: 28).

With regards to the analysis of theme in speech representation, Thompson and Thompson (2009: 51) assert that themes in the projecting clauses highlight the source of the information, while those in the projected clauses signal the content of the study being reported. This idea emphasizes that the nature of projection is indeed of a logico-semantic character. From one point of view, the projecting clause is structurally and semantically dominant, but, from a different perspective, it is the content of the projected clause that has a greater weight, in informative terms. Thus the projecting clause serves as a kind of auxiliary 'source tag', especially perceivable in clauses where the projecting clauses come at the end or are parenthetical. This can be seen in Example 50 below, in which the projected clause occupies the initial position, immediately followed by the reporting clause where the statement in the reported clause is attributed to an academic:

(50) *A sexual charge between pupils and teachers is sometimes a feature of good teaching, an academic has claimed…* (Thompson and Thompson 2009: 51)

As Thompson also asserts (1996a: 161), theme analysis of quotes is usually straightforward, and suggests that the reporter makes a theme choice in the reporting clause but, at the same time, recycles the theme originally uttered by the speaker in the reported clause, the 'quote'. This can be seen in Table 11 below.

She	Whispered:	"I	don't like it at all"
Reporting clause		Reported clause, quote	
Theme	Rheme	Theme	Rheme

Table 11. Reporting clause in thematic position (quote)

There is no problem if the reporting clause precedes the reported clause, or if they are placed in the reverse order. Each of them has its own theme, and therefore each contributes to thematic progression of the text or conversation where it appears, in an individualized way. In the example below (Table 12), the order of the previous example has been reversed. It is the reported clause in a direct style (quote) that comes in the initial position, followed by the reporting clause that provides the reader with the information of who and how the utterance was articulated.

"I	don't like it at all"	She	whispered:
Reported clause, quote		Reporting clause	
Theme	Rheme	Theme	Rheme

Table 12. Reported clause in thematic position (quote)

Regarding the analysis of theme in reports, Fries (1994) proposed the notion of 'T-Unit', an independent clause together with all the clauses that are dependent on it. This is an analysis which will entails that the reported clause form a T-Unit with the reporting clause in such a way that theme doesn´t need to be shown separately. Conversely, as Thompson puts forward (1996a: 162), there is an alternative analysis in which theme is taken as a separate message on a separate

level than the projecting clause, and thus analyzed independently in each of the clauses, in an identical way as done in quotes. As seen in Tables (13) and (14) below, reports have a hypotactical dependence on the reporting clauses, that serve as introductory elements which provide information that help to better understand of the information to come. In Table 13, the reporting clause has been chosen as the departure point of the sentence, and the author of the utterance reported (*Martin*) as theme. This stresses that it is relevant for the author to assure that the situation be taken more seriously. Conversely, in Table 14, the reported clause is located first, being theme position occupied by *The global crisis*, the event that centers the focus of the construction, and the speaker being located in rhematic position, making no emphasis on his identification (*he*).

Martin	assured	that the situation	had to be taken more seriously.
Reporting clause		Reported clause	
Theme	Rheme	Theme	Rheme

Table 13. Reporting clause in thematic position (report)

The global crisis	will have completely ended by 2013,	he	announced yesterday.
Reported clause		Reporting clause	
Theme	Rheme	Theme	Rheme

Table 14. Reported clause in thematic position (report)

In a deeper analysis carried out by Thomas (1983), she began by examining the choices available for writers when making this kind of speech reporting. She

specifically focused on different thematic options available for writers in independent clauses, and how their distribution and election can be useful for identifying patterns and features of reporting in the corpus of articles studied, in this case, research academic articles.

Thomas departed from the hypothesis that thematic elements determine the syntactic form of reports, making it possible to categorize reports according to their individual thematic option. Different syntactic forms of reports also condition their function in the context in which they occur. Moreover, the thematic choice of individual reports is connected to the other reports in the text, creating a thematic organization which reveals the connectedness of all reports in a particular context (Thomas and Hawes 1997: 31). She identified five different possibilities for the writer to set the initiation of the clause: agent as theme, a text reference term as theme, a content term, a predicated theme and bound clauses acting as theme. All five types of theme recognized in the corpus studied also constitute five different types of reporting in academic research articles, and will be taken into account later in the study developed.

In this chapter, we have attempted to review the most relevant aspects of theme by gathering perspectives from different authors, and taking into account the relevance of theme for the development of meaning in reporting speech. We can conclude that thematic elements seem to have a relevant role both in signalling reports and discourse organization through the interconnection between reports. In addition, thematic elements establish the point of departure of the clause, freely chosen by the writer. As Firbas (1986) and Fries (1981) have explained, theme is also an indicator of the initiation of 'communicative

dynamism' which occurs in clauses and advances the 'newsworthiness' of the clause, emphasized on the endmost elements.

At this stage of the study, we have already reviewed the most relevant literature on the three notions that will be used as the basis for the analysis that will be developed in subsequent chapters of this doctoral dissertation. On the one hand, we have examined the peculiarities of newspaper and science popularization discourse, focusing on the special characteristics of science popularization articles. These are articles with scientific content but printed in a publication aimed for a general, non-specialized audience where the journalist acts as a mediator, bringing the scientist´s knowledge to the general newspaper readers. On the other hand, and as a second ground for this study, we have explored in further detail one of the most common phenomena in science popularization discourse: the attribution of facts conveyed to voices different from that of the journalist. Reports and quotes have been studied in the different classifications made on the issue, from a systemic functional perspective as the guiding perspective to deal with this form of epistemic modality. Thirdly, in this chapter we have brought to the forefront the importance of theme structure, reviewing in depth the roles of theme and rheme and how they contribute to the thematic progression of texts and explicating explicit information. Bearing all of these three elements in mind, in the following chapters we will present an analysis of the thematic choices made by science popularization journalists from the British newspaper *The Guardian,* in the quotes and reports introduced in their texts to attribute scientific facts to external voices.

Part B: Empirical Study

Chapter 5

Study description

_____

Chapter 5

Study description

5.1. Introduction

5.2. Departure point for the study

5.3. Hypotheses and objectives

    5.3.1. Hypotheses

    5.3.2. Objectives

5.4. Corpus

    5.4.1.Compilation

    5.4.2. Description

5.5. Methodology

    5.5.1. Computational linguistics and *WordSmith Tools*

        5.5.1.1. Quantitative and qualitative data

        5.5.1.2. Corpus processing software: *WordSmith Tools*

    5.5.2. Analysis' phases

    5.5.3. Tags

    5.5.4. Annotation and analysis layers

5.6. Final Remarks

Some words are lonelier than others
J. R. Firth

5.1. Introduction

Recently, there has been a growing concern over the use of language in different contexts and different text types, a field where corpus and discourse studies play a major role. Corpora provide researchers with real instances of language in context seen as valuable data to get meaningful results in the frame of corpus linguistics and discourse analysis. Corpus linguistics, described as "[…] a way of investigating language by observing large amounts of naturally-occurring, electronically-stored discourse, using software which selects, sorts, matches, counts, and calculates" (Hunston and Francis 2000:14-15), has meant a qualitative

_____

change to our understanding of language (Tognini-Bonelli 2001:1) by enabling researchers to analyse vast quantities of real data in a systematic way.

The study presented in this doctoral dissertation, has been born out of the aforementioned advances in corpus linguistics. Its main purpose is to contribute to a better characterization of explicit attribution within language events' reports in the context of British press discourse, and, more specifically, in science popularization articles from the newspaper *The Guardian*.

To obtain data on explicit attribution in language events' reports, we have carried out an analysis on the resources that are most commonly used to enclose external voices in the text. This analysis will provide us with patterns and frequencies of appearance that will lead us to observe tendencies and draw conclusions on the anatomy of this phenomenon. Patterns and distribution of quotes and reports will be examined in the corpus, together with complex cases of attribution, namely double and multiple cases. This is done in order to get information on the procedure that is most commonly used by journalists in this type of text. More specifically, we will find out how present these resources are in first paragraphs of the texts in the corpus, a presence (or absence) that will differentiate these sections from the rest of the text. At the same time, we will focus our attention on one of the alternative resources used to explicitly incorporate into the text, the voice of somebody else different from the journalist, the adverbial *according to*, and its different patterns of frequency and distribution in the texts. This will provide us with information on thematisation patterns in language events' reports and data on its specific presence in initial sections of the articles that form the corpus, together with data on the attributes that most

frequently accompany this particle. Finally, we will turn our attention towards reporting verbs, which will contribute to the obtaining of relevant information on the most common processes in language events' reports, and the kind of meanings that they typically transmit.

In the following pages, we will provide the basis of this research and develop a detailed description of the compilation process and parameters of the analysis of the corpus of science popularizations from the British newspaper *The Guardian* that we have taken as object of study.

5.2. *Departure point for the study*

As it has been already described in Chapter 3, numerous studies have focused on the language of newspapers, yet few of them have concentrated their efforts on describing science popularizations. Science popularization articles have recently attracted the attention of linguists as objects of study in order to obtain relevant data on the re-contextualization of knowledge that is carried out in these types of texts. Their growing presence in national newspapers and, as a consequence, the wide readership of these texts, has also enhanced their likelihood of reflecting the textual features that are conventions in the community of speakers they represent.

As a result, numerous studies carried out in the last few years have provided us with relevant information on the characteristics of science popularization

articles, and, more specifically, on the elements and structures that most frequently appear in language events' reports of science popularizations.

García Riaza and Elorza (in press) carried out a study on science popularization articles of the British newspaper *The Guardian* which aimed at getting information about the characteristics of first paragraphs in science popularizations in English, where sources of attribution are incorporated into the text, and the presence (or absence), frequency and distribution of explicit attribution sources in the corpus studied. For this purpose, they analysed patterns which are typically encountered in first sections of science popularizations in English and thematised elements in first paragraphs of the articles.

Of the attribution sources found, the study revealed that it is more frequent in science popularizations in the British press to refer to people rather than to material entities, such as studies, research processes or publications, as authorized sources in the text. The study also observed a clear tendency for them to be introduced at the beginning of the text.

In relation to the position that explicit attribution sources typically occupy in sentence and paragraph structure, García Riaza and Elorza found that while attribution sources are mainly located in rhematic position in sentence structure, figures referring to paragraph structure show that explicit attribution sources appear in rhematic position as frequently as they do in medial one. Concerning reporting verbs, they concluded that attributional meanings are predominantly conveyed through verbal processes of 'saying', expressed by means of both quotes and reports. They also identified a stronger tendency towards quoting

rather than reporting in the corpus but this tendency is reversed in first paragraphs, where all of the verbal processes of 'saying' encountered are reports.

In relation to the structures that most frequently convey attributional meaning, the most relevant pattern they encountered in thematic position was that of the explicit source being conflated with the theme of the sentence. In rhematic position, the most repeated structure which conveys attribution is 'according to', data which support the findings made by García Riaza (2010).

Additionally, Elorza (2011) focuses on reporting verbs in science popularization articles and the different positions they occupy with regards to the Subject of the clause in which they appear. With this study, she intends to portray the role and characteristics of reports through attribution in *The Guardian* and *El País* newspapers, with special attention on the meaning of thematization of elements in this text type across languages.

She finds that neutral verbs are the most common option in the British newspaper, as opposed to those showing how the message fits in or the attitude of the journalist towards the report. Conversely, the most frequent option in the Spanish newspaper is that of showing the speaker's purpose and should therefore be considered the non-marked option. In regards to the structure primed by reporting verbs in each of the publications, she finds that popularizations in *El País* use a greater variety of verbs, many of which are more flexibly used both in S+V and V+S structures. She highlights this as a possible reason that texts in the Spanish newspaper are less conventionalized than those in *The Guardian*. Reporting verbs showing the speaker's purpose were also found to be more

_____

frequent in *El País* than in *The Guardian*, where the verb *say* was detected as the non-marked option. Elorza related the higher percentage of the verbs showing the speaker's purpose in Spanish to a more explicit presence of reports' evaluation than in the English one. Furthermore, it was detected that journalists in *El País* texts rely on indirect report more often than their colleagues at the British newspaper, who more frequently use direct quotation.

As mentioned earlier, García Riaza (2010) focuses on a specific linguistic device used by authors to popularize science in initial sections of science popularization articles from the British press: the epistemic stance adverbial *according to*. She aims at contributing to a better characterization of language events' reporting in newspaper discourse by analyzing how frequently the adverbial *according to* appears in first paragraphs, as compared to its presence in the whole corpus, also identifying the elements which most frequently accompany the particle according to in right position.

She observed that first paragraphs concentrate an important number of occurrences of *according to* in the whole corpus, an adverbial which was found not to be typically followed by a personal reference, a team or scientist to whom the scientific discovery is attributed, but by a reference to a material entity. In relation to the typical positions that this particle occupies in first paragraphs, *according to* showed its rhematic character with a higher number of occurrences in final position.

Regarding the elements that most frequently accompany *according to*, she observed that the adverbial occurs followed by an NP. However, also detected

was a high occurrence of Prepositional Phrases conveying different meanings, such as authorship, location or identification of the subjects participating in the study. In addition there were occurrences of relative clauses, which expand the meaning of the antecedent complemented, and time adverbials, which tend to appear at the very end of the sentence, expressing the freshness of the news reported.

Positioning herself a step forward, García Riaza, (in press) aims to understand how the phenomenon of explicit attribution differs in frequency in science popularization articles of the newspaper *The Guardian*, as compared to a different text type within the same publication: editorials.

With this study, she wanted to test if explicit attribution, conveyed through language events' reports in the text, was present in corpus science popularizations of the British newspaper *The Guardian* in a higher degree than in a corpus of editorials from the same newspaper. This would confirm that the particle is not genre-dependent, but text type dependent and strongly associated with popularization in the contemporary press, also obtaining information of its most frequent collocates in the corpus.

The frequency of occurrence of the particle *according to* in both corpora revealed that the adverbial is more characteristic of science popularization articles than newspaper editorials, where there are scarce occurrences of it. The discourse marker *according to* revealed itself as a typical construction in science popularization articles, characteristic of the plurality of voices and attribution of

_____

scientific facts present in the aforementioned text type, rather than an element present in editorials.

At the same time, reinforcing what she had already stated (García Riaza 2010), more occurrences of *according to* were found with a material rather than a personal reference This leads us to conclude that entities are the most common collocates of the discourse marker, with special incidence of entities including 'poll', 'figure', 'publication' or 'report'.

The corpus she studied also provided new insights into the positions that are most commonly occupied by the discourse marker *according to*, which showed its rhematic preference in the corpus of science popularization articles, and a high presence of this element in medial position, especially in the corpus of editorials.

Taking a cross-cultural approach, García Riaza and others (2011) carried out an analysis of the discourse of science popularizations that focused on language events' reports in English, Spanish, Italian and German. They based their study on Thompson's (1996b) classification of resources used to construct language events' reports: *voice, message*, *signal* and *attitude*, to look for recurrent patterns in the way language events' reports are constructed in each sub-corpora.

The analysis resulted in a variable degree of specificity when participants are introduced as external sources of attribution in the language events being reported. The typical structure found for introducing external attribution in English is that of a noun phrase consisting of a participant acting as head followed by an optional post-modifier, which frequently has the form of a prepositional phrase or a relative clause. The most common way to enclose the message in the

study carried out by García Riaza and others (2011), is by quoting. However, a relevant number of cases of paraphrases have also been detected, which means that it must also be considered as a rhetorical device characteristic of the science popularizations they studied.

Regarding reporting signals, they classify the occurrences of the corpus into two groups, those with an explicit reporting verb and those in which the verb was omitted. In cases where there is no reporting verb, the most common form used to construct the language event report is that of an adjunct in English, Spanish and Italian (*according to, según, secondo* respectively), and by means of the Konjunktiv I in German.

Additionally, García Riaza and Pérez Veneros (in press) conducted a comparative analysis of heads from *The Guardian* and *El País* newspapers to obtain data on the presence (or absence) of explicit attribution in initial parts of science popularization articles. They analysed headlines and leads in order to contribute to the previous knowledge in the field, adding information on initial parts of science popularization that had never been described before. They detected that the presence of explicit attribution is much higher in the British newspaper than in the Spanish one. This fact highlights the character of newspaper discourse, not only as a vehicle to convey textual and cultural features of a certain community, but also as a medium to disclose linguistic conventions of the two countries studied.

Whereas in *El País* no cases have been encountered in which the explicit attribution occurs both in headline and lead, the newspaper *The Guardian* shows a

_____

tendency to duplicate the attribution in headline and lead, even though sometimes there is no full coincidence between them (*experts-psychologists*). No trace of the epistemic stance adverbial *according to* has been found in the corpus of heads. This fact along with García Riaza`s (2010) assumption about the high presence of this adverbial in first paragraphs has led to the possible conclusion that due to a journalistic convention, *according to* is more likely to appear in the text rather than in the head, priming its presence in the initial section of the article.

Regarding transitivity, García Riaza and Pérez Veneros point out that the presence of participants in both newspapers is quite distinct, as the Spanish one shows a tendency to incorporate material references, while the British one primes the existence of personal references, in different degrees of detail. In the majority of the cases, both newspapers chose to include explicit attribution by means of indirect report with a *verbum dicendi*, rather than including a direct report.

Furthermore, Elorza et al. (2011) aim to identify and describe patterns of the way attribution and averral (Hunston 1999) are used in science popularization articles across different European languages, namely English, Spanish, Italian and German. For that purpose, identification, analysis and comparison processes were performed on the 111 language reports detected in the different corpora. The main focus was placed upon reporting verbs and they were divided according to Thompson's (1994:33ff) categories of the different functions that can be performed by a reporting signal. As a result, it was noted that in *The Guardian* newspaper, a great majority of the verbs can be classified as 'neutral', followed by a relevant number of verbs which accomplish the functions of identifying the attributee's purpose and transmitting something that was thought by the attributee

(mindsay). A low percentage of occurrences of 'neutral (writing)' reporting verbs were also found, but no instances of verbs 'indicating how the message fits in' and transmitting the journalist's evaluation on the information were detected in the British newspaper. 'Say' was identified as the most frequent reporting verb, and therefore the non marked option, closely followed by 'suggest' and 'believe', while verbs such as 'declare', 'think' or 'write' show a lower percentage of occurrences.

Thematization patterns were also analysed by Elorza et al. (2011) in relation to the elements that are most prone to appear in initial position within the language event reported. In the case of *The Guardian* newspaper, the most common elements thematised are the findings, both in direct (quote) and indirect forms (report). In a lesser degree, scientists are placed in thematic position of language events in the British newspaper, an element which primes its construction as indirect report, rather than as its direct counterpart.

Data provided by these recent studies served as a basis of knowledge for a small-scale study, which took the outcomes of the research carried out in the aforementioned studies as a guide to establish its objectives and research questions. First, the aim was to learn more on the anatomy of initial sections of science popularization articles, and, more specifically, to test if there is a tendency for first paragraphs to be constructed by a single sentence.

To test this hypothesis, the average number of words and sentences was measured in first paragraphs of two similar publications: *El País* and *The Guardian*. As seen in Tables 15 and 16 below, first paragraphs of the newspaper

_____

El País tend to be significantly longer than those in the British newspaper with 86.5 and 34 words per paragraph respectively. Regarding the number of sentences that comprise first paragraphs in each of the cases, data suggests that the average number of sentences found in first paragraphs in *The Guardian* is one, as opposed to three sentences in *El País.*

LENGTH OF FIRST PARAGRAPHS	
NUMBER OF WORDS	**NUMBER OF SENTENCES**
34 (average number)	MEDIAN OF SENTENCES = 1 (HIGHER VALUE = 2, LOWER VALUE= 1) One sentence paragraphs=17 Two sentence paragraphs=8

Table 15. Average number of sentences in P1 in *The Guardian* newspaper

LENGTH OF FIRST PARAGRAPHS	
NUMBER OF WORDS	**NUMBER OF SENTENCES**
86.5 (average number)	AVERAGE NUMBER OF SENTENCES = 3 (HIGHER VALUE = 6, LOWER VALUE= 1) One sentence paragraphs: 2 Two sentence paragraphs:6 Three sentence paragraphs:11 Four sentence paragraphs:4 Five sentence paragraphs:0 Six sentence paragraphs:2

Table 16. Average number of words in P1 sentences in *El País* newspaper

_____

Data provided in this section have contributed to the study carried out in this doctoral dissertation. Results and outcomes have been taken as a departure point to take a step further in the characterization of explicit attribution in language event's reports from the newspaper *The Guardian*.

5.3. Hypotheses and objectives

In the development of this study, we depart from the idea that language events' reports are a frequent phenomenon in newspaper discourse (Thompson 1996b: 505), and thus a fruitful source of information on how journalists give access to external voices in their texts. As we have previously seen, journalists act as bridges between science and the lay reader, whose scientific knowledge is restricted. Moreover, journalists writing science popularization articles also act as spokespeople for the authors of the scientific advances reported. In this sense, they do not only reconstruct the information for non-specialised readers to understand it, but also give access to the voices of scientists through a series of devices that will be the focus of attention in this study. Based on this departure point, hypotheses and objectives that have underlain this study are presented below.

_____

5.3.1. Hypotheses

Once the rationale of this study has been detailed in the previous section, we can establish the departure point of this doctoral dissertation by taking into account that language events' reporting has revealed itself as a common feature of science popularization discourse, in which the presence of explicit attribution is also a relevant factor. To attain a better characterization of the phenomenon of explicit attribution in science popularization articles from the newspaper *The Guardian*, we have formulated a series of hypotheses which have been divided in different groups:

- *Thematisation in first paragraphs*

We want to obtain information on the patterns which are recurrent in first paragraphs of science popularization articles from *The Guardian* newspaper, as well as the special patterns that are displayed in first paragraphs due to their initial character.

Our first hypothesis is that initial sections of newspapers, due to their high degree of informative nature and introductory character that give the reader new information provided by the text, will contain a higher proportion of explicit attribution cases than non-initial sections of the articles. The inclusion of scientists' voices is a relevant resource in the discourse of science popularizations that places a higher relevance on the information. If these voices are so important, we believe that the journalist will present the reader with the words uttered by

experts as soon as possible, and thus an occurrence of explicit attribution will be introduced at the beginning of the text[1].

For this reason, we attempt to describe the most relevant features of explicit attribution in initial sections by hypothesizing that:

- It is more frequent to find language events' reports in first paragraphs in which the verb is omitted than in those in which the verb is explicitly stated, due to the high presence of epistemic stance adverbials which do not need any additional reporting signal.

- The epistemic stance adverbial *according to* will be more frequent in initial sections than it is in the rest of the text, that is to say, in non-initial sections of science popularization articles occurrences of this adverbial will not appear in a high percentage.

- Reports will be preferably used over quotes in initial sections of the newspaper studied, due to the introductory nature of first paragraphs.

Testing those hypotheses will not only lead us to understand what is characteristic of first paragraphs of science popularization articles in the British press, but also learn which of the patterns obtained are characteristic of first paragraphs in a higher degree than in the rest of the text. Moreover, we will be able to obtain data on which patterns are exclusively found in initial sections and nowhere else in the text, if any.

_____

[1] The focus is set on the body of texts, although attribution could also be present in heads of newspaper articles (headline and lead).

_____

- *Verb presence and reporting verbs*

Regarding verbal processes in the corpus, we want to know the most frequent reporting verbs used, as well as to get information on the frequency of occurrences of other structures that do not require the presence of a verb.

We, accordingly hypothesize that:

- There is a stronger tendency for reporting verbs in the corpus to be explicitly stated in the language events reported.

The rationale is that from the reporting verbs encountered, there will be a group of verbs more frequently used than the rest of them, therefore allowing us to be able to list the most frequent verbs used to report language events in the corpus. We are also assuming that reporting verbs in *Sci_TG Corpus* can be classified according to their meanings, into existing taxonomies and families, such as Thompson's (1994) and its reformulation in Elorza (2011).

- *Epistemic stance adverbials*

With regards to the presence of epistemic stance adverbials, we hypothesize that most of the cases where there is no reporting verb correspond to cases in which the reporting signal is an adjunct such as *according to*.

Being an adjunct, *according to* will focus the message being placed in both thematic or rhematic position. Therefore, we hypothesize that although the

attribution source does not appear in thematic position in cases in which the epistemic stance adverbial is involved, the adverbial followed by the attribution source can be set as theme in the sentence, even if it is not conflated.

Regarding the epistemic stance adverbial, we also consider the following as working hypotheses:

- *According to* is a structure that allows the journalist not to set the source of attribution as subject, but as theme, and consequently, provide the source with the prominence of the initial position.

- The use of the epistemic stance adverbial *according to* in science popularization articles from the British newspaper *The Guardian* is a frequent and efficient resource to introduce the attributee before or after the message conveyed, without placing the attributee as subject of the sentence.

5.3.2. Objectives

By characterizing explicit attribution through epistemic stance adverbials such as *according to*, reporting verbs and patterns identified, as well as by elements of Halliday's transitivity system, we place ourselves a step forward in the description of science popularization writer's roles, and how they act as mediators between the scientists and the layman.

_____

The main focus of this study has been placed on the linguistic devices and discursive resources used to access external voices of science popularization texts, and our aim is to:

-   Seek to identify the resources most frequently exploited to incorporate external voices in the articles by the scientific correspondents in the corpus compiled.

-   Contribute to a better characterization of explicit attribution present in language events' reports in science popularization discourse from the British newspaper *The Guardian.*

-   Get objective data on the characteristics of initial sections in science popularization articles from the British press, in relation to how the phenomenon of language events' reporting is constructed.

To achieve these objectives, we will use several procedures that will help us test the hypothesis through the data obtained in the analysis. We will identify recurrent patterns in the way language events' reports are constructed in the corpus, specifically focusing on reporting verbs and their position with respect to the subject. Data on how frequent the particle according to is in P1, compared to its frequency in the whole corpus *Sci_TG* will also be extracted, to finally provide data on the presence of the epistemic stance adverbial 'according to' and identify the elements which most frequently accompany this adverbial especially one word to the right (Right 1 position).

_____

5.4. Corpus

As we have previously stated, this study focuses on science popularizations of the British press, and more specifically, on science popularization articles from the newspaper *The Guardian*. The object of our analysis is a corpus of science popularization articles which were collected from the electronic version of the newspaper available at www.guardian.co.uk.

This newspaper is a quality broadsheet aimed at a relatively educated audience in Great Britain (Kim and Thompson 2010: 62) and according to Wikipedia (http://es.wikipedia.org/wiki/The_Guardian), has the most numerous print run of the country with 350,000 daily copies. As a result, this publication has been frequently used in comparative studies such as Hyde (1999), Mahlberg, O'Donnell, Hoey & Scott (2007), García Riaza (2010, in press), Elorza (2011), Gianoni (2008) or Pounds (2010). As Williams (2009: 468) explains, *The Guardian* is a prestigious elite British newspaper, which was founded in 1821 under the name of *The Manchester Guardian*. It is owned by the Guardian Media Group, and politically aligned to the left.

The Guardian's science coverage is extensive both in the printed and in the electronic version, and there are a varied number of journalists, some of whom are also scientists, who are concerned with the popularization of scientific events in the newspaper. Names such as Ben Goldacre, who has published a book[2] compiling scientific columns published in the newspaper, or Ian Sample, the

_____

[2] Goldacre, B. (2009) *Bad Science: The Short Life and Weird Times of Cold Fusion.* London: Fourth Estate.

_____

journalist with the most numerous production in this corpus, constitute the top of the group and sign their articles as 'science editor' at different times along the period covered in this corpus. Together with the science editor, journalists who write for the scientific section of the newspaper usually employ the label 'science correspondent' to identify him/herself with both the section and the content. Some of these authors are: Alok Jha, James Randerson or David Adam.

It is relevant to remark that all texts compiled for the study are in electronic format, which means that they appear in an electronic environment that may condition their construction. Electronic and printed versions of *The Guardian* are not coincident (Elorza 2010). As an example of this, the majority of electronic texts are accompanied by an image that illustrates the discovery which is being reported, something which is not so common in printed articles, and may be due to space constraints and colour ink costs.

5.4.1. Corpus compilation

In this study, the compilation process has been carried out in two subsequent stages, dependent upon each other. In the first stage of the compilation process, articles that form the corpus were manually retrieved from the website of the British newspaper, using the search engine provided by the website. Several preliminary searches were carried out to identify the better procedure. Then, the search was performed using the newspaper archive, and refining the search engine using a set of characteristics shown as relevant in the first retrieval attempts. The section 'Science' was considered as a major selection word for the search. We

then performed a monthly revision of the articles, together with an identification of the journalist writing them, to ensure that no text from a different section was selected in the search process. At the same time, *The Observer* was set aside in the search process, as it is a Sunday edition which appears with different features under a different name. We also took into account articles published in *The Guardian Weekly*, a publication that summarizes the articles published during the whole week, which means that a total of 81 texts were incorporated into the corpus.

Once the whole corpus was compiled, a second stage of the compilation was carried out to ensure the cleanness and homogeneity of the corpus. In spite of the careful search process that was executed, some articles were repeated and we also observed that a small percentage of the corpus was formed by texts which cannot be fully considered popularizations. These include dispatches, press releases and repertories, which are short texts that provide readers with concise information, but whose lengths are not sufficient to study the phenomenon of attribution in detail.

In addition, images and their corresponding footnotes have been removed from the files to obtain a textual corpus, not a multimodal corpus which would have to be analysed using different procedures. Although the structure of texts has been preserved, heads have not been taken into account in the analysis, because, as it has been stated, newspaper heads (including headlines and leads) constitute a differential element in the text, acting as a hook to attract readers' attention. As Zapaterra (2008: 61) points out, there are two different parts in heads: the headline, which acts as tie between the publication and the reader, and the lead, which

_____

constitutes a bridge between the text and the head, and whose content, occupying

the second place, is more relevant than that of the head as it informs the reader of

the topic and intention of the article. Reah (2002: 16) adds that press heads use

short words to capture the interest of readers by using ambiguity, set phrases and

proverbs.

The corpus, which has been named '*Sci_TG*' alluding to the common

content and publication of all articles in it, contains 567 science popularization

articles published in the British newspaper during the entire years of 2006 and

2007.

5.4.2. Corpus description

Sci_TG Corpus has a total number of 567 texts, and, as seen in Table 17

below, reaches a number of running words around 360,000 with the number of

word types being slightly higher than 19,000.

	Sci_TG Corpus
Number of texts	567
Running words	363,636
Types	19,069
Type-token ratio	5.39

Table 17. Characteristics of *Sci_TG Corpus*

As it has been already stated, the corpus studied is composed of science popularization articles, but, within this text type, there are different texts that can be identified and classified according to their diverse lengths. We have identified three types of texts that differ from an article. We shall refer to them as 'short texts' from now on, to differentiate them form complete articles taken into account in the analysis, which are referred to as 'articles'. Articles, which on average have around 500 words, consist of a head, which is normally composed of a headline and a lead, together with metadata that locates the article in a date and page of the newspaper and identifies in detail the journalist who wrote the article. Below the head we find the body of the text, consisting of a series of paragraphs where the information is fully developed by the journalist, as we can see in the Example below.

_____

Sci_TG07102
15 August 2007, The Guardian, p.14
Science: Eating for two puts unborn child at risk of junk addiction / Home News
Randerson, James

Pregnant women who overindulge in junk food risk giving their child an addiction for a fatty, sugary and salty diet, according to researchers who studied the phenomenon in rats. The study counters the notion that mothers-to-be can safely overindulge because they are "eating for two".

Although the effect on rats may not necessarily be found in people, the research does back up a US study of 190,000 families published in 2005. It found that women who gained more weight during pregnancy than the US Institute of Medicine's recommended amount - 11.5 to 16 kilos (2 to 2.5 stone) - were more likely to have obese two to four-year-olds.

Neil Stickland at the UK's Royal Veterinary College and his team fed pregnant and breast-feeding rats on either a balanced diet or processed food such as doughnuts, muffins, biscuits, crisps and sweets. They then gave the offspring a choice of diets. The mothers who ate the high fat, sugar and salt diet gave birth to young with a greater preference for junk food. They also had a propensity to overeat and to put on more weight.

"They eat more and they eat an increasing proportion of junk food. It gets worse as the animal gets bigger," said Prof Stickland.

"Women may not consider pregnancy and breast-feeding as an opportunity to over-indulge on fatty, sugary and salty foods, on the misguided assumption that they are 'eating for two'," the team warn in the British Journal of Nutrition. The study was funded by the Welcome Trust.

Prof Stickland added that programmes to improve the diet of children at primary school, while important, may be too late. "I think the message is that we have got to start before that during pregnancy."

The research did not address how the changes in offspring occur. However, Prof Stickland speculated that a mother's junk food diet might affect the development of reward centres in the brain involved in the feeling of satiety and response to drugs. "The foetus is getting used to the high fat, sugary and high salt diet and seems to prime the reward centres in the brain so it needs more when it is born," he said, "It's an addiction if you like."

Example 51. Model of an article in *Sci_TG Corpus*.

We have selected the body of 567 articles to carry out our study and analyse attribution patters in a corpus of science popularization articles. As seen in Table 18 below, bodies represent the majority of running words in the *Corpus*.

_____

| | Sci_TG Corpus ||
	Body	Heads
Running words	260,848	14,877
Types	16,575	2,984
Type-token ratio	6.48	22.50

Table 18. Size of body and heads in science popularization articles of the corpus.

The final total number of running words that have been analysed is 260,848, representing the words contained only in the body of articles from *Sci_TG Corpus*. A deeper division to facilitate the analysis has also been carried out between first paragraphs and the rest of the paragraphs contained in the article (Table 19), thus creating one Sub corpus named *Sci_TG_P1 Sub corpus* and another named *Sci_TG_Rest Sub corpus*. This differentiation between initial parts of the articles and the remaining parts is based on the notion that frequencies of words and patterns do not only vary across different text types, but also within texts from a single variety. Therefore, words can have different frequencies within different sections of the text that allow us to distinguish them by word distribution (Sinclair 1991; Stubbs 1996; Hoey 2005; Mahlberg and M. O'Donnell 2008). Moreover, first paragraphs have been considered highly informative sections in texts (Mahlberg and O'Donnell 2008, Ho-Dac 2007), which is the reason why we want to get a deeper insight into the anatomy of first paragraphs in science popularization articles of the British newspaper.

As can be seen in Table 19 below, *Sci_TG_P1 Sub corpus* has a total of 4,692 distinct words, while *Sci_TG_Rest Sub corpus,* much bigger in terms of size, has 15,983 types, which does not represent such a bigger increase, taking

_____

into account that the number of words is much higher. Accordingly, the type-token ratio in first paragraphs is significantly higher (21.79) than that which corresponds to the rest of the text. This imbalance in the number of distinct words in each of the two parts of the texts studied leads us to the datum that due to their introductory nature and the high concentration of new information which they contain, first sections of science popularization articles have a much more varied range of words than subsequent sections of the texts. In different words, that there are more lexical words out of the total of running words in the first section than in the rest of the text.

	Sci_TG Corpus	
	P1	**Rest**
Running words	22,293	240,254
Types	4,692	15,983
Type-token ratio	21.79	6.78

Table19. Comparative representation of first paragraphs and rest of article in *Sci_TG Corpus*

To conclude, Figure 25 summarizes the structure of the corpus and clarifies the part which has been selected for the analysis; bodies of articles that form the corpus *Sci_TG Corpus.*
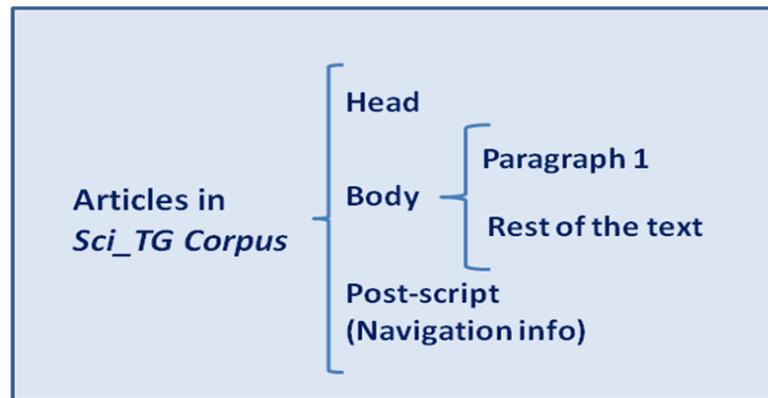
Figure 25. Structure of articles in *Sci_TG Corpus*

The first conclusion drawn from this study is that science popularization articles are shorter both in their total length as well as in the extension of the first paragraph of the articles. Editorials have an average of 100 words more per text than science popularization articles, and around 80 words if we restrict our scope to the first paragraphs only. From this data, we can state that science popularization text-type is typically shorter than other text types from the same publication, the British newspaper *The Guardian*. This first outcome on the characterization of science popularizations has cleared a path towards the comparison of different text-types within the same newspaper (hard news, soft news…), as well as across different newspapers.

The frequency of occurrence of the particle *according to* in both corpora has also revealed that it is more characteristic of science popularization articles than newspaper editorials, where there are scarce occurrences of it. We have found that there are 14 occurrences of the discourse marker in the corpus of science popularizations, a very small sample corpus, as opposed to 116 occurrences of the

discourse marker in the corpus of editorials. Although there are important differences in size between the two corpora, data have been calculated per 1 million words to allow the comparability of results.

As Ho-Dac stated (2008), first sections concentrate relevant information in texts, because they serve as discourse organizing sections. This was our departure point in finding that first paragraphs of science popularization articles concentrate a high number of occurrences of 'according to' in both corpus. First paragraphs of science popularization articles are therefore highly informative with respect to the rest of the text, although they also suffer higher variations in extension across text types.

The discourse marker *according to* has revealed itself as a typical construction in science popularization articles, characteristic of the plurality of voices and attribution of scientific facts present in the aforementioned text type. At the same time, surprisingly, more occurrences of *according to* have been found with a material rather than a personal reference. This leads us to conclude that entities are the most common collocates of the discourse marker, with special incidences of words such as 'poll', 'figure', 'publication' or 'report'. Journalists writing science popularizations have demonstrated having a tendency to relate scientific discoveries to the publications where they are popularized or to data which makes the invention reliable to the general public rather than presenting science.

The corpus studied have likewise provided us with new insights into the positions that are most commonly occupied by the discourse marker *according to*,

which has shown its rhematic preference in the corpus of science popularization articles, together with the high presence of this element in medial position, especially in the corpus of editorials.

The preliminary data we present here is a first step towards the characterization of news discourse started by Bell (1991) and focused on science popularization by authors such as Calsamiglia and López Ferrero (2003). However, they can also serve to raise awareness on the necessity of carrying out a deeper analysis in British press discourse. This could resolve some questions that remain unanswered, and provide relevant insights into the characterization of language events' reporting in science popularization discourse, the results of which are mentioned and presented in Chapter 6.

5.5. Methodology

In this study, we have followed a corpus-based approach, in which "a corpus can be used in different ways in order to validate, exemplify or build up a language theory" (Tognini-Bonelly 2001: 63). In this approach, and as Tognini-Bonelly (2001: 64) adds, corpus-based linguists adopt a confident position towards the relationship existing between theory and data, as they analyse the corpus through categories and extract the data accordingly.

As Francis 1993 points out, it is not only important that the corpus is the informant of corpus-based and corpus-driven approaches, but also that authentic

_____

data which comes from these approaches reinforce personal introspection, a

totally unreliable method when a scientific approach to language is longed for:

> The corpus is the major informant, providing the raw information we need in order
> to describe the language, and intuition is considered to be of secondary importance.
> Intuition may be useful to linguists in many ways, but for the purposes of saying
> exactly how language is used, it is notoriously unreliable. Even when people are
> describing their own language practices, there is often a difference between what
> they think they say and what they actually do say. (Francis 1993:139)

Theory is, therefore, dependent on evidence, and as Sinclair (1991: 4) states,

the examination of corpora leads to the acquisition of evidence of a quality that

was not attainable before. This evidence has to be either accepted or rejected, but

cannot be ignored, as it can give new insights into existing theories or theoretical

frameworks.

Moreover, according to Tognini-Bonelli (2001: 66), the corpus in corpus-

based approaches is useful to indicate where corrections and adjustments can be

made to the model adopted, as well as a source of quantitative evidence.

5.5.1. Computational linguistics and WordSmith Tools

There is no doubt that computers nowadays serve as a complementary tool

for a large number of the activities human beings can perform. Linguistic analysis

is not an exception and computers have assisted language researchers for some

time now. However, the function of computers in language research has shifted to

a central position with the development of new software packages, because they

enable analytical processes that would otherwise be tedious.

5.5.1.1. Quantitative and qualitative data

As Mahlbergh (2005: 43) points out, words are repeated and combined in many different ways in texts. These combinations can be observed through the use of corpus processing software such as *WordSmith Tools* that enable linguists to draw conclusions about the typical features found in the pieces of language studied. When using this kind of methodology, questions such as whether and how to combine quantitative and qualitative approaches need further consideration.

With the fast development of corpus linguistics since the 1980s, the relation of corpus linguistics with quantitative data was revealed as a very close one, in two different senses:

1.- Corpora are likely to be analysed in terms of frequencies of the words under study, creating statistical models of how language works.

2.- The setting of a quantitative model implies the basic requirement of having a corpus where it is possible to make calculations.

Corpus linguistics and quantitative linguistics are closely related in what Leech (1992: 110) considers a bidirectional connection. On one hand, the presence of a corpus entails the use of quantitative analysis because "once we have a large computer corpus, one of the most obvious things we can do with it is to derive frequencies". On the other hand, "if one wishes to set up a quantitative

_____

model of language performance, the most obvious requirement is a corpus from which quantitative counts can be made" (Leech 1992: 10).

'Qualitative research' is concerned with structures and patterns, with the way something actually is; 'quantitative research', however, focuses on how much /many there is/are of a particular characteristic or item. As noted by McEnery and Wilson, it concentrates on the quantification of properties encountered as well as on the creation of complex statistical paradigms to represent peculiarities detected:

> […] the use of quantification in corpus linguistics typically goes well beyond simple counting: many sophisticated statistical techniques are used which can both provide a mathematically rigorous analysis of often complex data – one might almost say, colloquially, to bring order out of chaos – and be used to show with some degree of certainty that differences between texts, genres, languages and so on are real ones and not simply a fluke of the sampling procedure. (McEnery and Wilson 2001: 81)

Quantification of data extracted from a corpus allows us to make generalisations and comparisons across corpora, amplifying the data obtained from a larger population as long as sampling and representativeness have been considered in corpus building. As McEnery and Wilson point out: "Quantitative analysis […] enables one to separate the wheat from the chaff: it enables one to discover which phenomena are likely to be genuine reflections of the behaviour of language or variety and which are merely chance occurrences" (McEnery and Wilson 2001: 76).

Qualitative approaches to corpus analysis have the advantage of accounting for the ambiguity present in human language (either deliberate or accidental), not coercing what could be a misleading elucidation. However, their disadvantage is

that outcomes cannot be generalised with larger samples with the same certainty with which results are extended in quantitative analyses, because specific findings are not tested to be statistically significant or not (McEnery and Wilson 2001: 76).

As inconveniences of the quantitative approach, we could state that they may lead the linguist to wrong conclusions as data should always be interpreted in view of contextual factors.

Balancing advantages and inconveniences of both approaches, we could conclude that the perspective taken depends on the desired objectives. We add, however that there is a third option that makes it possible to unite benefits of both approaches in an integrated frame of reference that widens perspectives and enriches outcomes (McEnery and Wilson 2001: 76-77). Moreover, Leech asserts that quantitative cannot be opposed to qualitative, arguing that they can be complementary (Leech 1992: 110). Biber (1988: 52) also states that quantitative analyses provide a solid empirical foundation for the findings, while qualitative analyses are required to interpret them. He concludes that both approaches have complementary strengths and weaknesses and for that reason, neither of them in isolation would give a complete description of language.

In conclusion, even if quantitative and qualitative methods still have a long way to go in corpus linguistics studies (Sinclair 1991: 3), (a progress that will produce new parameters and kinds of language data very different from the ones originally developed fifty or more years ago (e.g. Zipf 1935)), they are already being considered an essential source of empirical data for the research in the discipline, as Stubbs explains,

_____

> Corpus work depends on the interpretation of frequency and distributional data. Statistics provides ways of summarising complex numerical data, so that inferences can be drawn from them, showing which variation is likely to be significant or random, and which data sets are homogeneous or heterogeneous. (Stubbs 1993: 25)

5.5.1.2. Corpus processing software: WordSmith Tools

WordSmith Tools is a complete suite of analytical tools software for PC designed by Mike Scott , intended to assist text analysis of either a single text or a corpus, now at version 5.0. This computing tool has informally been described as a *Swiss Army Knife* for lexical analysis (Berber Sardinha 1996: 19-21) because it offers three different analytical tools to study the behaviour of words in language: Concord, WordList and KeyWords. By means of these three tools, the program is able to compute, list and retrieve specific words from texts or corpora of several million words in just a few minutes. As it is characterised by its author, *WordSmith Tools* is "an integrated suite of programs for looking at how words behave in texts" (Scott 2004-2007: 2).

This computer software has been selected to assist the analysis carried out in this study due to its capacity to quickly provide with *concordances*, lists of instances of a specific word in context that provide the linguist with information about the behaviour of a word in its environment or co-text.

In a concordance line, we select the 'node word' (the central word to be studied), and also the scope of the node word, that can reach as far as the linguists desire both on the right and left sides of the node word. As you can see in Figure 26, concordance lists can be arranged to show the node word selected in the middle, and the span indicated by the linguist on each side. Node words are presented in a different colour to facilitate the visualisation, together with information about their frequency of appearance and distribution of the word, both in the corpus and in the texts. In the screenshot below, '<RV>' is one of the tags chosen to compute reporting verbs in the corpus, and as shown, it appears as node word (blue), with a group of words in its span (green) on its right.
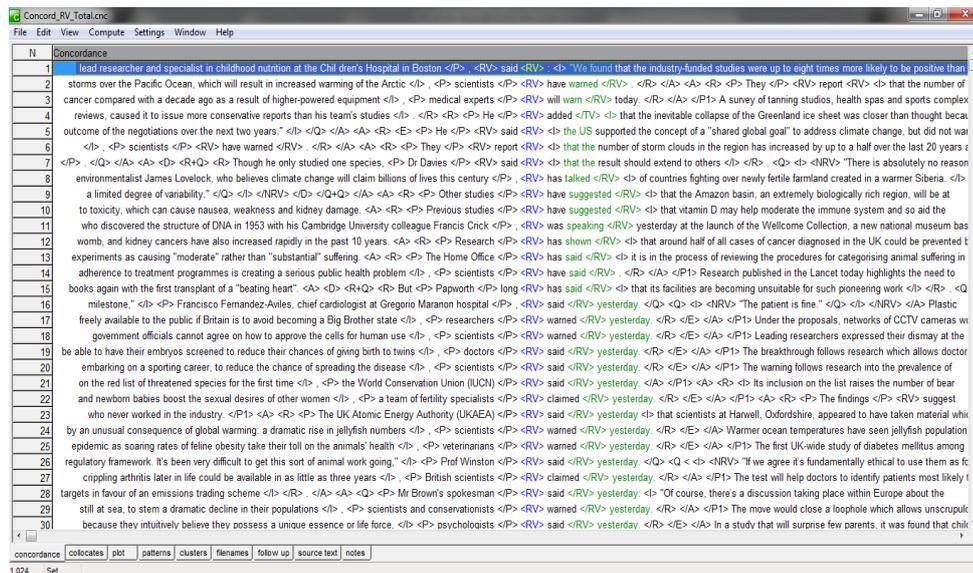
Figure 26. Example of a concordance list

Collocations shown by the concord tool in *WordSmith Tools* allow linguists to observe the tendency of certain words to occur in a specific company. Therefore, allowing them to achieve a better description of the language under

_____

study through the patterns that are most frequently repeated in the corpus. In this case, concordance lines have mainly been used to examine the type of participants that usually accompany the adverbial *according to*, and to observe the different structures that subjects and reporting verbs present in the corpus. Therefore, concordances have been very useful in the analysis of thematisation, because the visualization of node words also allows its identification in clause structure.

5.5.2. Analysis' phases

When the compilation stage of this research was accomplished, and the corpus was ready to be scrutinized, we began performing the tagging and analysis of the corpus.

The tagging phase was perhaps one of the longest in this study. We first reviewed the existing studies in the field (described in section 5.2.) and accomplished the aforementioned small-scale study which allowed us to get a more precise idea of which discourse units we wanted to analyse. We then designed tags to mark all elements of the texts that would be used as data in the analysis and they were manually introduced in the corpus.

After tagging the corpus, we carried out the analysis phase, which was accomplished in several different consecutive stages.

First, once we had obtained an overall impression of the corpus size, we moved on to study its structure and composition. A more refined search was performed to get information on the size of articles selected for the analysis,

getting data about bodies and heads of the corpus as well as about how bodies would be split into first paragraphs and the rest of paragraphs of the article for a more thorough scrutiny.

Secondly, we obtained data of the presence of explicit attribution, both in the whole corpus and in the subcorpora of first paragraphs, in the form of percentages that provide us with a general idea of facts such as how frequently do journalists use attribution in the form of projection (either quote or report) or by means of an adjunct such as *according to*.

Thirdly, we extracted concordance lines to study in deeper detail the pattern and position of attribution in the articles studied, as well as the characteristics of transitivity elements and the kind of entities which are typically introduced by means of the particle *according to*.

5.5.3. Tags

A tag is a term assigned to a piece of information in the form of metadata that classifies an item and allows it to be found again by browsing or searching. A corpus which contains tags is an 'annotated' corpus, while a corpus without metadata is known as a 'raw' corpus. In the corpus we have used for this study, a series of tags have been introduced to classify each of the elements of the text tagged.

We assigned a code to each of the articles in the corpus, to facilitate their storage and location. The code is formed by the letters *Sci_TG*, followed by

_____

number that identifies each text. Secondly, we differentiated between what we called short texts and the articles to be analysed, and tagged each text in the corpus with a label identifying its type.

Firstly, we accomplished a structural annotation process in which we assigned tags in different levels. Then, a more deep tagging was carried out and labels were assigned to the different parts of the articles to be analysed. In this tagging phase, we differentiated between heads, first paragraphs and bodies of the texts. This tagging, which identifies different parts of the text structure, has made it easier to analyse the texts in relation to the distribution of the phenomenon of attribution.

We have also designed a set of different tags that wholly cover the phenomena to be studied, and have included them in the corpus as metadata to help the analysis. The list of tags can be seen in Table 20 below. The first of the tags included in the table (<A>) labels the presence of explicit attribution, and is therefore included in the articles every time a case of attribution in a language events' report occurs. The second and third tags (<Q>, <R>) refer to quotes and reports, and label the two different simple attribution processes that have been inspected in the corpus. In addition to simple cases of explicit attribution, we have also found cases of complex attribution, which mark the presence of a case of double attribution (<D>) or cases of multiple attribution (<T>, <4> and <5>). As seen in the table below, we have not only tagged the presence of a double or multiple case of attribution, but also identified the structure that these cases have. We have also designed a series of tags (numbers 8 to 15) that reproduce the

structure of each of the cases encountered, such as <Q+Q>, <Q+R>, <R+Q+Q>, <Q+Q+Q>.

Tag number 16 (<E>) identifies cases of embedded clauses, which are also present in the text. Tags number 17 and 18 (<P> and <RV>) identify two of the elements of transitivity that have been analysed in the corpus, namely participants and processes (reporting verbs). In opposition to tag 18, tag number 19 (<NRV>) identifies those cases in which the reporting signal is not present, and no verb is explicitly present in the attribution case. Tag number 20 identifies the information that is conveyed in the language event reported (the message reported). Finally, number 21 (<AC>) identifies those cases in which *according to* appears in the article as a reporting signal.

	Tag	Meaning
1	<A>, </A>	*Attribution start, attribution end*
2	<Q>, </Q>	*Quote start, quote end*
3	<R>, </R>	*Report start, report end*
4	<D>, </D>	*Double case start, double case end*
5	<T>, </T>	*Triple case start, triple case end*
6	<4>, </4>	*Quadruple attribution start, quadruple attribution end*
7	<5>, </5>	*Quintuple attribution start, quintuple attribution end*
8	<Q+Q>, </Q+Q>	*Double case: quote and quote start, double case: quote and quote end*
9	<R+R>, </R+R>	*Double case: Report and report start, double case: report and report end*
10	<R+Q>, </R+Q>	*Double case: report and quote start, double case quote and report end*
11	<Q+R>, </Q+R>	*Double case: quote and report start, double case: quote and report end*
12	<R+R+R>, </R+R+R>	*Triple case: report, report and report start, triple case: report, report and report end*
13	<R+Q+Q>, </R+Q+Q>	*Triple case: report, quote and quote start, triple case: report, quote and quote end*
14	<R+R+Q>, </R+R+Q>	*Triple case: report, report and quote start, triple case: report, report and quote end*
15	<Q+Q+Q>, </Q+Q+Q>,	*Triple case: quote, quote and quote start, triple case: quote, quote and quote end*
16	<E>, </E>	*Embedded clause*
17	<P>, </P>	*Participant*
18	<RV>, </RV>	*Reporting verb*
19	<NRV>, </NRV>	*No reporting verb*
20	<I>, </I>	*Information*
21	<AC>, </AC>	*According to*

Table 20. Tags used in the annotation of *Sci_TG Corpus*

An instance of the tagging of a double case in the corpus is provided below in Example 51, where we can see how tags are introduced in each of the cases identified in a complementary way and occurring from the most general to the most specific one:

(51) *<A> <D> <Q+R> <Q> <I>* "This large number of grinding stones and other tools used to crush and grind ore shows that the site was a centre for organised gold production," *</I> <RV>* said *</RV> <P>* Geoff Emberling, co-leader of the expedition and director of the Oriental Institute Museum at the University of Chicago. *</P> </Q> <R> <I>* Fragments of gold ore may have been found among the gravel deposits in nearby wadis, or dry river beds, and taken to the site for processing *</I>, <P>* he *</P> <RV>* added *</RV> </R> </D> </Q+R> </A>*

5.5.4. Annotation and analysis layers

The analysis has been carried out at different levels where different layers have been assigned to the consecutive phases of the analysis carried out. These layers are mutually dependent and thus were analysed in consequent phases to base new data on existing knowledge.

Before starting the analysis using the layers designed for this study, we examined three different aspects that helped us identify the type of text and the part of the text (head-body) we were dealing with, following the structure shown in Figures 27 and 28 respectively.

_____
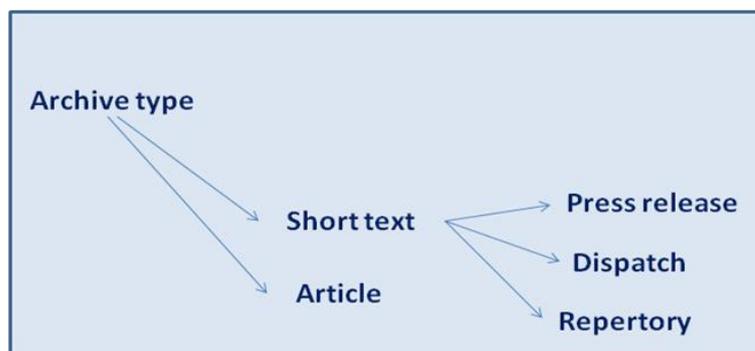
Figure 27. Archive type analysis in *Sci_TG Corpus*
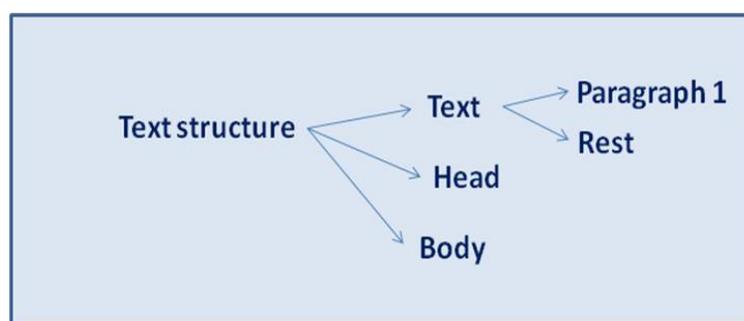
Figure 28. Text structure analysis in *Sci_TG Corpus*

Layer 1 deals with the identification of cases of attribution in the articles selected. As we have stated in the objectives of this study, one of our goals is to identify how frequent attribution is in science popularization texts of the British press. With the analysis of this layer, we obtain data about the number of occurrences of explicit attribution of language events in the corpus (Figure 29). This second layer is very relevant for the analysis, as it implies the identification of all instances that will be investigated in subsequent layers.

Figure 29. Layer 1. Language event report analysis in *Sci_TG Corpus*

Layer 2, on the other hand digs deeper into the anatomy of language event reports, as we can see in Figure 30 below. It sets its basis on the transitivity system which "specifies the different types of process that are recognised in the language, and the structures by which they are expressed" (Halliday 1985: 101). This process is said to be formed by three components:

- The process itself

- Participants in the process

- Circumstances associated with the process

Following Halliday's scheme, we have analysed the process in each of the attribution occurrences identified in Layer 1. The process itself has been identified with the reporting verb, which encompasses a great part of the meaning of the report or quote attributed to an external source. Participants have been identified as people or entities that are presented as authors of the information reported. Science popularization article journalists make participants responsible for the discoveries reported, and thus they may detach themselves from the information

_____

conveyed in the text. Participants are often reputed characters of scientific and social areas, well known by readers, who are portrayed in very different degrees of detail, as we will see in Chapter 6.
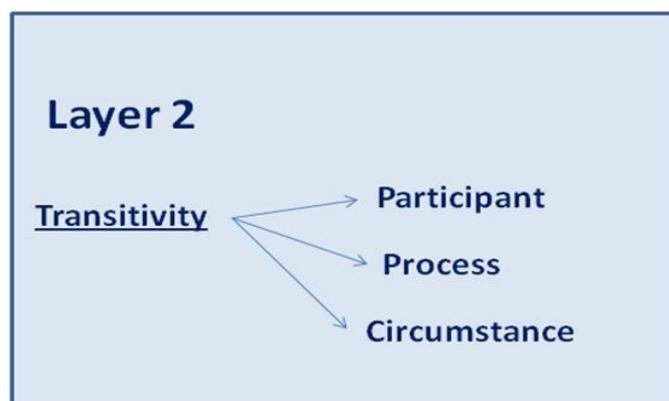
Figure 30. Layer 2. Elements of Transitivity in *Sci_TG Corpus*

Layer 3 considers the different logico-semantic relations that are established in clause combination. More specifically, we have focused on the analysis of the mode in which clauses of language event reports are projected. Two initial possibilities have been considered. On one hand, that of creating interdependent relations between clauses; relations that can be either paratactical or hypotactical and materialise in the texts as quotes and reports, or embedded; a relation in which projected clauses are 'rank-sifted' (Halliday and Mathiessen 2004: 467) to function as a Qualifier of a nominal group.

_____
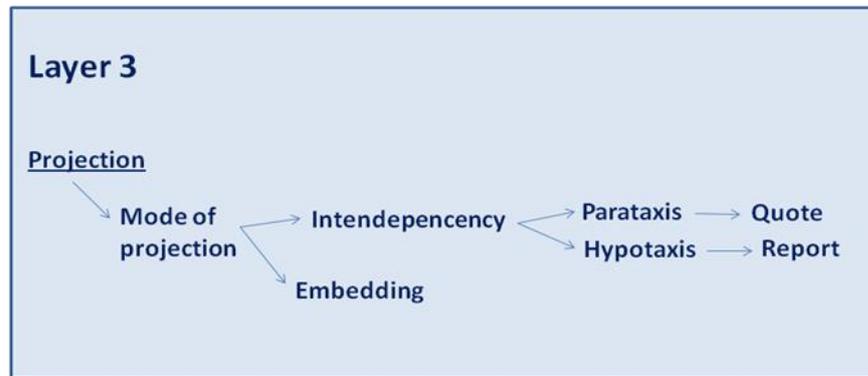
Figure 31. Layer 3. Types of projection in *Sci_TG Corpus*

Besides simple attribution cases, those of multiple attribution have also been identified while tagging the corpus. In order to identify, classify and analyse multiple cases spotted in the corpus, we have designed Layer 4, which takes into account double, triple, quadruple and quintuple cases, as seen in Figure 32 below.
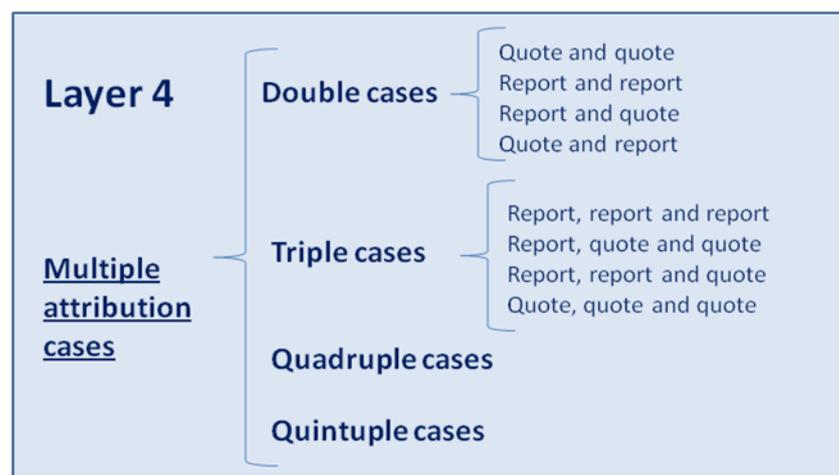
Figure 32. Layer 4. Types of multiple attribution in *Sci_TG Corpus*

_____

Layer 5 takes into account a frequent procedure of attributing language event reports in science popularization discourse, the epistemic stance adverbials. As Biber, Conrad and Finegan (1999: 871) describe, *according to* is an epistemic stance adverbial which signals the source of information in texts. That is the reason why we have carried out an analysis of the different patterns that *according to* shows in our corpus of science popularization articles, as seen in Figure 33 below.
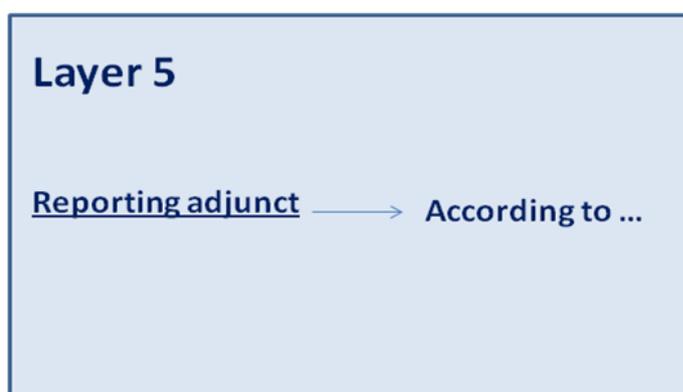
Figure 33. Layer 5. Epistemic stance adverbial *according to* in *Sci_TG Corpus*

Finally, Layer 6 considers the alternative possibility of not explicitly presenting a reporting verb in language events reports, and thus relying on readers to retrieve the process from language events previously reported in the text. Reporting verbs do not only help us to identify that a language event has taken place somewhere, but also to describe the way in which that event occurred. There is a clear semantic difference between 'argue' and 'comment' or between 'explain' and 'complain'. The absence of an explicit reporting verb also implies a lack of information, as readers do not have data on the journalist's perspective on

_____

how the language event happened. This is also a relevant pattern that we wanted

to take into account in the analysis of *Sci_TG Corpus* (see Figure 34 below).
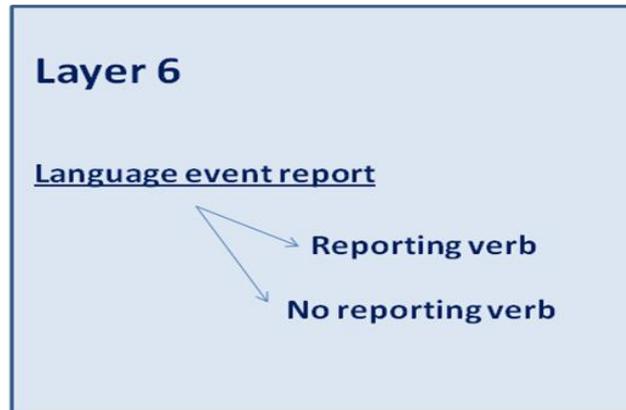
Figure 34. Layer 6. Use of reporting verb in language event reports in *Sci_TG Corpus*

Layers have been created to account for the different aspects of the

phenomenon analysed. As a result, tags and layers have been related in a Table 21

below, creating a single net under which the analysis of the corpus has been

carried out.

	Tag	Layer
1	<A>, </A>	*Layer 1*
2	<Q>, </Q>	*Layer 3*
3	<R>, </R>	*Layer 3*
4	<D>, </D>	*Layer 4*
5	<T>, </T>	*Layer 4*
6	<4>, </4>	*Layer 4*
7	<5>, </5>	*Layer 4*
8	<Q+Q>, </Q+Q>	*Layer 4*
9	<R+R>, </R+R>	*Layer 4*
10	<R+Q>, </R+Q>	*Layer 4*
11	<Q+R>, </Q+R>	*Layer 4*
12	<R+R+R>, </R+R+R>	*Layer 4*
13	<R+Q+Q>, </R+Q+Q>	*Layer 4*
14	<R+R+Q>, </R+R+Q>	*Layer 4*
15	<Q+Q+Q>, </Q+Q+Q>	*Layer 4*
16	<E>, </E>	*Layer 3*
17	<I>, </I>	*Layer 2*
18	<P>, </P>	*Layer 2*
19	<RV>, </RV>	*Layer 2,6*
20	<NRV>, </NRV>	*Layer 6*
21	<AC>, </AC>	*Layer 5*

Table 21. Relation between tags and layers in *Sci_TG Corpus*

_____

5.6. Final Remarks

In this chapter, we have set the basis for the analysis that has been conducted on *Sci_TG Corpus.* First, we have established the departure point for this doctoral dissertation by reviewing the most recent studies that have been carried out in the field, which provided us with relevant results that have been taken into account in the development of this doctoral dissertation. Secondly, we have formulated a series of research questions and hypotheses that have to be tested with this study, and which are related to the objectives that were set as goals of this doctoral dissertation. Thirdly, we have provided data about the object of study of this analysis: *Sci_TG Corpus*, describing its structure and compilation process. Fourthly, we have reviewed the methodology followed in this study, focusing our attention on the retrieving of quantitative and qualitative data and the computing tool that has assisted the analysis of the corpus. We then described the stages followed in the analysis and the tags that have been included in the corpus as metadata for the analysis. Finally, the annotation stage of this study has been explained, placing a special emphasis on the depiction of the different layers of analysis that have been used in this study.

Chapter 6

Analysis and results

Chapter 6

Analysis and results

6.1. Introduction

6.2. Attribution in science popularizations of *The Guardian*

    6.2.1. Simple reporting

        6.2.1.1. Rephrasing

        6.2.1.2. Quoting

    6.2.2. Complex reporting

        6.2.2.1. Double reporting

        6.2.2.2. Multiple reporting

    6.2.3. Adverbial reporting

        6.2.3.1. Presence of *according to*

        6.2.3.2. *According to* and transitivity

        6.2.3.3. *According to* and thematic structure

    6.2.4. Reporting verbs

        6.2.4.1. Tense in reporting verbs

        6.2.4.2. Classes of reporting verbs

One special feature of human language
is that it can talk about itself.
No other communication system has this power.
Musical themes can reach each other,
parrots can mimic other noises,
but they cannot refer to each other.
John Sinclair

6.1. Introduction

As John Sinclair states (1994: VI), the capacity of language to report, not only what people do but also what people say, immensely increases the power of communication and its flexibility. Many have attempted to create models to reconstruct what Thompson (1996a) calls 'language events' into reports, but decisions on how these reports are expressed ultimately depend on the situation of the original talk. We should bear in mind, as Sinclair explains, that it is difficult in human communication to achieve a complete and successful reproduction of the words of others, and that some change, of minor or greater degree and with no given rules, is always made on the discourse:

233

> The exact reproduction of previous talk is hardly ever attempted, except in telling jokes that depend on accents and voice quality. Mimicry in humans is usually considered offensive. In reporting there is always some change made – by adapting the talk to the present situation, by shortening it, by adding descriptive or evaluative phrases, and so on. No rules are given in this process, no more than there are rules that relate physical objects to the descriptions that are made of them in language. Instead, the author offers a rich variety of ways of making reporting effective. (Sinclair 1994: IV)

It is precisely the range of resources available in English to encode language events' reports that we want to portray in this chapter. We will give an account of the different patterns that have been observed within the phenomenon of attribution in *Sci_TG Corpus,* paying special attention to first paragraphs. These are highly informative sections of science popularization articles where we have found patterns that will help characterize this newspaper text-type.

6.2. Attribution in science popularizations of The Guardian

As a departure point for this analysis, we have analysed the number of cases of attribution that have been identified in the corpus. It is relevant to remark that, for this study, we have only considered cases in which the attribution is part of a language report. We have left aside the cases in which the reporter attributes the scientists or the experts discoveries and data unrelated to language events, as in Examples (51) and (52) below:

(51) *A huge wild pig that can grow to more than 1.2 meters (four feet) long and which inhabits the central Amazonian rainforest has been discovered by scientists.* (Sci_TG0735)

(52) *Not only are these objects more than simply dirty snowballs, as had been previously thought, scientists found materials in them that suggest they could have kickstarted life on Earth.* (Sci_TG0605)

Therefore, following Thompson (1994), we have exclusively paid attention to those cases in which there is an explicit reporting signal in the form of reporting verb, punctuation (colon, inverted commas) or the epistemic stance adverbial *according to*. As a result, we have found that there are 2,377 cases of explicit attribution in language events' reports of the corpus, as seen in Table 22 below:

	Running words	*Number of occurrences*	*Occurrences per 1,000 words*
Sci_TG Corpus	262,547	2,377	9.05

Table 22. Frequency of attribution cases in *Sci_TG Corpus*

There are no significant differences between first paragraphs and the rest of the text, where the number of attribution cases place around 9 every 1,000 words. Attribution is thus as frequent in first paragraphs as it is in the rest of the text, where journalists keep ascribing the words reported to external sources of attribution.

	P1	*Rest*	*Total*
Sci_TG Corpus	191 (8.04%)	2,186 (91.96%)	2,377

Table 23. Number of cases of attribution in *Sci_TG Corpus*

6.2.1. Simple reporting

As Thompson states (1994), there are three ways to refer to sequences of language uttered by somebody else. The first of the options is to repeat the sequence of while remaining loyal to the original words uttered. Little to no alteration is made to the original message when using this first option, as seen in example 53 below:

(53) *"That's a disaster and it is absolutely not in the public interest," said Jenny Hewison, a clinical researcher at Leeds University*. (Sci_TG06175)

The second option entails a rephrasing of words to adapt them to the reader and make them fit into the current narration. This rephrasing of language, carried out by the reporter, causes a change in the words originally uttered, which are no longer the speaker's or writer's, but the reporter's words, as shown in Example 54 below. This second type of reporting has traditionally been called indirect speech, thus establishing a dichotomy between direct and indirect speech that has been used as a rule in English teaching. Conversely, we will use the term 'rephrase' to refer to indirect speech in this chapter and we will call the traditional direct speech 'quotation' or 'quote'.

(54) *Dr Zhang   said that on his scaffolds cells survived longer and differentiated better, without the need for additional chemicals to encourage growth.* (Sci_TG0602)

The third option considered by Thompson takes into account cases where a language event is reported without using the original wording but with a high degree of integration in the journalist's narration. In these cases, there is a lexicalization of the reporting signal, which can appear as a reporting noun, as in Examples 55 and 56 below.

(55) *The government last week ordered an inquiry into claims that tissues and organs from nuclear workers who died were removed without the knowledge of their families over 30 years at Sellafield, Cumbria.* (Sci_TG07176)

(56) *Mr Blair's comments were made in a speech in Oxford, the fourth talk in a series on securing Britain's future. His main message was to encourage young people to engage with science.* (Sci_TG0659)

We have not considered this type of report in the study carried out, as it presents different features from the previous two and constitutes what Halliday and Mathiessen (2004: 395) call 'expansion' instead of a 'projection' structure, formed by the reporting structures studied. We will thus deal with rephrasing and quoting in subsequent sections, the two most frequent phenomena observed to introduce a language event in science popularization articles in *Sci_TG Corpus.*

The presence and distribution of each type of attribution in the corpus deserves a deeper analysis. On the one hand, as seen in Table 24 below, the total number of occurrences of each type is quite similar, because there is only a difference of 125 occurrences in a total of 2,377 cases of attribution, so we can

say that their overall presence in the corpus is comparable. On the other hand, taking into account the distribution of each structure in the corpus, we find more relevant differences between the location of these structures in first paragraphs or subsequent sections.

Sci_TG Corpus	P1	Rest	Total
Rephrasing	101 (14.11%)	615 (85.89%)	716 (100%)
Quoting	4 (0.48%)	837 (99.52%)	841 (100%)

Table 24. Frequency and distribution of rephrasing and quoting in *Sci_TG Corpus*

Even if none of them tend to be concentrated in first paragraphs of the science popularization articles studied, it is remarkable to note that in the case of quotes, their presence in initial sections is almost nonexistent, and they are consistently concentrated in later paragraphs. In the next two sections, we will pay special attention to each of these two types and the distribution they present in the corpus.

6.2.1.1. Rephrasing

Rephrasing is one of the most common phenomena found in the corpus. In *Sci_TG Corpus* this procedure has been recurrently chosen to introduce a language event in the text. Using rephrases, journalists do not aim at giving readers access to the voice of the scientists to directly convey their message. They re-write the original words of the scientists in such a way that they fit their

purposes, informing the lay reader by means of the reporting clause that introduces the language event as seen in Example 57 below:

(57) *Dr Bateson said that although the eyes were not real, they seemed to make people behave more honestly.* (Sci_TG06217)

As already mentioned, rephrases appear to be more frequent in the rest of the text rather that in first paragraphs. As highlighted in Chapter 5, existing differences are derived not only from the fact that first paragraphs are much shorter than the rest of the text, where they have a notable difference in size; but also by the fact that rephrases seem to be primed to appear in later parts of the text, rather than in first paragraphs (see Table 24 above).

To facilitate the comparability of the data obtained, we have taken a step forward and analysed how frequent this structure is per 100 cases of attribution. This standard measure enables us to find out if this type of attribution shows an uneven distribution in *Sci_TG Corpus*. As a result, we have found a total of 716 occurrences of rephrases in the corpus, which means that we could expect to find 301.21 occurrences of this structure every 100 cases of attribution in *Sci_TG Corpus*. If we look deeper into the two different parts that we have established for the analysis of the texts, we find that in initial parts of the text the number of occurrences of rephrases is 52.879 (per 100 occurrences of attribution), and 28.133 cases if the analysis is carried out in subsequent parts of the texts (see Table 25 below). As figures reveal, there is a difference in the concentration of rephrases along the text. The number of rephrases in first paragraphs is higher than in the rest of the text, so we can say that there is a tendency in *Sci_TG*

Corpus to appear more frequently in initial sections. This datum can be explained by the fact that the journalist uses the first paragraph as an introduction where he/she sets the contexts of the texts, thus starting the narration with his/her own words before giving voice to other sources of attribution.

		Cases of attribution	Occurrences of rephrases	Occurrences per 100 attribution occurrences
Sci_TG Corpus	*P1*	191	101	52.879
	Rest	2,186	615	28.133
	Total	2,377	716	301.21

Table 25. Significance of rephrases' occurrences

It is also relevant to note that there is no homogeneity in the way rephrases observed in the corpus are constructed. Among the occurrences of reports observed in the texts, we can also differentiate various patterns of construction according to the order in which the reporting and reported clauses are presented to the reader. On the one hand, we find occurrences where the reporting clause stands in thematic position. In other words, the reporter has chosen to give more emphasis and importance to the participant, placing him/her in thematic position, rather than to the message itself. When the rephrase appears in thematic position it is usually constructed as a *that-clause*, either with the relative explicit or omitted, as can be seen in Examples 58 and 59 below:

(58) *The scientists said **that** the results, published yesterday in the Proceedings of the National Academy of Scientists, are explained by the way in which the sun's*

rays are absorbed or reflected by different parts of the world (Sci_TG07192*)* [Emphasis added]

(59) *Prof Marsh said the disc was likely to be the remains of a passing asteroid.* (Sci_TG0606)

On the other hand, we have also encountered occurrences of reporting clauses after the reported clause, which seems to be very common in the language of science popularization articles from the British newspaper *The Guardian* which has been analysed in this study, as we shall see later. The reporting clause appears in rhematic position and is normally separated from the reported clause by means of a punctuation mark, usually a comma, which makes this reporting clause stand in inverted position and add information establishing a dependence on the reported clause. As can be seen in Example 60 below, the information on the actual damage caused by global warming if trees are planted outside the tropic is placed at the beginning of the sentence, as a re-elaboration of the actual words uttered by scientists. It is not before rhematic position that we find the reporting signal, the verb 'believe' and the attributee, 'scientists':

(60) *It may have become the penance of choice for the environmentally conscious individual, but planting trees to offset carbon emissions could contribute to global warming if they are planted outside the tropics, **scientists believe*** (Sci_TG07192) [Emphasis added]

Additionally, we have found that not all cases of rephrases in the corpus are either thematic or rhematic. There are also occurrences in which the reporting clause appears in mid-position. This is to provide extra information between

commas, that breaks the rhythm of the message, while recovering the source of attribution from its previous appearance. This helps the reader remember the person who produced the original language event. An instance of a reporting clause in mid-position can be seen in Example 61 below, where the reporter has chosen to use the reporting clause as a continuative device, making the source of attribution explicit in the middle of the reported clause:

(61) *Bringing it back to Earth,* **she said**, *would take up valuable room on return flights that could be used for other equipment or samples from experiments* (Sci_TG06424) [Emphasis added]

These three reporting structures coincide with those already described by Thompson (1994: 9-10). Therefore, we can say that rephrases found in *Sci_TG Corpus* are similar to the rephrases found in other types of discourse, but the question here is whether a specific pattern can be observed in this text type.

To get more detailed information on the frequency of the appearance of each of these reporting structures, we have conducted a deeper analysis on first paragraphs of *Sci_TG Corpus*. In this analysis, we have studied the 101 occurrences of rephrases in initial sections of the *Corpus,* and found 28 cases of reporting clauses in thematic position. They occupy the initial part of the rephrase structure, and are receiving more importance by the reporter, as seen in the example below:

(62) *A wildlife expert has told how he convinced an African rebel army, which has fought a bloody struggle with the Ugandan government for nearly two decades, to sign up to a conservation project to save one of the world's rarest animals.* (Sci_TG06118)

We have also found that out of the 101 occurrences, only 9 occur in medial position, as shown in the example below, which reinforces its uncommon character already pointed out by Thompson (1994: 11).

(63) *Schizophrenia should be abolished as a concept, a group of mental health experts say today, because it is a catch-all term which does not define a specific illness and carries a stigma that destroys people's lives.*(Sci_TG0691)

Conversely, we have found that rhematic position is the most common location of rephrases in first paragraphs of science popularization articles studied. There awere 64 occurrences in which the reporting clause was found at the end of the reporting structure, separated by a comma, as previously described. Contrary to what Thompson asserted about this construction being less frequent than others in written discourse (1994: 11), it seems to be primed to appear in science popularization articles from the publication studied.

(64) *In a single generation, once thriving populations of deep sea fish have been driven to the brink of extinction by expanding fisheries, researchers say today.* (Sci_TG06449)

6.2.1.2. Quoting

Data obtained from the corpus have demonstrated that, rephrases are a common procedure to enclose explicit attribution in science popularization texts. However, quotes are used in a higher degree in order to to give access to

authorized voices in the text. Therefore, *Sci_TG Corpus* reporters have recurrently chosen this procedure to introduce a language event in the text. By using quotes, they give access to the voice of the scientists and directly convey their message, without re-writing the words written or uttered by the scientists. They construct a structure that makes the message fit into the reporting clause that introduces the language event, as seen in Example 65 below:

(65) *"The Earth would be extremely singed - the oceans will boil dry and we'll be left with a rather unpleasant rock," said Prof Marsh.* (Sci_TG0606)

As we can see in Table 26 below, quotes show a clear tendency to appear more frequently in the rest of the text rather than in first paragraphs. Existing differences, as also seen in the case of rephrases, are derived from the fact that a there is a notable difference in size between first paragraphs and the rest of the text. They are also derived from the fact that quotes seem to be more intensely primed to appear in later parts of the text, rather than in first paragraphs. This tendency can be the result of the journalist's desire to provide the reader with a departure point for the text, and therefore does not include original instances of language events, but rephrases the information in the initial sections of the text. In subsequent sections, the journalist will be able to give further details, to develop the idea that has been put forward in the first paragraph, and to also provide readers with more specific information that enables the inclusion of particular examples and more precise data.

		Cases of attribution	Occurrences of quotes	Occurrences per 100 attribution cases
Sci_TG Corpus	P1	191	4 (0.48%)	2.093
	Rest	2,186	837 (99.52%)	38.288
	Total	2,377	841 (100%)	35.381

Table 26. Significance of occurrences of quotes in *Sci_TG Corpus*

Taking into account the overall figures of *Sci_TG Corpus,* we found that there are 2.093 cases of quotes in initial sections of the articles per 100 attribution cases, and 38.288 cases if the analysis is carried out in subsequent parts of the texts. As we can extract from the variation in the figures, there is a great difference in the tendency for quotes to appear in different sections of the text. First paragraphs of the texts contain a small percentage of the occurrences of quotes, while they are located, in a much higher number, in later paragraphs of the texts.

In addition, it is relevant to note that in this case there is no homogeneity between quotes analysed in first paragraphs of the corpus. Among the occurrences of quotes observed in the texts, we can, distinguish various patterns of construction according to the order in which reporting and reported clauses are presented to the reader. However, in this case, the presence of quotes in first paragraphs is so low that no relevant patterns can be extracted. We have found three occurrences where the reporting clause stands in thematic position, and the first element in the clause provides readers with the identification of the attributtee

who uttered the words and is given access to the text, as seen in Example 66 below:

(66) *Producer Joe Oppenheimer: "People do love meerkats but there is far more going on than them just being cute and furry and standing on their hind legs."* (Sci_TG0666)

Conversely, we have found only one case in which the message is emphasized over the person speaking, and thus the exact words uttered are presented in thematic position, as seen in Example 67 below:

(67) *"The age thing?" says Mark, a middle-aged father who spent his 30s and 40s vaguely wanting children but working and travelling and developing complicated interests instead.* (Sci_TG06289)

Two cases that we have also observed in the corpus are those in which the reporting signal is not so easily identifiable, and when 'free speech' is used. This can be done either by not making any reporting verb explicit, as in Example 68 below, where the reader has to infer that it is a videogamer who utters the message conveyed; or by omitting punctuation marks that open the quotation. As we can see in Example 69 below, the only signal marking that a quotation comes after the reporting clause is the colon which separates the reporting clause from the message.

(68) *It is the ultimate excuse for every video-gamer accused of spending too long hunched over a console: "I'm not addicted. I'm just honing my surgical skills."* (Sci_TG07237)

(69) *Peter Grunert, deputy editor of Top Gear magazine, said: Apart from looking slightly freakish, it's slow, in fact it's almost an incredibly lazy alternative to walking.* (Sci_TG07159)

On the other hand, *Sci_TG Corpus* also contains cases in which the language event report contains more than one source of attribution, sometimes referring to the same participant. In Example 70 below, two reporting verbs are associated with the same language event, and thus the message is introduced by a reporting verb (*described*), and then followed by a second one (*said*) that stands in rhematic position and has a continuative function:

(70) *The environment secretary, David Miliband, yesterday **described** the encyclopedia as a "unique endeavour. What could be more appropriate for a world that is more and more connected, where issues of climate change, issues of biodiversity, join us together over national and natural boundaries," he **said*** (Sci_TG07167) [Emphasis added]

Taking into account the thematic development of one of the texts studied, as seen in Example 71 below, we find that the first paragraph not only sets the tone of the articles, but acts as an introduction and shows the first attribution case, which uses the adverbial *according to* as the reporting signal that identifies the source of attribution: a study by ecologists. In subsequent sections of the text, the journalist provides the readers with more detailed instances of language reports and more specific information about the ecologists. Ken Caldeira and his colleague Govindasamy Bala are marked as attribution sources, occupying both rematic and thematic positions along the text. In the tenth paragraph, conversely, a new attribution source is mentioned in the text, John Coequyt, of Greenpeace

USA, but in paragraph eleven, the journalist recuperates Prof. Caldeira as the

attribution source.

(71) Sci_TG0612
15 December 2006, The Guardian, p.18
Science: Planting trees to save planet is pointless, say ecologists / Home News
Jha, Alok. San Francisco

Planting trees to combat climate change is a waste of time, according to a **study by ecologists** who say that most forests do not have any overall effect on global temperature, while those furthest from the equator could actually be making global warming worse.

"The idea that you can go out and plant a tree and help reverse global warming is an appealing, feel-good thing," said **Ken Caldeira of the global ecology department at the Carnegie Institution of Washington in Stanford, California, a co-author of the study.** "To plant forests to mitigate climate change outside of the tropics is a waste of time."

The carbon dioxide used by trees for photosynthesis helps cool the Earth by reducing greenhouse gases in the atmosphere. But forests also trap heat from the sunlight they absorb.

Professor Caldeira and his colleague Govindasamy Bala, of the Lawrence Livermore National Laboratory, also in California, said that outside a thin band around the equator, forests trap more heat than they help to get rid of by reducing CO 2.

Their research comes in the wake of criticism from scientists of forestry schemes to offset carbon emissions, which they argue let consumers carry on polluting with a clear conscience. The schemes are big business; within three years, the market is expected to reach pounds 300m.

Dr Bala's study, the first simulation to link carbon dioxide and the heat- absorbing effect of trees, found forests had different effects on global warming depending on latitude. He is to present his findings at the American Geophysical Union's meeting in San Francisco today.

"North of 20 degrees [latitude] forests had a direct warming influence that more or less counterbalanced the cooling effect of carbon removal from the atmosphere," said **Prof Caldeira**. Past 50 degrees, forests warmed the Earth by an average of 0.8C. But in the tropics forests helped cool the planet by an average of 0.7C.

Dr Bala explained that forest canopies, because they are relatively dark, absorbed most of the sun's rays heating falling on them. Grassland or snowfields, however, reflected more sun, keeping temperatures lower. Planting trees above 50 degrees latitude, such as in Siberia, could cover tundras normally blanketed in heat- reflecting snow.

In the tropical regions, though, water evaporating from trees increased cloudiness, which helped keep the planet cool.

John Coequyt, of Greenpeace USA, said: "We have always come down on limiting the credits you give to countries and companies that use forestry policy to mitigate climate change."

Prof Caldeira said planting trees was a diversion, letting consumers pollute more. **He** said it would be better to transform the way energy was derived and used, for instance through investment in renewable and carbon-free electricity generation.

© Guardian News & Media Ltd 2006

Moreover, corpus analysis has revealed other quote constructions which are not as canonical, but also present in science popularization articles in the newspaper *The Guardian* and, therefore, deserve our attention. We have found cases in which the quotation is interrupted by the reporting clause, which stands in the middle of the reported clause, the message. This type of construction is described by Thompson (1994: 27-28) as 'discontinuous report', asserting that it is not a report neither a quote strictly speaking, but something that is closely related to them. In canonical quotes, reporting and reported clauses are grammatically linked to form a unit. However, in the case of discontinuous reports, the reporting is carried out through a "grammatically separate structure which could stand alone" (Thompson 1994:27), thus making the report a discursive structure of a higher rank than the sentence, as seen in Example 72 below:

(72) *"The danger is that people will take the message from this that red meat is unsafe," said the commission's spokesman, Guy Attenborough.* [Q1] *"But they're talking about giving people the equivalent of two 8oz steaks a day, seven days a week. Anyone whose diet is that unbalanced is going to have problems."* [Q2] (Sci_TG06425)

This structure will receive further attention in the next section of this chapter, where more complex structures found in the corpus are analysed in detail.

6.2.2. Complex reporting

All cases that have been detailed so far can be said to be cases of 'simple reporting'. In simple reporting there is a specific reporting signal for each message reported. However, this is not the only method observed in the corpus to encode external voices in the text.

In addition to simple reporting, and as García Riaza and others (2011:4) have found, the corpus presents cases in which the reporting of the language events is carried out in a sequence in which the reported message presents different degrees of embedding in the journalist's narration. Sometimes they present both a quotation and a rephrasing, as is the case in Example 73 below:

(73) *Dr Bodani said he was, "pleased and very pleasantly surprised" to have won. "I wanted to make electricity sort of come alive. I was struck by this eternal power, this force that was created over 13 billion years ago when the universe was young."* (Sci_TG 06283)

This kind of language event reporting, as we will shortly see, is a common feature of science popularization articles from the British newspaper studied. It provides the research in this area with new insights into the existing conception of language reporting, which is now revealed as a much more complex process.

'Complex reports' observed in the Corpus differ from Thompson's discontinuous reports (1994: 27-28) in the sense that they imply a formal discontinuity in the quote or report structure, as in Example 74 below.

(74) *"The RSPB is fundamentally supportive of renewable energy," Mr. Stowe said, "but risking irreplaceable wildlife sites for the sake of energy generation is not a sustainable option."* (Sci_TG06299)

Thompson describes 'discontinuous reports' as being neither direct nor indirect report structures in which a reporting clause and a reported clause make up one sentence. This type of reports can be introduced by a reporting signal which can be carried out by a separate structure, able to stand on its own. Similarly, complex reports detected in the corpus show that reports, quotes or a mixture of both are combined to form single language events' rephrases. Moreover, Thompson (1994:28) states that the message in discontinuous reports may be a quote introduced by a reporting signal, such as a verb, as seen in Example 75 below where we see a dialogic structure:

(75) *'It's appalling.' Only the director agreed: 'Yes, we know.'* (Thompson 1994: 28)

That-clauses can also appear in discontinuous reports. The reporting verb (if present) is followed by an empty noun group which is the object of the verb but does not carry the message, as can be seen in Example 76 below:

(76) *She realized <u>something</u>: that she could not put Mr. Turner off.* (Thompson 1994: 28) [Emphasis original]

Reporting nouns can be used as reporting signals with quotes and *that-clauses*, as in Example 77 below, as well as a particular type within journalism,

used in the reporting of quotes in interviews to present the interviewee's answer as a direct quote (see Example 78 below).

(77) *Over two decades ago he posed a question: 'How does it feel to be on your own?'* (Thompson 1994: 28)

(78) *Did playing a tyrant make him act like a tyrant? 'No, but all of those things I just put off to the side, really'* (Thompson 1994: 28)

6.2.2.1. Double reporting

The most common type of this kind of reporting observed is that of 'double cases' (García Riaza et al. 2011:4) which are defined as those "in which the message is constructed in two ways, reformulating the information already given". This is apparent in Example 79 below, where we find a rephrase and then a quote which seems to provide the words of Dr. Baines:

(79) *Dr Baines said that storms could help scientists understand what goes on deeper in planets. "When you have storms they tend to dredge up materials from deep down below, and so if you want to see what's in the deep part of a planet then you can look in a storm system and see tracers of the material deep down."* (Sci_TG0647)

A good number of rephrases and quotes, described in the previous section as single reports, also combine to form strings of reporting in which a single language event is transmitted to the reader. Punctuation plays a key role in the identification of these cases, as reports which belong to a double case are included in the same paragraph, constituting a reporting unit of suprasentential nature:

(80) *"It looks like a hurricane, but it doesn't behave like a hurricane," said Andrew Ingersoll, a member of Cassini's imaging team at the California Institute of Technology. "Whatever it is, we are going to focus on the eye of this storm and find out why it's there."* (Sci_TG0647)

In the analysis performed on the *Sci_TG Corpus*, we have found that double cases are not primed to appear in first sections of science popularization articles. As we can see in Table 27 below, only two cases of double reporting were detected. Conversely, in subsequent section of the texts, double cases are quite common. We found a total of 547 which will be analysed in greater detail in order to identify recurrent patterns.

		Double cases	*Occurrences per 100 attribution cases*
Sci_TG Corpus	*P1*	2	1.047
	Rest	545	24.931
	Total	547	23.012

Table 27. Double cases in *Sci_TG Corpus*: occurrences and significance

Although figures extracted from the analysis of the corpus lead us to believe that double cases are present in the corpus to a large degree, we have to bear in mind that they are a small part of the whole amount of words found in science popularization texts. As with attribution, the fact that double cases seem to be characteristic of science popularization articles of the British press does not imply that they have to be massively present in the texts. As seen in Table 28 below, the significance of double cases with respect to the total number of words is quite low

253

but, again, the presence of double cases in the rest of the text is noticeable and higher than in first paragraphs. In addition, in first paragraphs, only 2 double cases have been found, which represents 1% of the cases of attribution in first paragraphs (191). Data in *Sci_TG Corpus* show that double cases are found less than 25 times per every hundred occurrences of attribution observed.

From a formal perspective, complex reports consist of two or more elements where the message is conveyed, which can be expressed either by quoting, rephrasing or both.

We have found that there are two different types of double report in the corpus. There are quotes and rephrases combined with each other in no particular order, and, then there is either a quote or a rephrase combined with a different element: either *according to* or an embedded clause.

Regarding the first type, four distinct combinations occur in the corpus with a very heterogeneous distribution, as presented in Table 28 and Figure 35 below.

		Q+Q	*Q+R*	*R+R*	*R+Q*
Sci_TG Corpus	*P1*	1	Ø	1	Ø
	Rest	322	30	8	170

Table 28. Occurrences of double cases of the first type in *Sci_TG Corpus*
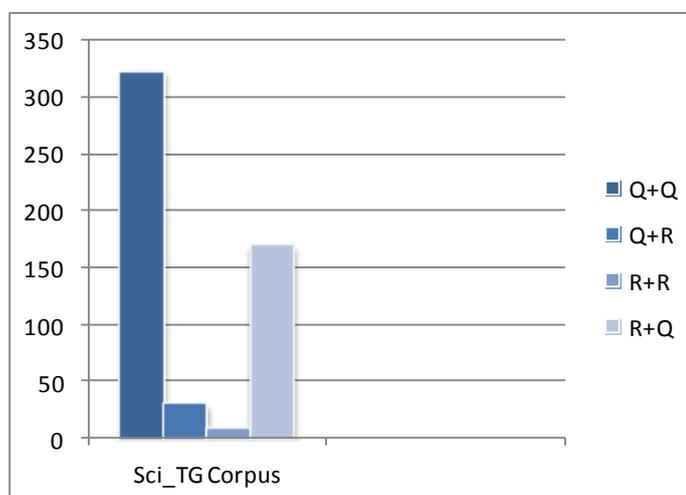
Figure 35. Distribution of double cases of the first type in *Sci_TG Corpus*

- Double quotation

The most frequent double case is that of two quotes following one another (*Q+Q*), a combination that appears a total of 323 times in *Sci_TG Corpus*, representing 59.26% of the total of double cases encountered in the corpus. Of this number, only two cases are found in the first paragraph.

The combination of two quotes is normally carried out by the link of a single reporting clause. This comes at the end of the first quote and acts as a bridge between the two quotations and, thus, between the two messages reported. As seen in Example 81 below, the structure followed is [Sentence 1: (REPORTED CLAUSE 1 + REPORTING CLAUSE) + Sentence 2 (REPORTED CLAUSE 2)]:

(81) *"This is probably the $64bn question when it comes to stem cells, but it's not a simple question and there will be no simple answer," <u>said Professor John Burn, head of the Institute of Human Genetics at Newcastle University</u>. "What is useful about this is that it will allow us to compare cells that have been rewound with genuine embryonic stem cells, so we can assess which are best."* (Sci_TG06259)

The reporting clause is not repeated in the vast majority of cases, as the reporter probably considers that the proximity between the two reported clauses make the reporting clause between them extensive to both. There are very few cases in which the reporter has chosen to introduce a second reporting clause in the structure, either after the first reporting clause (more frequent), as seen in Example 82 below, or after the second reported clause. This creates structures such as [REPORTED CLAUSE 1 + REPORTING CLAUSE 1 + REPORTING CLAUSE 2 + REPORTED CLAUSE 2+] or [REPORTED CLAUSE 1 + REPORTING CLAUSE 1 + REPORTED CLAUSE 2+ REPORTING CLAUSE 2].

(82) *"We now have the technology to fish down below two kilometres, something that has been brought in on a large scale in the North Atlantic in the past 15 years," <u>said Dr Myers.</u> <u>He added:</u> "What's alarming is we're eliminating fish we don't know much about at all and they are often the biggest conservation concern. If you're fishing for cod and you fail to catch any, you can quit fishing for it and hope stocks recover, but no one pays attention to the by- catch, so they can be fished out without anyone picking up on it."* (Sci_TG06449) [Emphasis added]

- Mixed [Q+R]

Cases of a report following a quote represent 5.50% of the total of double cases encountered in the corpus, and are all located in paragraphs other than the first one. It is not very frequent to find a quote preceding a re-elaboration of language made by the journalist in a report, and they convey different functions that will be later described. An example of a report following a quote can be seen below:

> (83) *"Scientists are the best people to do science - but, when it comes to which large new facilities should be chosen above others and where they should be located in the country, the wider scientific community and industry should contribute to the decision-making process," said Leigh. He also said more should be done to use them to inspire young scientists.*(Sci_TG0726)

- Double rephrasing

Cases of two consecutive reports are scarce in the corpus analysed. Only eight occurrences have been observed in *Sci_TG Corpus,* just one of which is found in the first paragraph (12.5%). The rest of the occurrences (87.5%) are found in subsequent parts of the science popularization articles studied. An instance of this can be seen in Example 84 below. The combination of two reports is less popular among journalists who write in the scientific section of the British newspaper. They seem to prefer using at least one piece of verbatim message to construct a case of double reporting.

(84) *James Healy, a spokesman for the HFEA, said the fertility watchdog was confident that PGS was beneficial for the groups of patients it recommended should use it. But he said it would consider the research .*(Sci_TG06428)

- Mixed [R+Q]

In the case of a quote following a report, journalists seem to provide readers with the significance of the language event reported. By using their own elaboration of the words uttered by the person accessed, they later include the original message that supports journalist's explanation, as seen in Example 85 below. A total of 170 cases of *R+Q* have been identified in the corpus, all of which are found in non-initial paragraphs of the texts.

(85) *Claus Speth, a consultant forensic pathologist from Woodbury, New Jersey, said the study was extremely valuable because of its detailed approach. "Many medical examiners' offices don't take the time and don't have the money to [carry out] these studies."* (Sci_TG06388)

The second type of double cases encountered, the combination of a quote or a report with a different element, it is typically formed by a quote or a report and either an embedded clause or a reporting adjunct introduced by an epistemic stance adverbial (*according to*).

It is relevant to note beforehand that the significance of this type in relation to the total amount of double cases in *Sci_TG Corpus* is very low. There are only 15 cases of this second type of double cases, which represents a 2.75% of the total number of double cases (545), and whose distribution of occurrences can be seen in Table 29 below.

Sci_TG Corpus		E+Q	Q+E	AC+Q	R+E	AC+R
	P1	Ø	Ø	Ø	Ø	Ø
	Rest	2	2	9	1	1

Table 29. Occurrences of double cases of the second type in *Sci_TG Corpus*

Clauses sometimes appear embedded into reports or quotes. These embedded stretches of language normally seem to reproduce verbatim words (Example 86 below) and are enclosed between inverted commas. This allows the reader to identify that somebody else's words have been introduced into the text.

(86) *"This raises concern," said Sudhansu Dey at the Vanderbilt University Medical Centre in Nashville. If the same was true in humans then smoking cannabis might mean "that pre-implantation embryo development will be abnormal and that the embryo might get trapped in the oviduct instead of going to the uterus", he said.* (Sci_TG06177)

The most frequent combination of this second type of double language reports is that of a reporting adjunct followed by a quote. In this case, the reader is first provided with an explanation, in this case introduced by the epistemic stance adverbial *according to*, followed by the original message that expands the information given in the first element. An example of this can be seen below:

(87) *According to Cathy Turner, a consultant forensic scientist at the FSS, the rapid advances in DNA technology have transformed the role of forensic scientists. "We've gone beyond corroborating allegations to using DNA and other techniques to provide fresh intelligence," she said.* (Sci_TG06120)

From a functional perspective, double reports found in the corpus accomplish a series of different functions.

One of the main features to distinguish patterns within this type of double reports is to establish whether the second element in the reporting structure adds information to what the journalist has stated in the first or, simply reformulated the information already given.

Regarding first paragraphs, we only have two double cases in which two reports and two quotes respectively combine to form a complex structure. As we can see in Example 88 below, the second quote is included for additive purposes, and the journalist uses it to complete the meaning already expressed by the first one. This provides the reader with a more ample panorama of the situation described.

> (88) [Simon, a recent first-time father of similar age, worried a little more about reproductive biology in his 30s.] *But not about his own: "Always the consideration was my girlfriend's biological clock," he says. "You just think you can be Charlie Chaplin if necessary, and have a baby you're too old to pick up."* (Sci_TG06289) [Background information added]

With regards to the presence of these structures in subsequent sections of the texts, we find that they accomplish much more varied functions. In the majority of cases (taking into account the four major structural combinations $Q+Q$, $R+R$, $R+Q$ and $Q+R$), the second structure of the double report fulfills additive purposes (184 times, 33.75%). The second elements of these structures are used as additional information to complete the meaning provided in the first structure, and

also as continuative devices to keep the reader on the same topic that is being narrated, as seen in the examples below:

(89) *"I think what we have done with this report is reverse the burden of proof on to the fishing regulators and fishing industry,"* **said** *Dr Rogers. "If you want to fish them, really you are going to have to prove to us that it is not going to do significant damage to these habitats."* (Sci_TG0643)

(90) *Prof Dowling* **said**: *"What we've shown here is the kinds of technologies and trade-offs and advantages they might bring. There are significant technical challenges to be overcome if we're to see an aircraft based on that concept.". "Some of the individual technologies one might see on more conventional looking aircraft in the nearer future."* (Sci_TG0658)

(91) *"My fear is that if this is adopted, unacceptable fishing practices will continue and may be accelerated." He* **added** *that by the time people realised how dire the situation was the point of no return may have been passed.* (Sci_TG06136)

In Example 91, it is the verb itself which provides us with the idea of addition. However, in Example 92 this addition is possibly achieved due to stylistic reasons. The reporting signal, which is shared by the two elements in the reporting structure, affects two different quotations that are semantically linked together in a continuative structure.

(92) *"I don't think this test is very sensitive or very accurate," said Mohamed Taranissi, who carried out one of the studies and runs the Assisted Reproduction and Gynaecology Centre, a London fertility clinic." I think people need to understand that before they make their decisions."* (Sci_TG06428)[Q+Q]

Not all structures that have an additive function are exactly the same. We have found a few cases in which the meaning of the second structure is not only additive but also explanatory. In this group of double cases, the second element

normally clarifies the meaning of the second element, as seen in the examples below:

(93) *Colin Chiverton of the agency said: "We never envisaged it would take this long. As long as the firewater is stored it remains a hazard." He said the delays were down to legal issues - technical consultants have already agreed to use technologies such as reverse osmosis and activated charcoal to clean up the water.* (Sci_TG06195) [Q+R]

Apart from structures which add extra information to the meaning of the first element in the structure, we have also found cases (54 occurrences, representing 9.91% of the total number of double cases in non-initial sections) in which the information provided by the second element is different from that provided by the first one, but also necessary for the reader to interpret both elements in the structure. Several examples of this structure can be seen below:

(94) *Richard Stokes, leader of Slough borough council, believes the MoD is trying to push Orion through before the law on crown immunity changes in April. "After that time, they will need to apply for planning permission like everyone else and we're hoping we can delay the decision until that time," he said.* (Sci_TG06438) [R+Q]

Additionally, we have found that the combination of two elements in a reporting structure can have adversative meaning. There are 11 cases (2.02%) in *Sci_TG Corpus* in which the second element of the structure presents the reader with a counter-argument to what was asserted in the first element, as seen in the examples below:

262

(95) *Houses and buildings in northern Europe typically have windows to the west to make the most of meagre winter sun, he said. "But in warmer countries you will never find windows to the west because the sun just pours in all afternoon during the summer."* (Sci_TG07164) [R+Q]

Several examples have also been found of structures in which the second element complements the first one by providing an example or a comment regarding what has already been said. As seen in the examples below, the first element of the structure quotes statements by *Sir David King, a Purdue University spokesman and Simon Fishel* respectively, and the second element remarks on the first one. In Example 96 below, it is *Sir David King* who cites a country as an example of his first statement:

(96) *"Transport has never been cheaper and we all want to travel more," said Sir David King, the government's chief scientific adviser, who headed the review. He **cited** China as a worst case scenario.* (Sci_TG06430) [Q+R]

There are other cases in which the second clause in the reporting structure does not add more information, but is significant to the first element. As seen in Example 97 below, the second element is fulfilling the function of relevance typical of newspaper discourse (Bell 1991:157-158), in that the quote shows the direct effect on the reader's own life experiences. In other words, they are used to reinforce the idea that the scientific finding is a breakthrough:

(97) *According to Cathy Turner, a consultant forensic scientist at the FSS, the rapid advances in DNA technology have transformed the role of forensic scientists. "We've gone beyond corroborating allegations to using DNA and other techniques to provide fresh intelligence," she said.* (Sci_TG06120)

We have also observed cases in which the quote in the reporting structure is used as the basis for the evaluation that is carried out in the element that appears later. The analysis of *Sci_TG Corpus* has provided us with examples of reporting structures in which the second element provides the basis for the report that is presented in the first element. In the reporting structures below (Example 98), the first element is a rephrase that the journalist presents the reader with, while the second element presents the actual words that serve as basis for the first element:

(98) *Claus Speth, a consultant forensic pathologist from Woodbury, New Jersey, said the study was extremely valuable because of its detailed approach. "Many medical examiners' offices don't take the time and don't have the money to [carry out] these studies."* (Sci_TG06388)

Examples have also been found of reporting structures that establish a cause-consequence or consequence-cause relation between the two elements. In Example 99 below, the quote illustrates the consequence that is caused by the report which appears in first position:

(99) *Dr Jagust said that there had been an explosion of research into the molecular basis of Alzheimer's disease. "That may lead to effective drugs to cure or prevent the disease, and those drugs would likely be more effective the earlier they are given. That has given a new urgency to research in predicting as early as possible those who will go on to develop Alzheimer's," he said.* (Sci_TG06364)

Conversely, in the example below, the causes for the first quote are made explicit in the second clause that appears in the reporting structure:

(100) *"I don't think we yet know why the expected benefit has not come to fruition," Prof Cedars said. "Are we doing some damage to the embryo? Are we testing the wrong things? Are we testing the wrong people?"* (Sci_TG06428)

We have also found cases of a reporting structure in which the second element, a quote, reformulates what has just been made explicit in the preceding report:

(101) *The robotics experts were commenting on a report published by the Office of Science and Innovation's Horizon Scanning Centre in December. The authors of Robo-rights: Utopian dream or rise of the machines? wrote: "If artificial intelligence is achieved and widely deployed (or if they can reproduce and improve themselves) calls may be made for human rights to be extended to robots."* (Sci_TG07180)

Examples of the question-answer pattern have also been found in *Sci_TG Corpus.* In this case, as seen in Example 102 below, two quotes form a reporting structure in which a question is posed in the first element and its answer is provided in the second element:

(102) *"The question was: what was the reason for this association?" she said. "With this research, we have established a direct link."* (Sci_TG06425)

Finally, we have also found cases of 'meta-report', which is a reporting structure in which a report is inserted within another report in a 'meta-narrative' way:

(103) *"I was in the room when the president made the announcement [on the moon and Mars vision] at Nasa headquarters," says Cowing. "He never said, dear Nasa administrator, you have my approval to take anything that that agency does and sacrifice it to meet this one goal."* (Sci_TG06418)

6.2.2.2. Multiple reporting

Multiple reports have also been identified in the corpus. They consist of more than two elements combined in a single reporting unit. In *Sci_TG Corpus*, we have identified multiple reports formed by three, four and five elements, named 'triple', 'quadruple' and 'quintuple' reports respectively. The total number of multiple reports encountered in the corpus is 57.

- Triple cases

Triple reports are the most numerous multiple report cases found in the corpus. Maybe due to their similarity in size to double reports, they constitute an extended version of double reports which has shown a very strong tendency towards non-initial sections of the corpus. As seen in Table 30 below, there are no

occurrences of multiple reports in first paragraphs of the corpus. The absence of multiple reports from initial sections may be due to their length, as first paragraphs tend to be short and concise, while multiple reports provide readers with details that require quite lengthy paragraphs.

As we have seen with double reports, there are multiple combinations, mainly between the two principal phenomena in reporting, quotes and rephrases, but also with the presence of embedded clauses and reporting adjuncts combined with quotes and rephrases, as seen in Table 30 below.

	R+Q+Q	*R+R+Q*	*R+Q+E*	*Q+Q+Q*	*R+Q+R*	*Q+Q+R*	*Q+Q+E*	*AC+Q+Q*	*Total*
Sci_TG Corpus	37	3	1	7	1	1	1	1	52

Table 30. Different combinations of triple reports in *Sci_TG Corpus*

Examples 104 below shows case of *R+Q+Q*, the most common combinations found in the corpus:

(104) *Speaking at the annual meeting of the European Society of Human Reproduction and Embryology in Prague, Prof Berga said that people suffering stress often do not notice or admit to it. "These women didn't report feeling stressed. But when we spoke to them it was clear that they tend to have a loss of perspective, they think they can get more done in a day than is realistic and their sense of self- worth is linked to their achievements at work," said Prof Berga "The therapy was targeted to what was bothering them. We teach them to love themselves."* (Sci_TG06248)

From a functional perspective, Example 104 above presents the reader with a rephrase of the first element of the reporting structure, after which two quotes are projected. The first of the quotes repeats what has already been said in the first element and uses the neutral verb 'say' as the reporting signal. The second quote shares the reporting verb but does not repeat the same information. Conversely, it adds new information to what has been previously said, and thus we can schematize this triple report as: Rephrase + Repetition + Addition.

- Quadruple and quintuple cases

The presence of quadruple and quintuple cases in *Sci_TG Corpus* is very scarce (5 cases in total), but their sole presence implies that reporting in science popularization articles can be a very complex procedure in which many elements are involved.

As we can see in Table 31 below, there are only four cases of quadruple reporting and just one of quintuple reporting, none of which are located in the first paragraph.

	P1	*Rest*	*Total*
Sci_TG Corpus	Ø	4	4

Table 31.Occurrences of quadruple reports in *Sci_TG Corpus*

Among the four cases of quadruple reporting encountered only one single occurrence has been found of the combinations *R+Q+Q+Q; Q+Q+Q+R;* *Q+Q+R+Q* and *Q+R+Q+Q*. Example 105 shows a case of *Q+Q+ R+Q* report in

which, from a formal perspective, elements that follow the first quote are used by the journalist to provide readers with more information in an additive sequence of quotes and a rephrase.

(105) *"Hit prediction is possible," said Dr Whitman. "What you do is make a global recommendation - instead of recommending something to a person, you recommend it to a very large community and see how well it will react." But he is reticent about using the software in this way: "It's not really an interest of ours."* (Sci_TG06441)

In the case of the only quintuple report located (Example 106) reports are interspersed with quotes forming a structure $R+Q+Q+R+Q$. As seen in Table 32 below, the only case encountered does not occur in an initial section.

	P1	*Rest*	*Total*
Sci_TG Corpus	Ø	1	1

Table 32. Occurrence of a quintuple report in *Sci_TG Corpus*

(106) *Even so, he said that there was an important **market** for drugs that could speed recovery from wounds. [R1] "People don't pay too much attention to wound healing because most of us have no problem with it," he said. [Q1] "But there's a number of people that have problems." [Q2] He cited the example of diabetics, many of whom have problems with wounds because of poor circulation, and people who suffer from bed sores. [R2] "It's not cancer, HIV or heart attack, but there is a **market** for it." [Q3]* (Sci_TG06243) [Emphasis added]

From a formal point of view, the first element in the reporting structure above is a rephrase to which two quotes add information. Furthermore, a rephrase is used to emphasize the importance of the discovery for diabetics, thus

emphasizing the benefits for daily life. Finally, the last element in the reporting structure is a quote which repeats the information that has been conveyed in the first rephrase *[R1]*.

6.2.3. Adverbial reporting

As was extensively discussed in Chapter 3, reporting adjuncts introduced by epistemic stance adverbials are used by science popularization journalists as a resource in attributing language events reported in a very efficient and economical way.

Within stance adverbials, we have focused our research on the presence of *according to* in science popularization articles from *The Guardian* newspaper. Due to its reporting function, this adverbial is used by reporters to signal the reporting function of the person or group of people responsible for the language event reported, as shown in Example 107 below.

> (107) *The discovery of a mysterious solid form of natural gas off the east coast of Canada could bring one of Earth's biggest untapped sources of energy a step closer to commercial use, according to scientists.*(Sci_TG0613)

To carry out the analysis on this epistemic stance adverbial, we have established our focus on its presence in the corpus, the patterns it follows, the words that most frequently accompany it, and its typical location in texts: either thematic, mid or rhematic position.

6.2.3.1. *Presence of* according to

The first data we wanted to obtain on the epistemic stance adverbial *according to* is how commonly it is found in the science popularization articles studied. With the obtaining of these data, we get a good idea of how much the reporters use it and where they use it, thus revealing the relevance of this adverbial in the construction of the attribution process in this type of text.

	P1	*Rest*	*Total*
Sci_TG Corpus	84 (50%)	84 (50%)	168

Table 33. Presence of *according to* in *Sci_TG Corpus*

As Table 33 shows, *according to* is a relevant element in attribution processes in science popularization articles of the newspaper *The Guardian* in the period studied. Both first paragraphs and subsequent sections of the texts seem equally common place for this epistemic stance adverbial, but the significance of occurrences in each part is quite different. Attribution through the epistemic stance adverbial *according to* in the corpus is quite relevant in the case of first paragraphs. While the percentage of the cases of attribution which entail the use of the aforementioned adverbial doesn't seem to be very significant, the importance of the adverbial in first paragraphs is notorious. While 84 occurrences in first paragraphs represent 43,98%, the relevance of occurrences in the rest of the text is only of 3,84%. Using this as a basis, deeper studies into transitivity and

thematic structures were carried out on the data extracted from *Sci_TG Corpus* to obtain more information about the behaviour of this adverbial.

		Cases of according to	*Occurrences per 100 attribution cases*
Sci_TG Corpus	*P1*	84 (50%)	43,98%
	Rest	84 (50%)	3,84%
	Total	168 (100%)	7,07%

Table 34. Relevance of *according to* in *Sci_TG Corpus,* in relation to attribution cases

6.2.3.2. According to *and transitivity*

Regarding transitivity structure, we analysed occurrences of *according to* following Halliday's framework, taking into account the fact that clauses evolve simultaneously in a grammatical function that expresses the reflective and experiential aspect of meaning in language. This evolution, called transitivity, "specifies the different types of process that are recognised in the language, and the structures by which they are expressed" (Halliday 1985: 101). This process, as we have already seen in Section 3.3.2.1., is said to consist of three components:

- The process itself

- Participants in the process

- Circumstances associated with the process

(Halliday 1985:101)

In relation to Halliday's model, we analysed the participants that most frequently accompany occurrences of *according to*, in right-one position (hereafter R1), to obtain information on the characteristics of the attributes in first paragraphs of *Sci_TG Corpus*. For this purpose, *according to* was taken as 'node word' (central element) to obtain concordances and examine the words that most frequently occur in R1 position of the epistemic stance adverbial studied, dividing them according to their material (*study, poll, Ministry, museum…*) or personal character (*scientists, researchers, Dr. Gow…*).

From this inspection, we obtained data that show an imbalance between material and personal elements in the right position of *according to.* We concentrate on first paragraphs as we want to focus on the first occurrence of the epistemic stance adverbial in the texts. People, groups of people and studies or institutions are found as attributees of language events reported in the corpus, as seen in Table 35 below, but references to material entities dominate the first paragraphs of *Sci_TG Corpus.*

Attributees	Occurrences
People	29 (34.53%)
Material Entities	55 (65.47%)
Total	84

Table 35. Participants accompanying *according to* in first paragraphs of *Sci_TG Corpus*

It is relevant to note that the number of occurrences of an entity in the R1 position of the adverbial is notably higher than that of references to people or

groups of people. Personal references most repeated include *experts* and *researchers*, but above all, the majority of personal references are made to *scientists*. In the case of material references there is not such homogeneity, and instances in R1 vary between a *poll,* a *study* or a *report.* However, it is more common to find references to a specific *publication* where the scientists unveil the results of their research.

Contrary to what we could reasonably expect in other text types of newspaper discourse, people are mentioned in science popularization articles as authorized attributees of language events reported, as well as also material entities, which occupy a relevant place among the attributees, as seen in Figure 36 below.
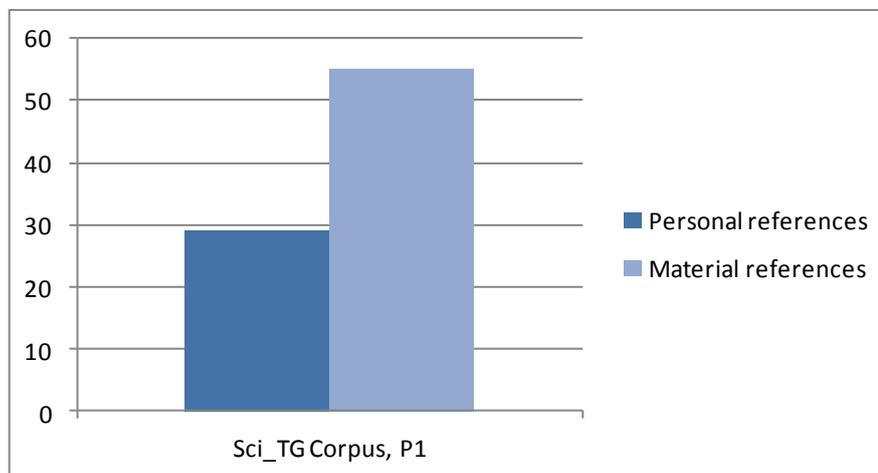
Figure 36. Distribution of personal and material references in first paragraphs of *Sci_TG Corpus*

There are further differences between both material and personal entities which appear as participants in the R1 position of *according to.* These differences are more specifically between the degree of detail provided by the reporter about

the attributee to whom the language event reported is being ascribed. For this purpose, we have carried out a comparative analysis of R1 references in first paragraphs and non-initial sections of the texts. As seen in Table 36 below where data on non-initial parts of the text are provided, specificity about people vary from the very vague reference to *"sources who have seen it",* to the much more specific *"Stephen O'Brien at the National Cancer Institute in Frederick, Maryland, US".* It is also relevant to remark on the high presence of what Thompson (1996b) calls 'unspecified others', such as *"scientists"*. This is in contrast to the cases in which the proper noun of the person attributed, is accompanied by information on his/her position, as in *"Sharon Garrison, Niac's co-ordinator at Nasa".*

Regarding material entities, very similar patterns have been found in *Sci_TG Corpus*. Some references provide readers with relevant data to identify the source, such as *"according to the report in Proceedings of the National Academy of Sciences, published yesterday".* In other cases the reporter chooses to omit most of the details, as in *"according to three new scientific studies".* It is also relevant, in this respect, to take into account that a deeper analysis could be carried out on thematic progression, considering attributees which appear in the text for the first time and those that are rather references to a previous antecedent. This could explain the differences observed in the quantity and type of information provided in each case.

Attributees	Cases
People	*According to sources who have seen it,...(Sci_TG06386)* *..., according to the UN's team of climate experts (Sci_TG06386)* *According to scientists in Germany,...(Sci_TG06403)* *According to Mr. Ingram,... (Sci_TG06389)* *..., according to scientists.(Sci_TG0705)* *According to Stephen O'Brien at the National Cancer Institute in Frederick, Maryland, US,...(Sci_TG06448)* *..., according to scientists who tracked a herd by satellite monitoring (Sci_TG06455)* *..., according to Sharon Garrison, Niac's co-ordinator at Nasa (Sci_TG06453)*
Entities	*..., according to a BBC poll (Sci_TG06431)* *According to Chemistry and Industry magazine…(Sci_TG06412)* *According to the new results,...(Sci_TG06413)* *According to a survey of thousands of couples who had babies during the 1990s,... (Sci_TG06451)* *..., according to the report in Proceedings of the National Academy of Sciences, published yesterday. (Sci_TG06455)* *..., according to the study in Proceedings of the National Academy of Sciences (Sci_TG0704)*

Table 36. Examples of material entities and people as attributees in *Sci_TG Corpus*

Things are slightly different when we examine first paragraphs of the articles that are part of the corpus. There is not a single personal reference which makes explicit the name of the scientist or experts to whom the discovery is

attributed. Information provided on the people who stand as attributees within the reporting unit is much less specific than the one provided in non-initial sections of the text. As seen in Examples 108 and 109 below, the level of details provided in first paragraphs is notably lower than that of non-initial parts of the text:

(108) *Earth's temperature could rise to levels far higher than predicted under the impact of global warming, according to the UN's team of climate experts.* (Sci_TG06386) [Rest]

(109) *Obesity can weaken the body's immune system and reduce its ability to fight off infections, according to scientists.* (Sci_TG0705) [P1]

In the case of material entities, a wider variety of information is found in first paragraphs, where, as in the rest of the text, there are important differences in the degree of detail provided in R1 references. As seen in Examples 110, 111 and 112 below, material attributees range from the succinct *three new scientific studies*, to the much more informative cases of *a long-delayed government-funded study* and *a new study which found that having your spine manipulated does nothing to speed up recovery.*

(110) *A screening technique used to select the best embryos during fertility treatment does not provide any benefit for patients, according to three new scientific studies.* (Sci_TG06428)

(111) *Environmentally friendly vehicles using hydrogen-based fuels and hybrid power sources will have little impact in preventing "dangerous and irreversible pollution" within 15 years, according to a long-delayed government-funded study.* (Sci_TG06429)

(112) *Going to a chiropractor to treat back pain could be a waste of time, according to a new study which found that having your spine manipulated does nothing to speed up recovery.* (Sci_TG0728)

6.2.3.3. According to *and thematic structure*

The last part of the analysis carried out on the occurrences of epistemic stance adverbials found in the corpus concerns the location where this adverbial is most commonly found: thematic, mid or rhematic position. As explained in detail in Chapter 4, Halliday defined theme as "the element which serves as a point of departure for the message; it is that with which the clause is concerned". He further defined it as "the starting point for the message, it is what the clause is going to be about"; and its counterpart, rheme, as "the reminder" of the clause (Halliday 1985: 38-39). Following his framework, we have obtained data on the different positions that the epistemic stance adverbial *according to* most frequently occupies.

As seen in Table 37 below, *according to* shows a different behaviour depending on the part of the text where it is located. When found in first paragraphs, *according to* tends to occur in rhematic position in a high percentage

of the cases (76.18%). Conversely, thematic position is not common at all, as just one occurrence of *according to* was observed in this position, and no cases were observed of absolute initial position, i.e. beginning the news story. Mid clause position does not seem to be a very typical position for this adverbial either, as only 19 cases (22.62%) have been identified in this position.

	P1			Total
	Theme	*Mid*	*Rheme*	
Sci_TG Corpus	1 (0.84%)	19 (22.62%)	64 (76.18%)	84

Table 37. Clausal positions of *according to* in first paragraphs of *Sci_TG Corpus*

In Examples 113 to 115 below we can see real instances of *according to* and R1 references in first paragraphs of the *Corpus*. Special attention has to be paid to Example 113, where we observe that the only occurrence of the epistemic stance adverbial in thematic position is not placed in absolute thematic position. This is because the adversative conjunction that is placed before it is not the first sentence in the paragraph. We should take into account that while *But* is a textual element that links the sentence with the previous text, *according to* is a purely interpersonal resource used in this case to construct the attribution.

(113) **But** *according to researchers in Manchester University, anyone can now produce the sound of a Stradivarius* (Sci_TG07178) [Theme. Emphasis added]

(114) *A maths professor has come up with a way of distracting children on heatwave car journeys - by inventing an equation which predicts when the first of them will crack and whine: "Are we nearly there yet?" The time-honoured query from the back seat can be calculated, <u>according to Prof Dwight Barkley</u>, by a piece of simple mathematics which in itself should keep a bored child quiet for a decent spell.* (Sci_TG06198) [Mid-clause. Emphasis added]

(115) *Mental disorders in children are on the rise, <u>according to a study of nearly 700,000 young people</u>.* (Sci_TG07253) [Rheme. Emphasis added]

Before dealing with reporting verbs in the corpus and as a general summary of the different types of attribution structures found in *Sci_TG Corpus,* we present Table 38 below with the data that was described throughout the last sections:

Sci_TG Corpus		*P1*	*Rest*	*Total*
Simple reporting	*Quotes*	4	837	841
	Rephrases	101	615	716
Multiple reporting	*Double cases*	2	545	547
	Triple cases	0	52	52
	Quadruple cases	0	4	4
	Quintuple cases	0	1	1
Epistemic stance adverbial	*According to*	84	84	168
Embedded clauses		0	48	48
Cases of attribution		191	2,186	2,377

Table 38. Attribution structures found in *Sci_TG Corpus*

6.2.4. Reporting verbs

Additional features of language events' reporting on which we want to focus, are the presence and characteristics of reporting verbs. Reporting verbs are one of the options available for the reporter to signal that the words uttered or written by somebody else are being incorporated into the text. However, they are not the only option, as inverted commas also set the beginning of a literal message being transcribed in the text.

As already mentioned in previous sections of this work, we are concentrating on projection structures as logico-semantic relations by which clauses function as representations of linguistic experiences. In this section, we are more specifically paying attention to verbal processes in the articles analysed which function as reporting signals in projection structures.

Taking this into account, we have analysed reporting verbs included into attribution cases located in the corpus, with the purpose of gaining a deeper insight into the use and patterns of this reporting signal in science popularization articles of the British press.

A total number of 2,268 verbs have been identified in *Sci_TG Corpus,* 2,160 of which are located in non-initial parts of the texts and 108 in first paragraphs. The difference between both parts of the text derives from the difference existing in the number of attribution occurrences in both sections, which is also notably higher in the rest of the text than in the initial parts of it.

As seen in Table 39 below, *Sci_TG Corpus* presents the percentage of cases of reporting verbs in the rest of the text at around 95% and a much lower proportion of reporting verbs in first sections. To a great extent, this difference can, to a great extent, be attributed to the high proportion of occurrences of reporting adjuncts introduced by the epistemic stance adverbial *according to*, which, as mentioned before, tend to concentrate more in the first sections of the text. *According to* is normally constructed in a structure that, as we have previously seen needs no reporting verb, and therefore their higher presence in initial sections explains the low number of reporting verbs observed.

	P1	*Rest*	*Total*
Sci_TG Corpus	108 (4.77%)	2,160 (95.23%)	2,268

Table 39. Occurrences of reporting verbs in *Sci_TG Corpus*

Apart from the absence of reporting verbs from the adverbial and some complex reporting cases described in previous sections, we have perceived the presence of one case in which no reporting verb exists. This case presents an uncommon structure, as seen in Example 116 below. It is formed by a first clause that acts as the background for the second clause. The first sentence is not the report of a language event, and thus not a case of attribution as we have considered here. However, the second has the form of a quote and verbatim reproduces the words of Dr. Fisher, who is, for the first time in the article, referring to the beginning of his research:

(116) *It began for Dr Fisher when he went to China as a medical student at the tail end of the Cultural Revolution, when Mao was still alive and the Gang of Four were in control. "I saw a woman on an operating table with her entire abdomen open, talking to the surgeons, with three needles in her ear. I became convinced something very interesting was going on."* (Sci_TG06274)

6.2.4.1. Tense in reporting verbs

Once we obtained data on the presence and distribution of reporting verbs, we carried out an analysis for the purpose of studying their behaviour. We analysed the verbal forms of reporting verbs encountered in the corpus, focusing both on their tense and their simple or complex forms.

As expected, most of the verbs are formed in the past tense (81.61%) while only 16.88% and 0.15% have been found in present and future tenses respectively. Continuous forms (present continuous, past continuous and gerunds) and complex forms are scarce in the corpus, as seen in Table 40 below, and only one occurrence of a reporting verb in the passive form was detected in the corpus.

	Present	*Past*	*Future*	*Non-finite clause*	*Passive*	*Total*
Sci_TG Corpus	387 (17.06%)	1,857 (81.90%)	4 (0.18%)	19 (0.84%)	1 (0.043%)	2,268

Table 40. Tenses of reporting verbs in *Sci_TG Corpus*

The fact that there is such a strong tendency for reporting verbs to be constructed in the past tense is due to the process of construction of the news

(already described in Chapter 2). The information reaches the journalist when something has already happened, and also when language events have also occurred. In the cycle of news, information quickly reaches the reporter and is almost instantaneously transmitted to the reader using virtual spaces such as web pages, but once the words to be reported have been uttered, the reporting verb that introduces the message is in past tense, as it is typically a narration about the past.

The present tense is the second most common tense, providing the reader with the impression of instantaneous transmission and a much more a-temporal character with the information conveyed (see Example 117 below).

> (117) *The researchers **believe** that Nanog, alongside other genes, kicks into action a cascade of complex biological machinery that forces cells back into their simplest state, before they have gone down the path of becoming one of the 200 cell types found in the body.* (Sci_TG06259)

The future tense provides, in this context, with a sense of prediction. The only four cases detected (0.18%), as Example 118 shows, seem to predict what medical experts will say, which entails what that the reporter has previously seen in a press release or a similar document where the experts advance some of the conclusions or findings that they will present.

> (118) *Sunbed users face nearly triple the risk of skin cancer compared with a decade ago as a result of higher-powered equipment, medical experts **will warn** today.* (Sci_TG07260)

Only one case of passive construction has been observed in a first paragraph of the corpus (Example 119), and refers to the information which will be unveiled

in a publication. This structure has a special thematisation structure, as it is *The dire state of the butterflies* (the object) that the journalist has chosen to place in thematic position, attracting the reader's attention and relegating the attribution source to a secondary mid position.

(119) *The dire state of Britain's butterfly species **is revealed** in a report today that highlights the counties where their decline is most marked.* (Sci_TG06199)

Additionally, we have detected 19 cases (0.84%) in which the reporting verb is part of a non-finite -ing clause which functions as an adverb of manner and is set at the beginning of the non-finite clause in thematic position. It has a similar "satellite" position to attribution constructed by means of reporting adjuncts (i.e. *according to* structures), as seen in Example 120 below:

(120) ***Writing*** *in the journal Nature today, the scientists raise fears that inevitably wetter weather will return the wetlands to their normal state in the next three to five years, boosting the amount of methane in the atmosphere by 10m tonnes a year.* (Sci_TG06100)

We can relate these examples to the use of reporting adjuncts, as they seem to fulfill similar structural functions. However, these examples suggest that journalists prefer to concentrate the emphasis of the clause on the circumstances of the event (mainly place and time), they seem to choose the reporting adjunct *according to* more often to accomplish this function, rather than constructions with non-finite clauses as in Example 120 above.

6.2.4.2. Classes of reporting verbs

As Thompson states (1994: 33), reporting verbs are one of the main ways in which the reporter provides the reader with information on the message and the circumstances under which it was uttered. Reporting verbs are a fruitful source of data on the type of message conveyed and the interpretation of that information by the reporter.

A total of 77 different verbs have been identified in the corpus, which are presented in order of their frequency of appearance in the corpus (see Figure 37 below):

SAY (1,702)	CITE (2)	HEAR (1)
ADD (97)	CONCEDE (2)	HIGHLIGHT (1)
BELIEVE (56)	CONFIRM (2)	INDICATE (1)
WRITE (46)	EXPRESS (2)	INVITE (1)
WARN (41)	FIND (2)	POINT (1)
TELL (36)	OUTLINE (2)	POSE [a question] (1)
SUGGEST (28)	PROPOSE (2)	PRAISE (1)
CLAIM (21)	PUBLISH (2)	PRESENT (1)
REPORT (20)	STRESS (2)	PROMOTE (1)
DESCRIBE (18)	WELCOME (2)	PRONOUNCE (1)
CONCLUDE (14)	ASSERT (1)	[As one scientist] PUT [it] (1)
ANNOUNCE (13)	ACCUSE (1)	RECALL (1)
REVEAL (10)	APPLAUD (1)	RECOMMEND (1)
ADMIT (9)	ASK (1)	REFUSE [to predict] (1)
ARGUE (7)	ATTACK (1)	REJECT (1)
URGE (6)	COME OVER [the speakers] (1)	REPEAT (1)
CRITICISE (6)	CONDEMN (1)	SHOW (1)
EXPLAIN (5)	CONTINUE (1)	SPECULATE (1)
INSIST (5)	CONTRADICT (1)	SUMMARISE (1)
PREDICT (5)	DECLARE (1)	TALK (1)
ADVISE (4)	DEFEND (1)	UNRAVEL (1)
POINT OUT (4)	DENY (1)	
AGREE (3)	DIRECT (1)	
CALL (3)	DISCUSS (1)	TOTAL 2,268
CALL FOR (3)	DISMISS (1)	
COMMENT (3)	ENVISAGE (1)	
STATE (3)	ESTIMATE (1)	
ACKWNOWLEDGE (2)	HAIL (1)	

Figure 37. List of reporting verbs in *Sci_TG Corpus*

As we can see in Figure 37 above, only 20 out of the 77 total verbs observed have a frequency of appearance (indicated between brackets) higher than 5, the minimum value considered relevant. In Figure 38 below we display the presence of those reporting verbs whose frequency is higher than 5, in a graph that highlights the high number of occurrences of the verb *say*, in respect to the rest of the verbs.
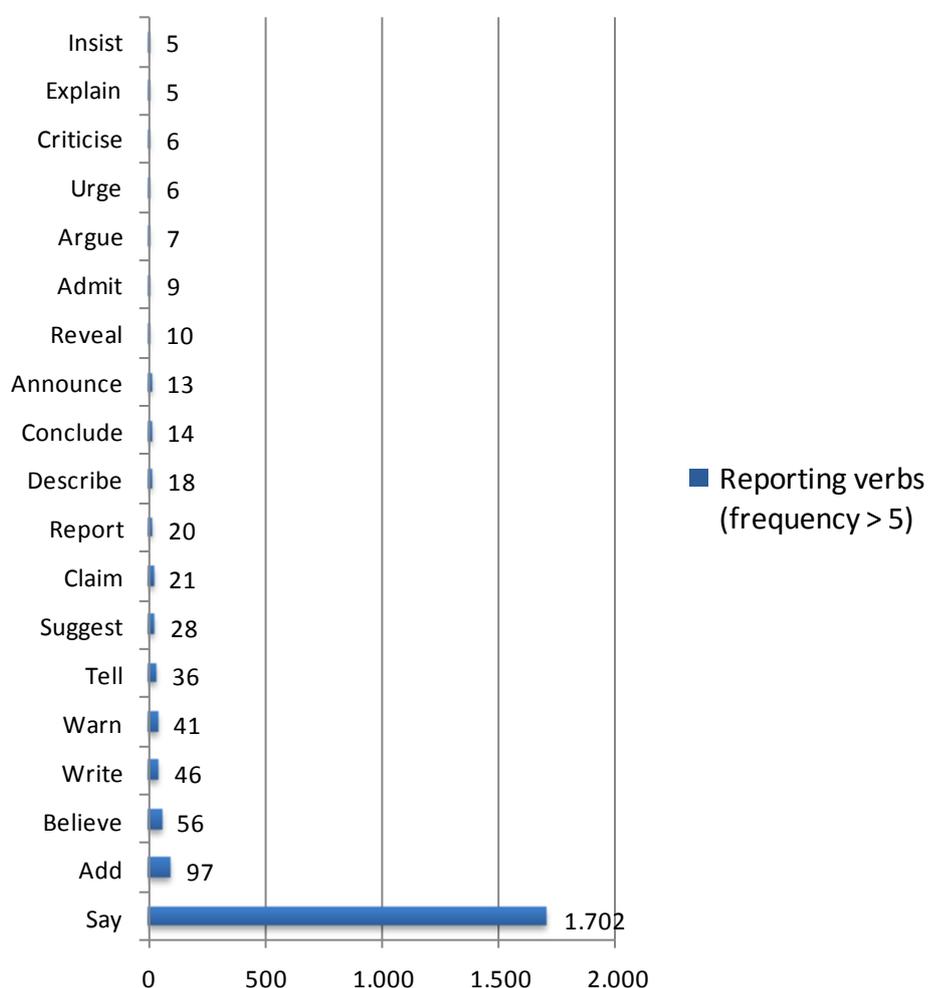

Figure 38. Reporting Verbs with more than 5 occurrences in *Sci_TG Corpus*

The most frequent verbs detected in the *Corpus*, apart from *say*, are *add* (97 occurrences), *believe* (56 occurrences) and *write* (46 occurrences). The high presence of *say* leads us to believe that it is the 'by-default' reporting verb, at least in science popularizations in the newspaper analysed (Example 121 below). This is not the case in popularizations in *El País* newspaper, for instance, where the most frequent reporting verb is *explicar* (explain) (Elorza 2011, García Riaza and others, 2011), as seen in Example 122 below taken from García Riaza and others. (2011):

(121) *"The widespread assumption that human evolution has slowed down because it's easier to live and we've conquered nature is absolutely not true. We didn't conquer nature, we changed it in ways that created new selection pressures on us,"* **said** *anthropologist Dr John Hawks, who led the study.*(Sci_TG0704) [Emphasis added]

(122) *"La visión de los primates atravesando sobre dos patas las distancias entre las zonas boscosas cada vez más apartadas ha mantenido su atractivo durante décadas de investigación,"* **explica** *White, para quien los datos obtenidos del ambiente en el que vivía Ardi socavan esta hipótesis y dan una visión nueva del nicho ecológico de los primero homínidos.* (EPMRE100527) [Emphasis added]

According to Thompson's classification of reporting verbs (1994:33-70), *say* is the basic reporting verb (1994:34), but a*dd* and *believe* represent two of the most repeated verbs used in reporting language events in the corpus. On the one hand, *add* is used by the reporter to show that the speaker has already said something, and is now supplementing the information with a language event (Example 123 below).

(123) *Prof Stringer* **added** *that the Ahob project was not the end of the story for the history of humans in Britain* (Sci_TG06125) [Emphasis added]

On the other hand, *believe* belongs in a series of verbs that fall under the term 'mindsay' (Example 124 below) which quote mental experiences that work in a similar way to those of hearsay, as they are classified in Systemic Functional Grammar (Halliday 2004: 201; Bednarek 2006: 643) and which have already been identified as reporting verbs of newspaper discourse by (Elorza 2011).

(124) *Scientists **believe** they have found a cure for the MRSA superbug after unearthing an existing drug on a computer database* (Sci_TG07265) [Emphasis added]

In respect to the verb *write* (46 cases), and as Thompson (1994: 35) states, it can be used as a neutral verb to report language events which have been expressed through a written medium, as seen in the example below:

(125) *"Rapid industrialisation and urbanisation in Asia have caused severe air pollution over many countries, including China and India. Long-term satellite measurements have revealed a dramatic increase in aerosol concentrations over Asia," **wrote** Dr Zhang yesterday in the Proceedings of the National Academy of Sciences* (Sci_TG07221) [Emphasis added]

The verbs *warn* (41 cases), and *suggest* (28 cases) *are* conversely non neutral reporting verbs that are used when the journalist wants to show the speaker's purpose (Thompson 1994: 36), as in Example 125 and 126 below:

(125) *It **warns** that in 20-40 years' time the Great Barrier Reef could be "severely damaged", forcing its closure, while other parts of Australia would be off-limits because of a rise in bushfires and insect-borne diseases.* (Sci_TG06108) [Emphasis added]

(126) *British scientists **suggested** that the dopamine levels could be due to women's bodies being better at eliminating the drug from the brain, an action which is linked to levels of oestrogen, the predominately female hormone.* (Sci_TG06331) [Emphasis added]

Regarding the verb *tell* (36 cases), Thompson (1994: 34) classifies it as a neutral verb which takes into account the hearer. In other words, while the neutral verb *say* simply shows that what someone said is being reported, without adding any information about the speaker's manner or purpose, *tell* includes information about the speaker and is normally used to report statements, instruction and suggestions, as seen in Example 127 below:

(127) *Randal Keynes, Darwin's great, great grandson, **told** the Guardian the project fulfilled the Darwin family's long-standing aim to have all of the scientist's work available for everyone.* (Sci_TG0681) [Emphasis added]

Once the frequency of reporting verbs was analysed, we examined the reporting verbs detected in *Sci_TG Corpus* to define their functionality.

In Section 3.3.2.1. (Chapter 3), we described a series of approaches which had taken reporting verbs as core elements in reporting processes. As it was already described, Caldas-Coulthard (1994: 306) classifies 'verbs of saying' in a taxonomy that considers them all according to their function in relation to the reported clause. Verbs are classified in three different categories: 'Speech-reporting verbs', 'Descriptive verbs', and 'Transcript verbs'. We focused on speech-reporting verbs, which can also be further divided into three categories. The first one, neutral structuring verbs, also termed 'glossing', introduces a report but does not evaluate it. Conversely, 'Asertive', 'Directive' and 'Expressive'

metapropositional verbs make the illocutionary force of the writer/speaker explicit by marking the stance towards the reported message. Finally, metalinguistic verbs, which, apart from making explicit the illocutionary force of the speaker, categorize the type of linguistic event that is being reported.

Additionally, we considered Halliday and Mathiessen's (2004: 448) model, which adds a complementary dimension to the descriptive parameters of a report. Since process types are considered by Halliday and Mathiessen as part of the 'transitivity' system through which the world and human experience are reconstructed as text, each process relates to a specific domain of experience in constructing meaning. Halliday and Mathiessen (2004: 172) identify different types of process types, namely behavioral, mental and verbal.

Below we present a table summarising our findings, which have been classified according to the aforementioned approaches:

		Type	Description	Cases in *Sci_TG Corpus*
Neutral verbs	**Neutral reporting signal**	Neutrally reporting what someone said	Do not add any further information on speaker's manner or purpose	*say, tell, ask* (interrogative)
Non-neutral verbs	**Verbs specific to speech function**	Showing the speaker's purpose	Add information on the intention of the speaker	*Acknowledge, admit, advise, announce, applaud, argue, assert, attack, call for, claim, concede, confirm, contradict, declare, defend, explain, highlight, insist, invite, point out, predict, present, propose, recall, recommend, refuse, stress, suggest, urge, warn*
		Showing the manner of speaking	Show the way in which something was said	*comment*
		Showing what was said through the reporting verb	Give in themselves an idea of what was said or written	*criticise, deny, discuss, dismiss, hail, praise, reject, welcome*
	Verbs specific to speaker	Indicating how the message fits in	Show that what has been reported fits in with the rest of the language event	*add, agree, cite, conclude, comment, continue, repeat*
		Drawing attention to the speaker's or writer's words	Words of the speaker or writer used to describe or name things	*call, describe, direct, envisage, estimate, express, indicate, outline, pose, promote, pronounce, put (it), report, reveal, show, state, summarise, unravel*
		Showing a tentative position of speaker	Expresses a non-confident position	*speculate, suggest*
		Showing whether a report is of speech or of writing	Type of language event report: spoken or written	*hear, publish, write, come over (the speakers)*
		Mindsay	Thoughts	*believe*

Table 40. Classification of reporting verbs from *Sci_TG Corpus*

Throughout this Chapter we have reviewed in detail all data extracted from the analysis of *Sci_TG Corpus,* and obtained patterns that are relevant in gaining a deeper insight into the anatomy of newspaper discourse. More specifically, we have gained insight into the phenomenon of reporting in science popularizations. We have drawn upon the theoretical basis set on the previous chapters in order to take a step further in the description of science popularization texts, by exploring the presence of attribution in science popularizations and the most frequent forms that it adopts. Reporting adjuncts and reporting verbs have received special attention as characteristic reporting signals which have provided us with relevant information on the different ways in which journalists introduce other voices into the text and create an authority that legitimizes his/her words. The location of reporting signals and clauses in the reports has also been a key node in this chapter, where we have also noted the relevance that journalists give to the different locations that elements occupy within science popularization articles.

Conclusions on the results of this study are exposed in the final part of this study, Chapter 7, where we have provided a general summary that gathers the most relevant points addressed in this study. We have also included final remarks and comments on the results obtained and the prospective work that could be done following this study.

Chapter 7

Conclusions

Chapter 7

Conclusions

7.1. Introduction

7.2. General conclusions

      7.2.1. First paragraphs

      7.2.2. Non-initial sections of the text

7.3. Pointers for future research

_____

La civilización no dura porque a los hombres sólo
les interesan los resultados de la misma:
los anestésicos, los automóviles o la radio.
Pero nada de lo que da la civilización
es el fruto natural de un árbol endémico.
Todo es resultado de un esfuerzo.
Sólo se aguanta una civilización
si muchos aportan su colaboración al esfuerzo.
Si todos prefieren gozar el fruto,
la civilización se hunde.

José Ortega y Gasset

7.1. Introduction

In this final part of the study, after having established a theoretical basis and carried out an empirical analysis using the framework provided by corpus linguistics and SFG, we bring together, in a few concluding pages, the essence of this research and the most relevant conclusions drawn.

Corpus linguistics, the research methodology followed in this study, has revealed itself as not only being a methodology, as many authors state, but also a theoretical approach to language analysis. This discipline has provided us with an ample theoretical foundation that keeps growing day by day. This is on account

_____

for the numerous investigations that are being carried out nowadays for the description of many languages, including Spanish (e.g. Parodi 2007), in the frame of this branch of knowledge. These are investigations that produce interesting outcomes which lead to a better acquaintance with the complexities of human language. Accordingly, corpus linguistics can neither be considered a purely methodological approach nor a theoretical discipline, but a hybrid and rich method for the study of language, whose methodological and theoretical dimensions benefit from one another.

A fact that we should not leave out in these conclusions is the actual relevance of empirical data in studies about language. Real data should not only only be the basis for all corpus linguistics study, but also be considered as a common source of information for linguists, as well as for researchers in other disciplines. Corpora can thus be conceived as an enormous repository of instances from where we can observe how language actually works and evolves, or how different registers and discourse types are used in specific situations. To support Sinclair's idea of invented examples as misconceptions of language, distortions made of languages, as well as his claim to use actual, fresh language produced by people as the basis for research in corpus linguistics, we have worked with real data extracted from an updated and reliable source; of communicative events which, through the newspapers in this case, convey meaning and transmit certain conceptions of the world. Much the same happens with how samples of language are collected and stored and the relevance of words in keeping their co-text.

_____

In this study, a corpus-based approach has proven to be a fruitful method of seeing how reporting is carried out in science popularization discourse. It has also been used as a technique to better understand language where the corpus is analysed in reference to previously defined classes. We have set those premises as our research method, analysing what language has to offer and using the extracted information to depict the phenomenon under study.

Media discourse has been taken as the sample of language where texts, and specially written outputs have been selected and analysed to find patterns and data about attribution which have been extensively described and explained in Chapters 5 and 6. Conclusions of this study are presented below.

7.2. General conclusions

Throughout the analysis carried out and presented in Chapter 6, language events' reporting has been unveiled as a common feature of science popularization discourse, as well as a productive source of data on types of language events and reporting signals used to report language events.

Explicit attribution in language events' reports is a conventional phenomenon in science popularization articles published in the newspaper *The Guardian*. We can accordingly conclude that it is frequent in this type of text, and can be considered a characteristic feature of science popularization articles in the publication and period of time studied.

_____

The encoding of external voices in newspaper discourse texts of *The Guardian* newspaper has been studied by means of the analysis of the types of reporting structures that are most frequently used in the corpus. A different behaviour has been identified between simple reports and complex reports, which has proven to be fruitful to establish a more delicate classification of this phenomenon.

7.2.1. First paragraphs

We know, from the analysis carried out, that the most frequent form of reporting found in first paragraphs of texts is that of simple reporting. Complex reporting is not a frequent phenomenon in initial sections. The analysis has revealed that complex structures are not primed to appear in first paragraphs of science popularization articles in *The Guardian*, as we have located only 2 cases of complex reports in initial sections of *Sci_TG Corpus.* Within simple reporting, it is more common to find rephrases in initial sections of the texts, rather than quotes which explicitly present the words uttered by a different source, as discussed in Chapter 5. The situation of first paragraphs in initial position of the text contributes to their special character. This is because they share part of the characteristics of headlines and leads, which do not provide extensive information but rather set the tone of the article (Zappaterra 2008: 61) by using concise, short words, addressed to catch the reader's attention and reach them in an efficient way (Reah 2002: 16). Therefore, the lack of the prominence of quotes in first paragraphs can be thus due to a reticence of journalists to present literal parts of

discourse at the beginning of the texts, but to elaborate on the introductory section of the article without making particularizations. The distribution of quotes and rephrases in initial sections of science popularization articles studied is not homogeneous at all, and it seems, as Vicente Mateu (2007: 234) already pointed out, that this is a deliberate choice, as indirect rephrases are much more objective, and want to provide the reader with an impartial vision on the language event reported which is presented at the beginning of the text.

Regarding adverbial reporting, we have concentrated on the reporting adjuncts introduced by the epistemic stance adverbial *according to* and its relevance in the Corpus. In this respect, and as García Riaza (2010a, 2010b) already pointed out, *according to* is one of the particles most frequently used to introduce explicit attribution in first paragraphs of science popularization articles, much more common than in subsequent sections of scientific articles published by *The Guardian* newspaper. The analysed data, as discussed in Chapter 6, corroborates her results, since cases of *according to* are significantly more frequent in first paragraphs than in the rest of the text. *Sci_TG Corpus* shows that there are many more occurrences of the adverbial per a hundred attribution cases in initial sections.

According to is used in the texts studied as an element which identifies the source of information, but no trace is left by this epistemic stance adverbial on the circumstances under which the language event was produced. Metalinguistic information is normally provided by the reporting verb, which, in those cases, is not explicitly stated. We can thus infer that the low percentage of reports and complex reports in first paragraphs is balanced by the relevant presence of quotes

and epistemic stance adverbials, the two phenomena par excellence in first paragraph reporting.

In regards to the relationship between the epistemic stance adverbial and Halliday's transitivity structure (1985:101), and more specifically between *according to* and participants occurring in R1 position, it is relevant to note that the majority of participants who act as attributees of *according to* in R1 position of first paragraphs are material entities. These are mainly *publications, reports, surveys* or *polls* over personal references, which are typically either single scientists or teams of researchers whose voices are included in the text. As a consequence, the source of the information provided by the journalist is more frequently a textual document which had already been published, rather than a language event. In this view, the relation with the source of attribution is an inter-textual one.

The level of detail when providing readers with information is another of the feature that distinguishes initial sections of science popularization articles from the rest of the text. Details in R1 attributees of *according to,* vary to several extents. Again motivated by the desire of initial sections to be concise and brief, journalists do not go into great detail in these sections. There are no extensive references to people, but just brief references to groups of scientists and experts who have carried out different projects, as in *UN's team of climate experts.* With reference to material entities, elements in *according to* R1 keep being much more laconic than those in non-initial parts of the text. We can thus conclude that first paragraphs condition the explicitness of the reference following the epistemic

stance adverbial, and attributes are presented in a much less detailed structure that the one provided in non-initial sections of the text. The journalist writing the articles provides readers with as much information as necessary to avoid being vague in his/her narration, but, as said before, different patterns have been observed with regards to personal and material entities. When personal references appear in R1 position of the epistemic stance adverbial, they are characterised by their names and the institutions or fields they work in, thus providing the reader with information on who they are and where they develop their research. In the case of references to material entities, the most frequent information found is where and when the information has been published.

Concerning the relation of the reporting adjunct introduced by the epistemic stance adverbialand the thematic structure of the clause in which it appears, and based on the fact that the section of the text where the occurrence of *according to* is located conditions its appearance, we narrowed our study to focus on the different positions in which the adverbial occurs within first paragraphs. When found in first paragraphs, *according to* tends to occur in rhematic position in a high percentage of the cases. The epistemic stance adverbial is thus primed to appear at the end of the clause and adopt the form of an adverbial which affects the whole previous clause. Conversely, thematic position is not common at all, as just one occurrence of *according to* was observed in this position, and it does not even occupy absolute initial position, but is preceded by an adversative conjunction which stands in absolute initial position.

Mid clause position is not a frequent place for this adverbial either, as only 19 cases have been identified in medial clause location. When the epistemic

stance adverbial *according to* appears in mid position, both the adverbial and the attributee on its right are normally presented between commas, in the form of an adjunct that stands in the middle of a clause. The fact that this position breaks one of the reporting clauses into two separate chunks may be the reason that it's less commonly used.

We can conclude that the thematic structure in first paragraphs of this kind of text strongly primes the presence of attribution by means of epistemic stance adverbial *according to* at the end of the reporting structure. This implies that the message is placed in thematic position, and emphasized over the attributee. The high frequency of appearance of the epistemic stance adverbial and its accompanying attributee in rhematic position means that the information conveyed is more important in initial sections than the source of the words reported. The focus is thus placed on the message in opposition to the source of attribution, which appears in rhematic position. The source of attribution is not hierarchically as important as the information conveyed, and becomes a subtopic in first paragraphs that will be later recovered in non-initial sections of the texts. This is a subtle technique used by journalists to include data on the source of the information that can afterwards be recovered as theme in non-initial sections of the texts.

7.2.2. Non-initial sections of the texts

Regarding non-initial sections of the texts, we say that, although the presence of simple reporting is higher than that of complex reporting, relevant complex structures have also been found.

As noted before, the tendency for quotes to appear in subsequent sections of the text is a very strong one. Because a good number of quotes have been identified in non-initial parts of science popularization articles from *The Guardian* newspaper, most of them as part of complex reports.

Rephrases, identified as the second most frequent form of simple reporting, present a more homogeneous distribution in *Sci_TG Corpus.* They do not show any tendency to appear with higher frequency in any of the two parts of the text studied, in contrast to the behaviour of quotes, primed to appear in non-initial parts of the text. The homogeneity of rephrases along the whole text maintains the status of the journalist as narrator, as mediator between the scientists and the lay reader. The inclusion of quotes in the text involves a polyphony that starts after the first paragraph, where the introductory information for the text has been given to the reader.

Concerning complex reporting, we have dealt with a phenomenon that had not been identified in previous studies, namely complex attribution. In this phenomenon more than one rephrase or quote are combined with one another or with different elements to form a complex structure that works as a suprasentential unit, both semantically and syntactically as its parts are mutually dependent and refer to a single language event. We have identified a good number

_____

of complex cases in *Sci_TG Corpus,* and, although the phenomenon is not as frequent as simple reporting is, the data obtained allow us to place complex reporting with a place among the features that characterize reporting in science popularization articles from the British newspaper *The Guardian*.

The most frequent of the types of complex reporting identified is double reporting, which is primed to appear in non-initial sentences of the texts, where we have found the majority of instances.

Two different structures have been distinguished for double cases in the corpus. One is according to the combination of a report and a quote (type 1) and the other is the mixture between one of the two types of simple reporting with a third element (embedded clause or epistemic stance adverbial). Double reports from Type 1 show four different varieties (*Q+Q; Q+R, R+R; R+Q*) among which the most relevant one is that of combining two quotes. In contrast to the relative relevance of *Q+R* and *R+R*, *Q+Q* combination has been found many more times in the Corpus. Additionally, *R+Q* structure is the second most frequent double structure in *Sci_TG Corpus.*

Q+Q structures are normally linked by a single reporting clause and, therefore, the reporting verb and the reference to the person or entity attributed is just one, shared by the two reported clauses. The fact that journalists who write science popularization articles often choose this structure leads us to believe that journalists use this structure for two reasons. On the one hand, two reported clauses sharing a same reporting clause allow the journalist to reproduce long

_____

stretches of other speakers' messages in his/her own narration in a very economical way.

On the other hand, if two reported clauses sharing a same reporting clause behave as a single unit introducing a polyphonic effect in the text, the function of the reporting clause is to keep a link with the journalist's narration, thus giving continuity to the text in terms of voice hierarchy. In those cases, a continuative verb with additive meaning, such as *add* or *say*, can be found.

The second most frequent combination of double reports, *R+Q* is also a relevant example of how journalists model language to include external voices in science popularization articles. In this case, it is the report that opens the double structure, followed by a quote, which, in the majority of cases, has no reporting verb. As it has occurred in the case of *Q+Q* structures, just one reporting verb introduces a language event split into two reported clauses. The rephrase acts as a summary that introduces readers to the context in which the language event was produced and prepares them for a deeper understanding of scientists' or experts' words, which come later in the quote, thus acting as a transition between the journalist's voice and the attributees.

The second type of double cases, either embedded clauses or epistemic stance adverbials combined with a quote, conform a very reduced group of occurrences in the corpus, but their relevance also stands in the line of *Q+Q* and *R+Q* structures, since verbatim pieces of texts are linked to the language event being reported by means of a single reporting element. In the case of embedded clauses and quotes, the reporting clause which introduces the adjoined piece of

language affects both the embedded and the reported one. Something similar happens in the case of reporting adjuncts introduced by epistemic stance adverbials, the most frequent combination, where the attributee accompanying *according to* and the reporting verb accompanying the quote provide readers with the information on the attribution source necessary for the whole structure.

Although they appear in a much smaller quantity than single, and even double reports, multiple reports have also caught our attention as structures that characterize the discourse of science popularizations. Triple, quadruple and quintuple reports form a very special type of chained attribution that constitutes a suprasentential structure in itself, normally a paragraph. No occurrences of multiple reports have been observed in first paragraphs, which is mainly due to their excessive length for an initial context where, as previously mentioned, the narration tends to be influenced by the short and concise words of headlines and leads, and very general information is provided as a means of introducing the reader into the text. The most recurrent combination of reporting structures in *Sci_TG Corpus* is *R+Q+Q,* followed at a certain distance by *Q+Q+Q*. As an amplified reproduction of the double case *R+Q*, in *R+Q+Q* the first quote has both reporting and reported clauses, but the second quote benefits from the reporting clause of the first quote. This feature also supports the tendency towards the reporting clause acting as continuative element and the reports marking the transitions which serve to introduce the quotes to the readers.

The scarce presence of quadruple and quintuple cases in *Sci_TG Corpus* does not diminish the relevance of cases in which a long string of reports combine

to provide readers with a portrait, as complete as possible, of the language event reported. The occurrences found show the repeated pattern of *R+Q+Q+Q*, which also adds its relevance and significance to *R+Q* and *R+Q+Q* structures. No other functions have been identified here.

The length of complex structures found mainly in non-initial parts of the text are able to modulate how outstanding the source of attribution is, a potential which is not fully exploited in the texts studied, due to the relatively low presence of those structures in the corpus.

Regarding the epistemic stance adverbial *according to*, analysed as a technique to introduce attribution in the text, we can say that it is not as common in the rest of the text as it is in first paragraphs, where it has been identified as a the preferred resource to explicitly present the source of attribution in rhematic position. In regards to personal and material references which can be found in R1 position after *according to*, journalists follow a common pattern, and we find detailed identifications of scientists from whom name, surname, academic degree and place of employment are sometimes provided. In contrast, in other occasions succinct references to scientists in general or authors of the study where results are reported are only mentioned. Much the same happens *à propos* of material references, and we find a disparity of degrees in the level of details provided. This can be from the concise *three new scientific studies, a BBC poll* or *new results*, to the explicit *the report in Proceedings of the National Academy of Sciences, published yesterday*, which provides readers with as much information as is necessary to completely identify the circumstances surrounding the linguistic event reported.

_____

Turning our attention now to reporting verbs as one of the most relevant reporting signals found in the whole *Sci_TG Corpus,* we have detected that the number of verbs in non-initial parts of the text is considerably higher than that in initial paragraphs. This may be due to the fact that attribution by means of structures containing a reporting verb are not as economical as others, typically reporting adjuncts with *according to*, and hence they are not primed to appear in the introductory paragraph. Data are homogeneous in this respect, for the only exception of a single case that presents an uncommon structure where the first clause acts as an introduction to the second (a quote), but it is not a report.

As expected, tenses in reporting verbs show a clear tendency towards the past. A vast majority of the reporting verbs are constructed in past tense, and only a small percentage of the verbs are formed in present or future tenses. Other forms, such as those in non-finite clauses or passive structures are very scarce in *Sci_TG Corpus.* However, what is also remarkable is the presence of non-finite constructions which show a similar structure to attribution instances constructed by means of reporting adjuncts (i.e. *according to* structures) and passive constructions which act as resources for the journalist to relegate the attribution source to a secondary mid position.

Tenses in reporting verbs have also been classified as belonging to different categories that depend, on their meaning, as well as their frequency of appearance in the corpus. A total of 77 different reporting verbs have been encountered in the corpus, 20 of which appear more than 5 times. This small group of verbs highlights the specificity of the language of reporting in science popularization

articles, which tends to favour a small group of reporting verbs, followed by a bigger group of verbs which only appear rarely, as hypothesized in Chapter 5.

Say stands out because it is, by far, the most frequent verb on the corpus with 1,702 occurrences in its various tenses and forms. *Say* is considered as neutral verb in reporting a language event in the English language to. In opposition to verbs which carry some form of appraisal, *say* only conveys the meaning of a speaker uttering some words. The fact that *say* is the most frequent verb in the corpus also reinforces the presence of the objectivity that is pursued in this type of text, and at the same time it turns out to be the unmarked option in science popularization articles from *The Guardian* newspaper studied.

Other frequent options are *add* and *believe*, which represent two of the most repeated aspects of journalist's narration: adjoining information to pre-existing one and hypothesizing what the outcomes and future advances could be. *Add* also represents the continuous adjoining of details by the journalist, and *believe* represents the different degree of speaker's certainty towards science, both being non-marked options in their additive-continuative and mindsay functions respectively.

The second group of verbs in the corpus, those whose frequency of appearance is lower than 5, constitute other options. Verbs such as *write, warn, tell* and *suggest* contain metalinguistic information as well as conveying the report of a linguistic event. They describe the way in which the language event happened according to the journalist and provide the reader with some information on the circumstances around it. *Write* informs the reader about a source of information

which has been read rather than heard, while *warn, tell* and *suggest* inform the reader of the speaker's intention when uttering the words. As a consequence, we can say that in popularization articles of *The Guardian* newspaper neutral verbs are primed over those implying a further meaning, and the clear non-marked option in the corpus is the verb *say*.

Additionally, reporting verbs have also been classified according to their meaning, and attending to different pre-existing classifications that have been taken as a basis to categorize them. Verbs have been classified according to their degree of neutrality in first instance, in order to later refine the classification and categorize the verbs in relation to the element of the report to which they mainly refer: speech function, speaker or message. This is only a tentative classification that could only cautiously be extrapolated to different contexts or corpora, as the assignation to each of the groups has been based upon cases which are found in the corpus studied.

7.3. Pointers for future research

Along this study, we have obtained relevant data on the behavior and characteristics of explicit attribution sources found in language event's reports of science popularization articles from *The Guardian* newspaper.

As stated in Albert Einstein's words at the beginning of this work, new questions, new possibilities and new perspectives should always appear in front of

our eyes to feed the advance of science. In this study, we have been able to test out hypothesis and to achieve a good degree of consecution of the objectives established at the beginning of the study. However, many questions have arisen in the process and this study has opened the way to carry out future research.

A comparative analysis between different British publications, or a contrastive analysis between Spanish and British publications could not only widen the perspective provided by this study, and also prove to be fruitful in deepening the patterns identified and inspecting, locally and in greater detail, some of the less frequent elements observed in the corpus, with a bigger-size corpus.

In addition, deeper attention could be paid to reporting nouns, lexical elements which perform the same function as a reporting verb and introduce somebody else's speech. These types of reporting signals, as Thompson (1994: 23) asserts, are present in quotes, being the message carried out by the quote, and rephrases, where the message is carried out by 'that'-clauses, 'wh'-clauses and 'to'-infinitive clauses, as in *Their research comes...* (Sci_TG0612).

Moreover, we have studied common patterns in first paragraphs of science popularization articles, establishing a difference between what is found in initial section and what is, and is not, so commonly found in non-initial sections of the texts. Differences among parts within non-initial sections however, have not been explored. We may as well develop a description of how final paragraphs, as closers of the message conveyed in the text, behave in science popularization articles, or whether other pragmatic structural parts can be distinguished.

_____

Teachers and students of academic writing, and specially those working in the field of journalistic writing in English may benefit from this study and future research along this line, as it provides them with new patterns and characteristics that they can use as models for this type of text. Science popularization articles can now be better described in light of findings provided in this study, with the limits of the applicability of the conclusions reached, which are circumscribed to the type of texts studied with its specific characteristics and restrictions. Therefore, special caution should be used when extrapolating them to different contexts, where the different objectives or planning of the research may make the corpus and analysis used in this work an invalid method of obtaining relevant results.

Works cited

Works Cited

A

Adams Smith, D. E. 1987. "The process of popularization – rewriting medical research papers for the layman: discussion paper", *Journal of the Royal Society of Medicine*, 80: 634-636.

Anderson, Lloyd B. 1986. "Evidentials, Paths of Change and Mental Maps: Typologically Regular Asymmetries". In Chafe, W. and J. Nichols (eds.) *Evidentiality: The linguistic coding of epistemology*. Advances in Discourse Analysis Series, vol. XX. New Jersey: Ablex: 273-312.

B

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

Bakhtin, M. M. 1981. *The dialogic Imagination* (translated by C. Emerson & M. Holoquist) Austin: University of Texas Press.

Bednarek, M. 2006. "Epistemological positioning and evidentiality in English news discourse: A text-driven approach". *Text & Talk. Interdisciplinary Journal for the Study of Discourse* 26(6): 635-660.

Bell, A. 1991. *The language of News Media*. Oxford: Blackwell.

Bell, A. 1995a. "Language and the Media". *Annual Review of Applied Linguistics*, 15: 23-41.

Bell, A. 1995b. "News Time". *Time and Society*, 5: 33-63.

Berber Sardinha, A. P. 1996. "Review of WordSmith Tools". *Computers & Texts* 12: 19-21.

Berry, M. 1975. *Introduction to Systemic Linguistics*, *Vol 1: Structures and Systems.* London: Batsford. (Reprinted by the Department of English Studies, University of Nottingham, 1989.)

_____

Berry, M. 1987. "The function of Place-Names". In T. Turville-Petre and M. Gelling (eds.). *Leeds Studies in English New Series*, *XVIII*. Leeds: University of Leeds.

Berry, M. 1989. "Thematic options and success in writing". In C. Butler, R. Cardwell and J. Channel (eds.). *Language and Literature – Theory and Practice*. Nottingham. University of Nottingham.

Berry, M. 1996. "What is Theme? – Another personal view." Draft of article in Berry et al. 1996.

Berry, M, C. S. Butler, R. P. Fawcett and G. Huang (eds.). 1996. *Meaning and Form: Systemic functional interpretations*. Meaning and Choice in language: Studies for Michael Halliday n°3. Norwood: Abblex.

Biber, D. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D. and E. Finegan. 1989. "Styles of stance in English: lexical and grammatical making of evidentiality and affect." *Text* 9.1. (Special Issue on the pragmatics of affect): 93-124.

Biber, D., S. Johnasson, G. Leech, S. Conrad and E. Finegan. 1999. *The Longman Grammar of Spoken and Written English*. London: Longman.

Brown, G. and G. Yule. 1983. *Discourse Analysis*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Burkett, D. W. 1973. *Writing science news for the mass media*. Houston, Texas: Gulf Publishing Company.

Bybee, J. and S. Fleichman. 1995. *Modality in grammar and discours*e. Amsterdam: John Benjamins.

C

Calaresu, E. 1998. "Le trappole interne alla dicotomia discorso diretto/discorso indiretto". In *Orientamenti pedagogici 45* (vol. 269): 869-886.

Caldas-Coulthard, C.R. 1994 "On reporting reporting: the representation of speech in factual and factional narratives." En Coulthard, M. (ed.). *Advances in written text analysis*. London: Routledge: 295-308.

Calsamiglia, H. 1997. "Divulgar: Itinerarios discursivos del saber." *Quark* 7: 9-18.

Calsamiglia, H. 2003. Popularization Discourse. *Discourse Studies* 5(2): 139-146.

Calsamiglia, H. and C. López Ferrero. 2003. "Role and position in scientific voices: Reported speech in the media." *Discourse Studies* 5(2): 147-173.

Calsamiglia, H. and T. Van Dijk. 2004. "Popularization discourse and knowledge about the genome." *Discourse and Society*, 15(4): 369-389.

Carli, A. and E. Calaresu. 2007. "Language and Science". In M. Hellinger and A. Pauwels. (eds.) *Handbook of language and Communication: Diversity and*

_____

Change. Handbooks of Applied Linguistis, 9. Berlin/ New York:Mouton de Gruyter: 523-551.

Carter , B. 1986. "An examination of Items in Theme Position as an Entry Vehicle for Text Processing of Six banking Articles." MA Dissertation. University of Birmingham.

Cassany, D., C. López, and J. Martí. 2000. "La transformación divulgativa de redes conceptuales científicas. Hipótesis, modelo y estrategias", *Revista Iberoamericana de Discurso y Sociedad,* 2(2): 73-103.

Chafe , W.L. 1976. "Giveness, contrastiveness, definiteness, subjects, topics and point of view". In C. N. Li (ed.). *Subject and Topic*. New York: Academic Press: 25-55.

Chafe, W. 1986a. "Evidentiality in English Conversation and Academic Writing." In Chafe, W. and J. Nichols (eds.). 1986. *Evidentiality: The linguistic coding of epistemology*. Advances in Discourse Analysis Series, Vol. XX. New Jersey: Ablex: 261-272.

Chafe, W. 1986b. "Writing in the perspective of speaking." In C. R. Cooper, and S Greenbaum (eds.). *Studying Writing: Linguistic approaches*. Beverly Hills: Sage: 12-39.

Chafe, W. and Nichols, J. (eds.). 1986. *Evidentiality: The linguistic coding of epistemology*. Advances in Discourse Analysis Series, vol. XX. New Jersey: Ablex.

Channell, J. 1999. "Corpus-Based Analysis of Evaluative Lexis." In Hunston, S. and G. Thompson (eds.). *Evaluation in Text*. Oxford: Oxford University Press: 38-55.

Ciapuscio, G. 1997. "Lingüística y divulgación de ciencia." *Quark*, 7: 19-28.

Collins, H. and M. Scott, 1997. "Lexical Landscaping in Business Meetings." *DIRECT PAPERS* 32. São Paulo and Liverpool: Pontificia Universidade Católica de São Paulo CEPRIL / University of Liverpool AELSU.

Conrad, S. and D. Biber. 1999. "Adverbial Marking of Stance in Speech and Writing." In Hunston, S. and G. Thompson (eds.). *Evaluation in Text*. Oxford: Oxford University Press: 56-73.

Cook, G. 1989. *Discourse*. Oxford: Oxford University Press.

Cornillie, B. 2009. "Evidentiality and epistemic modality. On the close relationship between two different categories." *Functions of Language* 16(1): 44-62.

Cortazzi, M. and Jin, L. 1999."Evaluating Evaluation in Narrative". In Hunston, S. and G. Thompson (eds.). 1999. *Evaluation in Text*. Oxford: Oxford University Press: 102-120.

Coulthard, M. 1994. (ed.). *Advances in written text analysis*. London: Routledge.

Croft, W. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.

_____

D

Daneš, F. 1970. "Functional Sentence Perspective and the Organization of Text." In *Papers of Functional Sentence Perspective: First International Symposium.* Berlin/ New York: Mouton de Gruyter: 106-108.

Daneš, F. 1974. "Functional Sentence Perspective and the Organization of Text." In Daneš, F. (ed.). *Papers of Functional Sentence Perspective.* The Hague: Mouton de Gruyter: 106-128.

Davies, F. 1988 "Reading between the lines: Thematic choice as a device for presenting writer viewpoint in academic discourse." *ESPecialist* 9 (2): 173-200.

Davis, L. R. 1993. "Critical analysis of the popular media and the concept of ideal subject position: Sports Illustrated as case study." *QUEST* Volume 45, Issue 2: 165-181.

Davies, F. 1994. "From writer roles to elements of text: interactive, organisational and topical." In *Reflections on Language Learning.* L. Barbara and M. Scott (eds.). Clevedon, England: Multilingual Matters: 170-183.

Davies, F. 1997. "Marked theme as a heuristic for analyzing text-type, text and genre." In *Applied Languages: Theory and Practice in ESP*. J. Piqué and D. J. Viera (eds.). Valencia: Universitat de Valencia: 45-79.

Downing, A. 2001 "Thematic progression as a functional resource in analysing texts." In M.T. Caneda and J. Pérez. (eds.). *Os estudios ingleses no contexto das novas tendencias*. Circle of Linguistics Applied to Communication, 5. Vigo: Universidad de Vigo: 32-42.

E

Elorza, I. 2010. "A Cross-Cultural Analysis of Popularizations in English and Spanish Newspaper Discourse." *ES. Revista de Filología Inglesa* 31: 85-109.

Elorza, I. 2011. "The encoding of authorial voice in science popularizations in the press: a corpus-based cross-cultural study." Paper presented at *12th International Pragmatics Conference*. Manchester, 3-8 July 2011.

F

Fagan, A. and P. Martín Martín. 2004. "The use of critical speech acts in Psychology and Chemistry research papers." *Ibérica* 8: 125-137.

Fairclough, N. 1989. *Language and Power*. London & New York: Longman

Fairclough, N. 1992. *Discourse and Social Change*. Cambridge: Polity Press.

Fairclough, N. 1995. *Media Discourse*. London: Edward Arnold

Firbas, J. 1964. "On defending the Theme in Functional Sentence Analysis." *Travaux Linguistique de Prague* 1: 267-280.

Firbas, J. 1975. "On existence/appearance on the scene in functional sentence perspective." In *Prague Studies in English*, 16: 47-70.

Firbas, J. 1982. "Has every sentence a theme and a rheme?". In J. Anderson (ed.). *Language, form and linguistic variation*. Amsterdam: John Benjamins: 97-115.

Firbas, J. 1986. "On the dynamics of written communication in the light of the theory of functional sentence perspective." In C.R. Cooper and S. Greenbaum (eds.). *Studying Writing: Linguistic Approaches*. Beverly Hills, CA: Sage: 40-71.

Forey, G. 2002. *Aspects of Theme and Their Role in Workplace Texts*. Unpublished Ph.D., Department of English Language. Faculty of Arts. University of Glasgow.

Forey, G. and G. Thompson (eds.). 2009. *Text type and texture*. In honour of Flo Davies. London/Oakville: Equinox.

Fowler, R. 1991. *Language in the news. Discourse and Ideology in the Press*. London & New York: Routledge.

_____

Francis, G. 1993. "A corpus-driven approach to grammar − principles, methods and examples." In Baker, M., Francis, G. and E. Tognini-Bonelli (eds.). *Text and Technology. In Honour of John Sinclair.* Amsterdam: John Benjamins: 137-156.

Fries, P. H. 1981. "On the status of Theme in English: arguments from discourse." *Forum Linguisticum* 6: 1-38. [Reprinted 1983].

Fries, P. H. 1983. "On the status of Theme in English: arguments from discourse. Reprinted revised form." In J. S. Petöfi and E. Sözer (eds.). *Micro and macro Connexity of texts*. Hamburg: Helmut Buske:116-152.

Fries, P. 1994. "On Theme, Rheme and discourse goals." In Malcolm Coulthard (ed.). *Advances in Written Text Analysis*. London: Routledge and Kegan P: 229-249.

Fries, P. H. 1995. "Themes, methods of development and texts." In R. Hasan and P. H. Fries (eds.). *On Subject and Theme: From the Perspective of Functions in Discourse*. Amsterdam: John Benjamins: 317-359.

Fries, P. H. 2009. "The textual metafunction as a site for a discussion of the goals of linguistics and techniques of linguistic analysis." In Forey, G. and G. Thompson (eds.). *Text type and texture. In honour of Flo Davies.* London/Oakville: Equinox: 8-44.

_____

G

García Negroni, M. M. 2008 "Subjetividad y discurso científico-académico. Acerca de algunas manifestaciones de la subjetividad en el artículo de investigación en español." *Signos* 41(66): 5-31.

García Riaza, B. 2010a. "Explicit attribution sources expressed by 'according to' in a corpus of science popularization articles from *The Guardian*." *Interlingüística XXI*: 157-168.

García Riaza, B. 2010b. "Mapping Attribution across Texts: the particle 'according to' in science popularizations and editorials from *The Guardian.*" Paper presented at *Mapping Language across Cultures Conference (MLAC10)*, Salamanca, 5-7 July 2010.

García Riaza, B., Nencioni, A., Eorza. I., Iglesias, N. and M. Pérez Veneros. 2011. "Cross-cultural analysis on the discourse of science popularizations: language event's reports in Spanish, English, Italian and German." *Proceedings of the XXIX AESLA Conference.* Salamanca 4-6 May, 2011.

García Riaza, B. and M. Pérez Veneros. (in press). "'Los científicos afirman que…': Marcas explícitas de discurso referido en encabezados de artículos de divulgación científica de la prensa española y británica." *Interlingüística XXII*. Ediciones Universidad de Salamanca.

García Riaza, B. and I. Elorza. (in press). "The emergence of external sources of attribution in science popularizations in *The Guardian.*" In Molina Plaza,

S. and A. Roldán (eds.). *NEW PERSPECTIVES ON LSP: The contribution of genre research.*

Garret, P. and A. Bell. 1998. *Approaches to Media Discourse*. Oxford: Blackwell.

Gianoni, D. S. 2008. "Popularizing features in English journal editorials." *English for Specific Purposes,* 27: 212-232

Gil Salom, L. 2000-2001. "El discurso de la ciencia y la tecnología: el artículo científico de investigación *vs.* el artículo de divulgación científica." *RESLA*, 14: 429-449.

Goldacre, B. 2009 *Bad Science: The Short Life and Weird Times of Cold Fusion*. London: Fourth Estate.

Gosden, H. 1993. "Discourse Functions of Subject in Scientific Research Articles". In *Applied Linguistics*, Vol. 14, nº 1. Oxford: Oxford University Press: 56-75.

H

Halliday, M. A. K. and R. Hasan. 1976. *Cohesion in English*. London & New York: Longman.

Halliday, M. A. K. 1895. *An Introduction to Functional Grammar*. London: Edward Arnold.

_____

Halliday, M.A.K. and J.R. Martin. 1993. *Writing Science: Literacy and Discursive Power.* Pittsburgh: University of Pittsburgh Press.

Halliday, M.A.K. 2004. *The language of Science*. Jonathan J. Webster (ed.). London and New York: Continuum.

Halliday M.A.K. and Christian M.I.M. Matthiessen. 2004. *An introduction to functional grammar.* 3rd revised edition. London: Edward Arnold.

Hartley, J. 1982. *Understanding News*. London: Methuen.

Hernando Cuadrado, L. A. and Hernando García-Cervigón, A. 2006. *Lengua y comunicación en el discurso periodístico de divulgación científica y tecnológica*. Madrid: Fragua.

Ho-Dac, L. M. 2008. "Discourse organisation though Theme position." In Steiner, E. and Neumann, S. (eds.). *Proceedings of the Nineteenth European Systemic Functional Linguistics Conference and Workshop*. Saarbrücken: Germany. Available at:

http://scidok.sulb.uni-saarland.de/volltexte/2008/1691/.

Hoey, M. and M. O'Donnell. 2008 "Lexicography, Grammar, and Textual Position." In *International Journal of Lexicography*, Vol. 21 No. 3: 293-309.

Hoey, M. 1993. (ed.). *Data, Description, Discourse*. London: Harper Collins.

Hoey, M. 2005. *Lexical Priming*. London: Routledge.

Huddleston, R. 1991. "Further Remarks on Halliday's Functional Grammar: A Reply to Matthiessen & Martin." *Occasional Papers in Systemic Linguistics* 5: 75-129.

Hunston, S. 1989. *Evaluation in Experimental research articles* PhD Thesis. University of Birmingham.

Hunston, S. 1994. "Evaluation and organization in academic discourse." In M. Coulthard (ed.). *Advances in written text analysis*: 191-218

Hunston, S. 1999."Evaluation and the Planes of Discourse: Status and Value in Persuasive Texts." In Hunston, S. and G. Thompson (eds.). *Evaluation in Text*. Oxford: Oxford University Press: 176-205.

Hunston, S. and G. Thompson. 1999. *Evaluation in Text*. Oxford: Oxford University Press.

Hunston, S. and G. Francis. 2000. *Pattern Grammar. A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: John Benjamins.

Hyde, J. R. 1999 "Lo demás es literatura" –Predicción y adversatividad en textos argumentativos en inglés y español." En Fernández González, J., Fernández Juncal, C., Marcos Sánchez, M., Prieto de los Mozos, E. y Santos Río, L. (eds*.). Lingüística para el Siglo XXI. Vol. II*. Salamanca: Ediciones Universidad de Salamanca: 935-941.

Hyland, K. 1996. "Writing without conviction: hedging in science research articles." *Applied Linguistics* 17(4): 433-454.

Hyland, K. 2000. *Disciplinary discourses*. New York: Longman.

Hyland K. 2010. "Constructing proximity: relating to readers in popular and professional science." *Journal of English for Academic Purposes*, doi:10.1016/j.jeap.2010.02.003.

I

Irwin, A. and B. Wynne (eds.). 1996. *Misunderstanding Science*. Cambridge: Cambridge University Press.

K

Kenward, M. 1988. "Parasites in the press." *Parasitology Today*, 4(2): 31-32.

Kim C. K. and G. Thompson. 2010. "Obligation and Reader Involvement in English and Korean Science Popularizations: A Corpus-based Cross-cultural Text Analysis." *Text & Talk* 30(1): 53-73.

L

Labov, W. 1972. *Language in the Inner City*. Philadelphia: University of Pennsylvania.

Labov, W. 1984. "Intensity." In D. Schiffrin (ed.). *Meaning, Form and Use in Context: Linguistic Applications*. (Georgetown University Roundtable on

Language and Linguistics). Washington, D. C.: Georgetown University Press: 43-70.

Lavob, W. 1972. "Hedges: a study in meaning criteria and the logic of fuzzy concepts." *Proceedings of the Chicago Linguistics Society* 8: 183-228.

Leech, G. 1992. "Corpora and Theories of Linguistic Performance." In Svartvik, J. (ed.). *Directions in Corpus Linguistics. Proceedings of the Nobel Symposium 82, Sockholm, 4-8 August 1991*. Berlin and New York: Mouton de Gruyter: 105-122.

Leech, G. 1983. *The principles of Pragmatics*. London: Longman.

Li, C. N. (ed.). *Subject and Topic*. New York: Academic Press.

Ljung, M. 2000. "Newspaper Genres and Newspaper English." In Friedrich Ungerer (ed.). *English Media Texts, Past and Present*. Pragmatics & Beyond New Series, nº 80. Amsterdam/Philadelphia: John Benjamins.

M

Mahlberg, M and M. O'Donnell. 2008. "A fresh view of the structure of hard news stories." In Steiner, E. and S. Neumann (eds.). *Proceedings of the Nineteenth European Systemic Functional Linguistics Conference and Workshop*. Saarbrücken: Germany. Available at http://scidok.sulb.uni-saarland.de/volltexte/2008/1700/

_____

Mahlberg, M., M. O'Donnell, M. Hoey & M. Scott. 2007 "Exploring the textual positions and functions of lexical items in hard news stories." Paper given at the *4th Corpus Linguistics Conference*. University of Birmingham.

Mahlbergh, M. 2005. *English General Nouns. A corpus theoretical approach.* Studies in Corpus Linguistics n° 20. Amsterdam/Philadelphia: John Benjamins.

Maingueneau, D. 1999. "Ethos, scénographie, incorporation." En R. Amossy (ed.). *Images de soi dans le discours. La construction de l'éthos.* Paris: Delachaux et Niestlé: 75-102.

Maldonado González, C. 1999. "Discurso directo y discurso indirecto." En Bosque y Demonte (eds.). *Gramática descriptiva de la Lengua Española.* Madrid: Espasa Calpe S.A. - Real Academia Española: Colección Nebrija y Bello: 3549-3595.

Martin, J.R. 1992. *English text. System and Structure.* Philadelphia: John Benjamins.

Martin, J. R. 1999. "Beyond Exchange: APPRAISAL Systems in English." In S. Hunston and G. Thompson (eds.). *Evaluation in Text*. Oxford: Oxford University Press: 142-175.

Martin, J. R. and D. Rose. 2003. *Working with discourse: meaning beyond the clause*. London/N.Y.: Continuum.

Martin, J. R. and P. R. R. White. 2005. *The language of Evaluation. Appraisal in English*. London & New York: Palgrave Macmillan.

Mathesius, V. 1939. "On so-called functional sentence perspective." *Slovo a Slovensnost* 7: 169-180.

333

_____

McCabe, A and K Heilman. (2007). "Textual and Interpersonal Differences between a News Report and an Editorial." *Revista Alicantina de Estudios Ingleses*, 20: 139-156.

McEnery, T, and A. Wilson. 2001 [1996]. *Corpus Linguistics*. Edinburgh Textbooks in Empirical Linguistics. Edinburgh: Edinburgh University Press.

Minelli de Oliveira and Pagano, A. 2006. "The research article and the science popularization article: a probabilistic grammar perspective on direct discourse representation." *Discourse Studies* 8(5): 627-646.

Moirand, S. 1997. "Communicative and cognitive dimensions of discourse on science in the French mass media." *Discourse Studies* 5(2): 175-206.

Montemayor-Borsinger, A. 2003. "A comparison of thematic options in novice and expert research writing." In *Estudios de Lingüística Aplicada*, 21 (037). Distrito Federal, México: Universidad nacional Autónoma de México: 33-51.

Montemayor-Borsinger, A. 2009. "Text-type and texture: the potential of Theme for the study of research writing development." In Forey, G. and G. Thompson (eds.). *Text type and texture. In honour of Flo Davies.* London/Oakville: Equinox: 108-124.

Mushin, I. 2001. *Evidentiality and Epistemological Stance. Narrative retelling*. Amsterdam/Philadelphia: John Benjamins.

Myers, G. 1990. "Making a discovery: narratives of split genes." In Nash, C. (ed.). *Narrative in Culture. The Uses of Storytelling in the Sciences, Philosophy, and Literature*. London: Routledge: 102-126.

Myers, G. 1994. "Narratives of science and nature in popularizing molecular genetics." In M. Coulthard (ed.). *Advances in Written Text Analysis*. London: Routledge: 179-190.

Myers, G. 2003. "Discourse studies of Science Popularization: Questioning the Boundaries." *Discourse Studies* 5 (2): 265-279.

N

Ndahiro, A. 1998. *Theme, Rheme, Given and New in written Discourse: Evidence from Annual Business Reports*. Unpublished PhD Thesis. University of Liverpool.

Niemeyer, S. and R. Dirven (eds.). 1997. *The language of emotions: Conceptualisation, Expression, and Theoretical Foundation*. Amsterdam: John Benjamins.

Nuyts, J. 2001. *Epistemic modality, language and conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins.

O

O'Keeffe K. J. 2002. "The Quality Press in Ireland: A Linguistic Analysis of Two Text-Types." In A. Sánchez Macarro (ed.). *Windows on the World: Media Discourse in English*. Valencia: Universitat de València: 239-263.

Ochs, E. and B. Schiefflen. 1989. "Language has a heart." *Text* 9(1). (Special Issue on the pragmatics of affect): 7-25.

Ozturk, I. 2007. "The textual organization of research article introductions in applied linguistics: variability within a single discipline." *English for Specific Purposes* 26(1): 25-38.

P

Parodi, G. 2007 "Specialized written discourse at university and professional domains: Composition of a corpus." *Revista Signos,* 40(63): 147-178.

Perfetti, C. A. and S. R. Goldman. 1975. "Discourse Functions of Thematization and Topicalization." In *Journal of Psycholinguistc Research* Vol. 4, Nº 3: 257-271.

Piazza, R. 2009 "News is reporting what was said: techniques and patterns of attribution." In L. Haarman and L. Lombardo (eds.). *Evaluation and*

_____

stance in war news. A linguistic Analysis of American, British and Italian television news reporting in the 2003 Iraqui war. Research in Corpus and Discourse Series. Wolfgang Teubert and Michaela Mahlberg (eds.). London/N.J.: Continuum.

Pounds, G. 2010. "Attitude and subjectivity in Italian and British hard-news reporting: The construction of a culture-specific 'reporter' voice." *Discourse Studies*, 12(1): 106-137.

R

Reah, D. 2002. *The Language of Newspapers*. Abingdon: Routledge.

S

Scott, M. 2004-2007. *WordSith Tools. User's manual.* Oxford: Oxford University Press.

Scott, M. 2004-2005. *WordSmith Tools, Version 4.0*. Oxford: Oxford University Press. [Software].

Scott, M. 1997. "PC Analysis of Key Words –And Key Key Words." *System*, 25(2): 233-245.

Searle, J. R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Shaw, P. 1992. "Reasons for the correlation of voice, tense and sentence function in reporting verbs." *Applied Linguistics* 13: 302-319.

Sinclair, J. and M. Coulthard. 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* Oxford: Oxford University Press.

Sinclair, J. McH. 1981. "Planes of discourse." In S. N. A. Rizvi (ed.). *The Two-Fold Voice: Essays in Honour of Ramesh Mohan*. Salzburg: University of Salzburg: 70-89.

Sinclair, J. McH. 1986. "Fictional words." In R.M. Coulthard (ed.). *The Two-fold Voice*. Essays in Honour of Ramesh Mohan. Saltzburg: University of Saltzburg: 70-89.

Sinclair, J. McH. 1988. "Mirror for a text." *Journal of English and Foreign Languages* 1: 15-44.

Sinclair, J. McH. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. McH. 1994. "Trust the text." In M. Coulthard (ed.). *Advances in written text analysi*s. London: Routledge:12-25.

Sinclair, J. McH. 2004. *Trust the Text*. London and New York: Routledge.

Stubbs, M. 1993. "British Traditions in Text Analysis. From Firth to Sinclair." In M. Baker, G. Francis and E. Tognini-Bonelli (eds.). *Text and Technology. In Honour of John Sinclair.* Amsterdam: Benjamins: 1-33.

Stubbs, M. 1996. *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture*. Oxford and Cambridge, Mass.: Blackwell.

Swales, J. 1990. *Genre analysis*. Cambridge: Cambridge University Press.

T

Tadros, A. A. 1985. *Prediction in text*. University of Birmingham: English
    Language Research.

Tadros, A. A. 1993 "The pragmatics of text averral and attribution in academic
    texts." In M. Hoey (ed.). *Data, Description, Discourse.* London: Harper
    Collins.

Thetela, P. 1997. "Evaluated Entities and Parameters of Value in Academic
    Research Articles." *English for Specific Purposes*, 16(2):101-118.

Thomas, S. and T. Hawes. 1997. *Theme in Academic and Media Discourse*.
    Monographs in Systemic Linguistics, 8. Nottingham: University of
    Nottingham.

Thompson, G. 1994. *Collins Cobuild English Guides 5: Reporting*. London:
    Harper Collins.

Thompson, G. 1996a. *Introduction to Functional Grammar*. [imp. 2004] London:
    Arnold.

Thompson, G. 1996b. "Voices in the text: discourse perspectives on language
    reports." *Applied Linguistics* 17(4): 501-530.

Thompson, G. 2001. "Interaction in Academic Writing. Learning to argue with the reader." *Applied Linguistics* 22(1):58-78.

Thompson, G. and S. Thompson. 2009. "Theme, Subject and the unfolding of the text." In G. Forey, and G. Thompson (eds.). *Text type and texture. In honour of Flo Davies.* London/Oakville: Equinox: 45-69.

Thompson, G. and Y. Ye. 1991. "Evaluation in the reporting verbs used in academic papers." *Applied Linguistics* 12(4): 365-382.

Thompson, G. and Thetela, P. 1995. "The sound of one hand clapping: The management of interaction in written discourse." *Text & Talk. Interdisciplinary Journal for the Study of Discourse* 15(1): 103-128.

Thomson, G. and Zou, J. 1999. "Evaluation and Organization in Text. The Structuring Role of Evaluative Disjuncts." In S. Hunston, and G. Thompson (eds.). *Evaluation in Text*. Oxford: Oxford University Press:121-141.

Tognini-Bonelli, E. 2001. "Corpus Linguistics at Work." *Studies in Corpus Linguistics* 6. Amsterdam/Philadelphia: John Benjamins.

V

van Dijk, T. 1985. *Discourse and communication: new approaches to the analysis of mass media discourse and communication*. Berlin/N.J.: Walter de Gruyter.

_____

van Dijk, T. 1988. *La ciencia del texto: un enfoque interdisciplinario*. Barcelona: Paidós. [2ª Ed.]

van Dijk, T. 1993. *Elite discourse and racism.* Newbury Park, CA: Sage.

van Dijk, T. 1998. *News as Discourse*. Hillsdale, NJ: Erlbaum.

van Dijk, T. 2009. *Discurso y Poder*. Contribuciones a los estudios críticos del discurso. Barcelona: Gedisa.

Vande Kopple, W. J. 1991. "Themes, Thematic Progressions, and some Implications for Understanding Discourse." In *Written Communication*, 8: 311-347.

W

Whageng, P. 1997. "Keywords in Theme and Rheme." In *Language and Discourse*, 5: 43-69.

Wierzbicka, A. (ed.). 1990. *Australian Journal of Linguistics* 10(2). (Special Issue on the semantics of emotions).

Williams, J. 2009. "Variation of cancer metaphors in scientific texts and press popularizations." In *Proceedings of the fifth Corpus Linguistics Conference*. University of Liverpool. July 2009.
Available at http://www.cs.bham.ac.uk/~amw/pdfVersions/Williams.pdf

Z

Zapaterra, Y. 2008. *Diseño Editorial. Periódicos y revista*s. Barcelona: Gustavo Gili. [Translation from Zapaterra, Y. 2007. *Editorial Design: for Print and Electronic Media*. London: Lawrence King].

Zipf, G. K. 1935. *The Psychobiology of Language*. Cambridge, Mass.: MIT Press.

Zou, J. L. 1991. *The functions of disjuncts in the organization of written English discourse.* Unpublished M.A. dissertation. University of Liverpool.

Annex

Annex 1

ABSTRACT

HEADLINE

LEAD

G20: David Cameron warns world must tackle east-west trade imbalances

UK prime minister fears return to protectionism and trade barriers of Great Depression era

Patrick Wintour in Seoul
guardian.co.uk, Thursday 11 November 2010 09.23 GMT
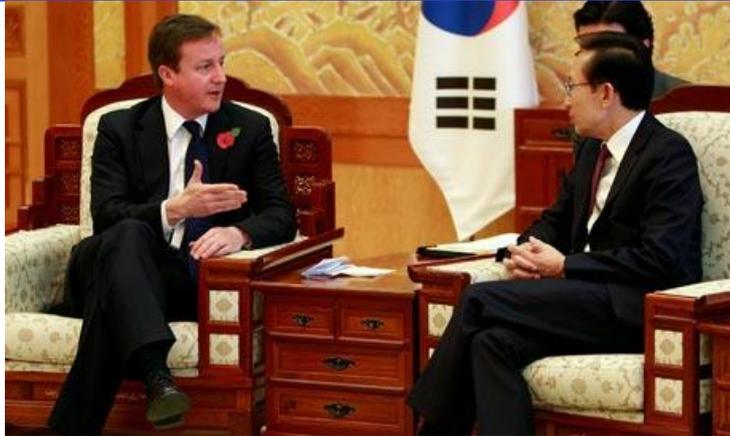
The east's cash versus the west's debt is the 'biggest issue' of the G20 says David Cameron. Photograph: Chung Sung-Jun/Getty Images

STORY

EPISODE 1

EVENT 1

EVENT 2

EVENT 3

World leaders must resolve the trade imbalances that are destabilising the global economy or risk a repeat of the Great Depression, David Cameron warned today.
Speaking at the start of the G20 Summit in Seoul, the prime minister said the real test for the summit would be its ability sort out the big battle between the indebted west and the surplus economies of the east.
"The fear we should all have is a return to what happened in the 30s: protectionism, trade barriers, currency wars, countries pursuing beggar my neighbour policies – trying to do well for themselves but not caring about the rest of the world. That is the danger," Cameron said.

"Now on the big battle, the biggest issue of all is the cause of the last crisis – a wall of money in the east, a wall of debt in the west. We've got to deal with that imbalance and I think it's a real test for this summit and one that Britain will play a very positive part in trying to make sure we really look at these imbalances and deal with them."

EPISODE 2

EPISODE 3

EVENT 1

EVENT 2

The G20 summit is being dominated by concerns that currency wars will break out between the world's largest economies. America is demanding that China allows its currency to appreciate, while the Chinese insist that any such move will create mass joblessness in China and so dampen worldwide growth.
US president Barack Obama and Chinese premier Hu Jintao were holding talks earlier today in Seoul. Before the meeting started, Obama said that the US and China were "making progress" on economic issues.

EPISODE 4

EVENT 1

EVENT 2

Obama's treasury secretary, Tim Geithner, told CNBC this morning that China would eventually have to yield to pressure from the financial markets and allow the yuan to rise in value.
"Those market forces are just a reflection of confidence that you're going to see strong growth in China, strong productivity growth in China, if you resist those market forces, that pressure is not going to go away," Geithner said.

"It's just going to end up in higher inflation or higher asset prices and that'll be bad for China."

EPISODE 5

Julia Gillard, prime minister of Australia, warned her fellow leaders to avoid descending into currency wars, in which country's would devalue their own currencies to drive exports.

EPISODE 6

"Australia believes currency should be market-based, that is our position. We should be trying to ensure how we move from one stage of global growth to another," Gillard said.

EPISODE 7