

THE EMERGENCE OF EXTERNAL SOURCES OF ATTRIBUTION IN A PILOT STUDY OF SCIENCE POPULARIZATIONS FROM *THE GUARDIAN* NEWSPAPER

BLANCA GARCÍA RIAZA*
IZASKUN ELORZA
Universidad de Salamanca

ABSTRACT. *The discourse of news media in English has traditionally been an attractive and fruitful source of information for discourse analysts and corpus linguists. However, studies on science popularizations as a text type of newspaper discourse are still scarce. What we know often comes from contrastive analyses between scientific texts (typically academic research articles) and science popularizations from newspapers (Hyland 2010; Myers 1990, 1994; Skorczynska 2001), and yet popularizations present features typically associated to newspaper discourse. The research presented here is a pilot study on how external sources of attribution are introduced in the text in popularizations from The Guardian. The presence of external sources of attribution has been scrutinized in relation to the thematisation of elements from the perspective of Systemic Functional Grammar (Halliday 1985; Thompson 2004). The identification of patterns of how attribution is introduced in this text type has provided insight of how information is constructed in science popularizations and the crucial role of attribution sources.*

KEY WORDS. *Thematisation, science popularisations, attribution sources.*

RESUMEN. *Tradicionalmente, el discurso de las noticias en inglés ha resultado una valiosa y productiva fuente de información para los analistas del discurso y de lingüística de corpus. Sin embargo, este tipo de estudios es aún limitado. Lo ya conocido suele proceder de análisis contrastivos entre textos científicos (típicamente artículos de investigación académicos) y artículos periodísticos de divulgación científica (Hyland 2010; Myers 1990, 1994; Skorczynska 2001), aunque los artículos de divulgación presentan rasgos típicamente asociados al discurso periodístico. Presentamos aquí un estudio piloto sobre cómo se introducen las fuentes externas de atribución en artículos de divulgación publicados en The Guardian. La presencia de fuentes externas de atribución se ha examinado respecto a la tematización de elementos desde la Gramática Sistemico-Funcional (Halliday 1985; Thompson 2004). La identificación de patrones de cómo la atribución se introduce en este tipo textual proporciona una profunda visión de cómo la información se construye en los artículos de divulgación científica y el papel crucial de las fuentes de atribución.*

PALABRAS CLAVE. *Tematización, divulgación científica, fuentes de atribución.*

1. INTRODUCTION

The discourse of news media in English has traditionally been an attractive and fruitful source of information about language for both discourse analysts and corpus linguists, mainly due to the possibility for analysts of building their own corpora (O’Keeffe 2002: 260) as well as to the availability of large corpora of newspaper text already compiled (Scott 1997).

As media discourse covers a great variety of linguistic and social phenomena, news genres have been and are analysed from a huge variety of perspectives and approaches, which in the case of popularizations may range from an analysis of how biotechnological news are framed in the media (Marks et al. 2007) to the study of metaphors of cancer in the British press (Williams Camus 2009). However, even if extensive research has been carried out, we still lack an appropriate description of them, especially in terms of their linguistic features.

Newspaper genres are sensitive to social changes and popularizations are a good example of this. Popular science started to draw the attention of linguists by the end of the nineteen eighties, and especially after the analyses carried out by Myers (1990), because of its relationship to scientific discourse and their potential to compare different recontextualizations of scientific knowledge. From the perspective of the academia, the interest in popularizations is also due to scientists’ awareness of the need to present their findings to the public and not to hide “behind an impenetrable screen of jargon and arrogance” (Kenward 1988: 31).

However, as Hyland highlights, it would be misleading to assume that popularizations address a homogeneous audience:

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).

Science popularization articles, widely known as popularizations, can be taken as a genre within scientific discourse characterized by its audience, who is called “the layman” by Adams Smith (1987: 636) and “the average citizen” by Calsamiglia and López Ferrero (2003), and also by its features, which can be expected to reflect the textual features that are conventional practice at a given community (Kim and Thompson 2010: 54) but which may not be conventional in academic research articles (Gil Salom 2000-2001: 430). To obtain a reliable description of those typical features is the general aim of linguistic analyses and descriptions of this genre.

Among other features, Adams Smith (1987: 634) describes how, unlike what happens in academic research articles, in popularizations the topic takes second place to editorial treatment, so that a new context is frequently added in order to frame the information from a different point of view. In a contrastive study carried out by Myers

(1990, 1994) a further distinction is made. Myers argues that, in popular science, facts are endowed with an authority which is 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). This narrative style is called by Myers a ‘narrative of nature’. In contrast, he finds that the narrative style of most scientific research reports emphasizes “the concepts and techniques through which the scientist conceives of and delimits nature”, calling this type of narrative style a ‘narrative of science’. Other features described include that, in popularizations, there are typically two main kinds of actors (scientists and organisms) and that the stories narrated are about discoveries involving observation in the field or laboratory. However, Myers does not go further to describe differences between the role played by the two different kinds of actors involved.

From a qualitative perspective we are, therefore, dealing with texts which omit most of the detail of the materials, methods and technical procedures which are found in scientific research articles; texts which are written in a less formal tone in order to be more appealing to the newspaper readers. Popularizations also show patterns which include the presentation of findings as a breakthrough, typically by constructing a rhetorical contrast between past and present, and the exemplification of a Problem-Solution pattern which is made explicit in the headline (Adams Smith 1987: 635).

In addition, from a quantitative perspective Adams Smith adds that popularizations tend to be shorter than their corresponding source articles. It is clear that a finer-grained description should be desirable in order to provide a more accurate picture of those genres, so the relevant information here is not whether popularizations are shorter than research articles but rather how shorter or what their typical length may be. Empirical analyses aiming at attested descriptions show the underlying problems we must face when trying to answer these questions, which may be illustrated with two examples.

Fagan and Martín Martín (2004) have analyzed contrastively research articles of psychology and of chemistry, finding that chemistry texts tend to be much longer than psychology ones (3,000 and 1,100 words per article respectively). Those data show that the length of research articles varies too much between disciplines for an average length to be established. In addition, Ozturk (2007) has compared the relative length of research articles in two journals within the same discipline (applied linguistics) and has obtained data not only suggesting that 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) but also that the variability *within* each journal is also very high, with lengths ranging from 426 to 2,429 words in the journal *Studies in Second Language Acquisition* and from 145 to 1,771 words in the *Journal of Second Language Writing*. More recently, Elorza (2010) has developed a cross-cultural comparison of two science popularizations, one in Spanish and the other in English, which narrate the same scientific finding, getting to the conclusion that, even within the same genre (science popularization articles as published in comparable newspapers), these texts are highly dependent on the strategies used by each writer in order to achieve their communicative

goals. A consequence of this for further analysis is that a perspective which takes into account the fact that the writers of these texts approach their production from the perspective of the newsroom rather than from the perspective of the scientist seems to be more fruitful for explaining their features.

In our corpus of popularizations, the average number of words per article is 470. However, considering the extreme variability of length in research articles, it seems more productive to compare this figure to other newspaper genres. Comparing the relative length of texts in twelve different sections as distinguished by editors and journalists of the *Guardian* newspaper, Mahlberg (2007: 207) quotes an average of 706 words per article. If we discard the section of ‘Letters to the Editor’, which contains texts produced by readers of the newspaper, the relative average in each section ranges from 547 words of the ‘Overseas News’ section to 1,590 words of the ‘Features: Weekend’ section. In view of these data, our popularization texts are not only supposedly shorter than their corresponding scientific articles, but also shorter when compared to most of the other genres within the same newspaper. This relative shortness is relevant to the aims of our study because it is reasonable to expect that the textual features that are conventional practice in popularizations will be restricted by a limit in length availability which may impose some kind of specialization of the typical patterns of this kind of narration, not just in the frequency of their use, but especially in their distribution in text.

2. TEXT-INITIAL PARAGRAPHS IN POPULARIZATIONS

Popularization articles start with a headline consisting of a noun phrase (e.g. “*Genetic mutation linked to depression*”) or a clause either finite or non-finite.

The headline is sometimes followed by a lead (24 % of the texts in our corpus), e.g. “*Researchers find inheritance factor in succumbing to life’s woes*”. After this, the narration starts with an opening paragraph which typically consists of one single sentence (more than 75 % in our corpus). What is relevant here is that the source of attribution of the news story is typically introduced in the text in this first paragraph (about 90 % in our corpus), although sometimes it has already been introduced in the headline (about 28 % cases in our corpus).

3. CORPUS DESCRIPTION

The study we are presenting in this article focuses on science popularization articles in *The Guardian* newspaper, as samples of this genre in English language. This newspaper was chosen because it is a quality broadsheet aimed at a relatively educated audience in Great Britain (Kim and Thompson 2010: 62), so that popularizations are regularly published, and also because it is often used for descriptive studies of newspaper discourse in English (e.g. Hyde 1999).

A small corpus of science popularization articles (*Sci_TG_Pilot Corpus*) (about 12,000 running words) has been compiled from the electronic version of the newspaper (www.guardian.co.uk) from the years 2003-2007, which are taken as samples of the popularizations published during those years. The retrieval of the texts has been carried out by means of the search-engine provided by the website of *The Guardian*, the main criterion established for the compilation of the corpus being to ensure the homogeneity and reliability of the corpus.

The expected procedure for retrieving popularizations from *The Guardian* website is typically to search the Science section and download the printable version of the selected texts (to further convert them into an only-text format for analysis by means of some kind of corpus-processing software). However, as a preliminary scrutiny showed, this section does not only contain popularization articles but also other kinds of text, such as press releases and readers' contributions in blogs, so up to four of the texts previously retrieved had to be discarded as they did not belong to the target text type.

It had been set in the planning of the design of the study that a sample of texts would be appropriate to detect patterns of attribution, as this phenomenon is pervasive. Therefore, our aim for the compilation was that all the texts retrieved should be representative of the genre as realized in *The Guardian* newspaper. For this reason, the first criterion for the text-selection process was based on the authorship of the texts, by means of which only texts authored by scientific reporters were compiled. Although in several newspapers, such as *El País*, there are no indications as to whether the author of a popularization article is a scientist or a reporter, *The Guardian* newspaper includes, together with the name of the author, information about this. Therefore, we compiled only texts which are authored by a science correspondent (such as those signed by Alok Jha) or by a science editor (such as Tim Radford).

4. OBJECTIVES

Our research aims to:

- Get information about the characteristics of first paragraphs in science popularizations in English.
- Gain a deeper understanding of the patterns which are typical in first sections of science popularizations in English.
- Obtain information about the thematisation of elements in first paragraphs of science popularization articles.
- Retrieve data about the presence (or not), frequency and distribution of explicit attribution sources in the corpus studied.
- All these, with the ultimate goal of revealing how attribution sources emerge in the narration of scientific popularizations in the British press as represented by *The Guardian* newspaper.

5. RESEARCH METHODOLOGY

The methodology used in this study gathers techniques from both discourse analysis and corpus linguistics, taking benefits from the combination of two complementary approaches to the study of language. In this way, the analysis carried out has considered both quantitative and qualitative data.

Once the corpus had been compiled, a manual scrutiny of the texts was carried out in order to get familiarized with how the narration is constructed in the reporting of the scientific findings in *The Guardian* newspaper. As our main concern in the study was the phenomenon of attribution, the emphasis of the intensive manual analysis carried out was placed in this. This analysis allowed us to detect a certain degree of variation in the structural composition of the texts, observing that some –but not all– contain a lead after the headline, which affects to when the external source of attribution will be introduced in the text. In this phase it was also possible to observe that first paragraphs tend to be short, so that when an external source of attribution is introduced here, this is done economically or, in other words, it seems that options employing very few words are preferred by the authors.

A second phase of the study focused, then, on the study of the resources employed for introducing the external sources of attribution, as well as on the analysis of their distribution within the paragraph, so that we could study the phenomenon of attribution in relation to thematisation as well. In this phase, a quantitative approach was followed in order to identify which patterns had been employed more frequently by the authors of the popularizations to refer to external sources of attribution. To this purpose, the corpus processing software *WordSmith Tools v.4.0* (Scott 2004-2005) was employed, enabling us to obtain concordance lines which provided data on the distribution of explicit attribution sources and their positions within sentence structure. In the following sections, these results are presented and discussed.

6. RESULTS

In this section, we present and discuss the results of the research in two different parts: characteristics of the initial sections of science popularization articles, and patterns encountered in them. These parts also bear a direct correspondence with the stages in which the analysis has been carried out.

6.1. *The characteristics of initial sections in science popularizations in English*

Initial elements in newspaper articles fulfill a special role that is reflected in the textual structure as well. As Reah (2002: 16) points out, headlines consist of words which are not chosen merely to convey meaning economically, but also with the persuasive function of attracting the readers' attention. In this sense, they may be considered "relevance optimizers" (Dor 2003). Besides, headlines and leads normally condense the

whole point of the story (Bell 1991: 149), establishing a departure point for the narration in the text body. On their part, leads are described as “packed with information and news appeal, but as short as possible and clearly understood” (Bell 1991: 176), thus acting as bridges between the headlines and the texts.

Taking into account this introduction to the different functions of initial parts of science popularization texts, we have observed real samples in our corpus to look for shared patterns or characteristics. We have been able to identify two different types of texts, according to the initial sections they present.

On the one hand, a group of 6 articles (28.57 %) present a lead, which connects the headline and the first paragraph, also setting the tone of the article when informing the reader about its purpose (Zapatero 2008). Three of the articles assigned to this first group contain concise and short leads that express the most relevant information on the scientific finding through a finite proposition, behaving as a kind of initial summary that precedes the first paragraph of the texts, performing the same function as the abstract in a personal narrative (Bell 1991: 149). Leads in this first group are also characterized by their propositional character, not being ended by a full stop punctuation mark, as we can see in Example 1 below.

```
(1) 18 July 2003, The Guardian, p.9
<H> Genetic mutation linked to depression </H> / Home News
Tim Radford Science editor
<L> Researchers find inheritance factor in succumbing to life's
woes </L>
<P1> A common mutation in a single gene could make the difference
between fighting back against life's assaults and sinking into
clinical depression, according to research out today. </P1>
```

The remaining 3 texts containing a lead show the special feature of having a lead structured as an outline, i.e. a *bulleted lead* that introduces the topic not only summing up the information but also evaluating it, as can be seen below:

```
(2) 05 October 2005, The Guardian, p.12
<H> Science: Alcohol abusers should not get transplants, says
Best surgeon </H>
<L1> Liver shortage means help should be more targeted </L1>
<L2> Support needed to stop relapses into drinking / Home News </L2>
Sample, Ian
<P1> The surgeon who performed George Best's liver transplant
says urgent measures are needed to identify patients who are likely
to abuse alcohol after their operations, so they can be kicked off
hospital waiting lists. </P1>
```

On the other hand, in those cases where there is no lead (71.43 % of the articles) it is the first paragraph which acts as an introductory element instead, as can be seen in Example 3 below:

(3) 13 February 2004, *The Guardian*, p.6
 <H> Psychologist says maths can predict chances of divorce/Home
 News </H>
 Tim Radford Science editor
 <P1> A psychologist claims that a newly devised mathematical
 model can predict with 94% accuracy which couples will divorce -
 entirely on the basis of the first few minutes of a discussion about
 some disputed issue. </P1>

These first paragraphs tend to show a pattern in the number of sentences they consist of. A high percentage of them (76.19 % in our corpus) consist of a single sentence, an example of which is shown below:

(4) <P1> <S1> The overall success rate of fertility treatment
 has risen by 0.9%, according to the latest figures from the
 government's fertility regulator. </P1> </S1>
 (07 December 2007, *The Guardian*, p.22 *Science: Success rate of
 fertility treatment rising/Home News Randerson, James*)

This tendency to find one-sentence paragraphs is consistent with the relative shortness expected for this kind of genre. Table 2 below shows a summary of the composition of first paragraphs in our corpus, in terms of the number of sentences found.

One sentence	Two sentences	TOTAL
(76,19 %)	(23,81 %)	(100%)

Table 1. *Number of sentences in the first paragraph.*

Switching our focus of attention towards the presence or absence of attribution in the first sections of science popularization articles, we have to take into account that if these sections have been considered highly informative and thoroughly studied across different registers (Mahlberg and O'Donnell 2008), they can also be supposed to have relevant information in reference to how explicit attribution is constructed in science popularization articles. For this purpose, our approach to attribution follows not only Hunston's, when she asserts that explicit attribution can be said to be the presentation of discourse as deriving from someone different from the journalist (Hunston 1999: 178) and Conrad and Biber's, who define attribution as an explicit mark that identifies the source of information (Conrad and Biber 1999: 67), but also Bell's, when he posits that "[n]ews is what an authoritative source tells a journalist" (Bell 1991: 191), thus giving attribution sources a higher value than a mere referential function.

From this point of departure, we have analyzed the first introduction of explicit attribution sources in the first elements of science popularization articles that constitute the head of the text, in opposition to the body of it. In this respect, Ho-Dac (2008) studied the role of first elements (Theme) in the construction of discourse and stated that

the beginning of the message causes specific discourse processes to happen in the text, being thus initial sections considered key for the unfolding of texts. Findings of this analysis complete Ho-Dac's perspective and show that there is a strong tendency of first paragraphs to contain the first explicit reference to an attributee in this type of texts. Eleven of the texts which have explicit attribution show it for the first time in the first paragraph, which represents a 57.89 % of the total of texts in the corpus. It is also remarkable the fact that explicit attribution is found in only two leads (9.52 %) and six headlines (28.57 %). It is also remarkable that there are no occurrences where attribution can be found both in the headline and the lead simultaneously.

Regarding the explicit attribution sources found in initial sections of science popularization texts, it is also necessary to provide data about the entities or people to which the reference is made. A different pattern that has also been analyzed in the corpus is how attribution is explicitly expressed through processes in first paragraphs. According to Halliday and Matthiessen (2004: 175), there are three different components, organized in configurations that provide the models for building experience of what goes on.

1. A process unfolding through time
2. The participants involved in the process
3. Circumstances associated with the process

The concepts of process, participant and circumstance are semantic categories which seek to explain, in a general way, how reality surrounding us is construed as a linguistic structure.

Bearing this classification in mind, we have had a look at what the most frequent attributees in the corpus (actors) are, finding that 17 references (62.96 %) are made to *actors*, that is to say, we have found that in a 62.96 % of the occurrences, a scientific discovery is attributed to a person, usually a researcher or a scientist, or to a team of people who work together for the achievement of a same goal. Of these 17 occurrences, 4 are found in headlines, and only 3 in leads, while the vast majority is located in first paragraphs.

Ten times have we found explicit attribution sources referring to entities, material things which act as attributees of the scientific discoveries in science popularizations. Of these cases, 9 are again found in first paragraphs, with a 33.33 % of the occurrences, while there is only one case of entities as attributees in headlines and none in leads, as can be seen in Table 3 below.

Attributees	In headlines	In leads	In first paragraphs	TOTAL
Actors (People)	14.81 %	11.11 %	37.04 %	62.96 %
Entities (Material things)	3.70 %	0%	33.33 %	37.04 %
TOTAL	5	3	19	27

Table 2. *Type and distribution of attributees in initial sections of the corpus.*

Two examples are provided below of *actors* and an *entity* as attributees in rhematic position:

(6) 19 December 2003, The Guardian, p.5
<H> In a nutshell, this species is endangered / Home News
Tim Radford, Science editor </H>
<P1> The brazil nut - a traditional Christmas stocking filler, coated in chocolate - could become an endangered species if intensive harvesting continues, according to British and international researchers. </P1>

(7) 02 November 2006, The Guardian, p.11
<H> Science: Many endangered species stranded outside protected areas, study finds </H>
<L1> Challenge to 'hotspot' theory of conservation</L1>
<L2> At-risk animals more spread out than thought / Home News </L2>
Randerson, James

In the analysis of the corpus, we have also found two cases where there is no attribution in any of the first sections previously described in this study (headline, lead and first paragraph), and readers have to go as far as second and third paragraph to find an explicit source of attribution. In Example 8 below, the attribution source is found in thematic position at the beginning of the second paragraph, referring to the people who made the discovery reported, *scientists in Texas*.

(8) 29 July 2004, The Guardian, p.2
<H> Coke or Pepsi? It's all in the head / Home News </H>
Alok Jha Science correspondent
<P1> The long-standing conundrum of why Coke sells more than Pepsi despite being less popular in blind taste tests may have been solved. </P1>
<P2> Scientists in Texas used a brain scanning technique to carry out a hi-tech version of the Pepsi challenge and found that, when it comes to fizzy black drinks, brand love is just as important as taste. </P2>

In contrast to Example 8, where no explicit attribution was present, in Example 9 below we find the explicit attribution source in the third paragraph, making reference to the intellectual father of the discovery, who is not mentioned before in any of the preceding parts of this article.

(9) 21 June 2004, The Guardian, p.7
<H> High flyer may open the final frontier / Home News
Tim Radford, Science editor </H>
<P1> A new kind of spacecraft slung from an improbable high flyer could make history today. SpaceShipOne, carried into the stratosphere by a turbojet-powered sailplane called White Knight, will drop from its mothership, fire its rocket engines and soar to

an altitude of more than 60 miles on a mixture of rubber and laughing gas. </P1>

<P2> If all goes well, SS-1 will become the first privately funded passenger-carrying spaceship to break out of the atmosphere and touch the edge of space. Its pilot will experience weightlessness for up to three minutes before changing the shape of the plane's wings and gliding back to the runway, just one hour and 30 minutes after takeoff. </P2>

<P3> The project is the brainchild of Burt Rutan, who entered aviation history when he built Voyager, the first jet to make a non-stop voyage around the world. Rutan, backed by the Microsoft co-founder Paul Allen, wants to open the final frontier for a new generation of entrepreneurs and launch a new business: space tourism. </P3>

6.2. *Patterns typically found in first paragraphs of science popularizations in The Guardian newspaper*

One of the outcomes obtained concerning the patterns found in the articles is that where we find a *bulleted lead*, explicit attribution is present in both the headline and the first paragraph (cf. Table 3). In Example 2 above, it can be observed how the attribution source is present in the headline ('says Best surgeon') and also in the first paragraph in thematic position (*The surgeon who performed George Best's liver transplant says ...*).

Regarding the occurrences of explicit attribution found in the first paragraph of science popularization articles in our corpus, it is relevant to highlight that in 89.47 % of the cases, attribution is found in the first sentence of the first paragraphs, while only in 10.53 % of the texts explicit attribution sources appear in the second sentence (and last) of the first paragraph. The 100 % of the texts which have a first paragraph consisting of a single sentence show attribution in it, while among the 5 cases in which the first paragraph of the text consists of two sentences, we can find three occurrences of attribution in the first sentence, one of attribution in the second sentence, and one case of first paragraph where there is no attribution source. It is noteworthy to mention that first paragraphs in the corpus only consist of either one or two sentences, so there are no paragraphs of more than two sentences.

	ATTRIBUTION IN P1
First sentence	17
Second sentence	– Text 5 (sentence 2 out of 2) – Text 23 (sentence 2 out of 2)
TOTAL	19

Table 3. *Distribution of explicit attribution sources according to sentence in first paragraphs.*

Another pattern we have paid attention to in this research is the relative positions of explicit attribution sources within the sentence structure. We have found 10 occurrences (47.62 %) of sentences in which the explicit attribution source is found in rhematic position, occurring in the final position of the sequence (as in Example 10 below). These data mean that explicit attribution studied in our corpus shows a tendency to be placed at the end of the sentence, which is consistent with Bell's finding that "[i]n the press, attribution is often postponed" (Bell 1991: 193), although he highlights that this is especially the case after direct quotation and in our corpus, this is not necessarily so.

(10) 23 November 2005, The Guardian, p.8
 <H> Science: North-west and south-east get floods and droughts
 warning / Home News </H>
 Sample, Ian
 <P1> Towns and villages in the north-west and south-east of
 England can expect more severe winter flooding and lengthy droughts
 during the summer in coming decades, according to a report
published by climate scientists today.</P1>

The second most frequent position for explicit attribution sources to be found in the corpus is as theme (as in Example 11 below). Seven of the examples of attribution found in the corpus (33.33 %) occupy the initial position in the text, while only four times (19.05 %) do we find attribution in medial position of the sentence, not occupying thematic or rhematic positions in the sentences of the first paragraphs. A summary of distributions according to thematic structure can be seen in Table 5 below.

(11) 13 February 2004, The Guardian, p.6
 Psychologist says maths can predict chances of divorce / Home News
 Tim Radford Science editor
A psychologist claims that a newly devised mathematical model
 can predict with 94% accuracy which couples will divorce - entirely
 on the basis of the first few minutes of a discussion about some
 disputed issue.

In addition to the distribution of explicit attribution sources within the sentence structure of the sentences found in the first paragraphs, distribution has also been analyzed in relation to the relative position of attribution within the paragraph as a unit. Therefore, an analysis has been carried out to see whether there is a tendency of explicit attribution sources to occupy a certain position within the paragraph or, in other words, to see if attribution is primed (Hoey 2005) for a particular position in first paragraphs of popularizations.

In contrast to the majority of occurrences in rhematic position found in the observation of sentences, when focusing our analysis on paragraphs we find that there are 5 occurrences in thematic position (23.80 %), while there have been observed 8 occurrences of attribution in medial and rhematic positions (38.1 %).

In relation to Text 23, where we find the explicit attribution source in the second sentence of the first paragraph, instead than in the first, we find a case of medial

attribution, understanding by medial the one that is not explicitly marked as thematic or rhematic, and that, therefore, occurs in the middle of the sentence or paragraph as in Example 12 below.

(12) 21 November 2007, The Guardian, p.11
Science: Relax, it died 390m years ago - huge scorpions find
points to existence of giant spiders/ Home News
Claw find in quarry makes scientists think again
Arthropods may have been bigger than thought
Morris, Steven
It is enough to give people with arachnophobia a large dose of
the heebie-jeebies. Scientists have discovered the fossilised claw
of a sea scorpion that suggests the giant scorpions, spiders and
crabs that once crawled around the world were even bigger than
previously thought.

Referring again to the typologies put forward by Halliday and Matthiessen (2004: 175), on the components that provide the models for building experience of what goes on (A process unfolding through time; The participants involved in the process; Circumstances associated with the process), we have now focused our attention on the processes appearing in the first paragraphs of science popularization articles, specially on those having the aforementioned attributees as subjects or objects in the clause.

Following the Systemic Functional Grammar approach to language description, verbal processes can be classified in different categories, according to the meaning of the verbal group. Taking into account this classification (Halliday and Matthiessen 2004: 168-259) and analyzing what is present in the corpus, we have found several verbal groups, or processes. The first is a group of reporting verbs. This verbal group is used in the corpus to allow the writer to attribute information to different sources, which is done in two different ways:

1. By quoting, i.e. direct speech including the exact words uttered by the person to whom the attribution is being done, between inverted commas.
2. By reporting, i.e. indirect speech including a summary of the words uttered by the person to whom the attribution is being done.

There is an imbalance in the frequency of occurrences of the two types of 'saying' process mentioned, being more frequent to encounter direct speech quoting between commas (61.25 %), although a single case lacking the inverted commas has also been spotted. In contrast, there is a 38.75 % of cases when the journalist has chosen to report instead of quoting in the science popularization articles used to carry out this study, where a total number of 80 verbal processes (100 %) associated to explicit attribution sources have been identified. In the initial sections of the corpus (headline, lead, first paragraph), things are quite different and we only find nine occurrences of 'saying' verbs (11.25 %), all of them projecting reports. It is remarkable the difference between the high presence of quotes in science popularization articles and the absence of them in first

paragraphs, a pattern that leads us to conclude that direct speech quoting is not a very common practice in the initial sections (Headline, lead or first paragraph) of *The Guardian* science popularization articles studied.

	Sci_TG Pilot Corpus	Initial sections of Sci_TG Pilot Corpus
Verbal clauses projecting quotes	61.25 % (49) (one case without commas)	0
Verbal clauses projecting reports	38.75 % (31)	11.25 % (9)
TOTAL	100% (80)	11.25 % (9)

Table 4. *Frequency of quoting and reporting clauses in the first paragraphs.*

Among the verbs used to convey the ‘saying’ meaning in the corpus, we find, as expected, a high presence of *say*. This verb is the most frequent quoting and reporting verb in *Sci_TG Pilot Corpus*, and reporting one in the first sections of *Sci_TG Pilot Corpus*, followed by a group of verbs that also appear in the corpus, not as frequently as *say* though: *tell, express, assert, talk, speak, ask, question, command, order* or *add*. An example of it can be seen below:

(13) 13 February 2004, *The Guardian*, p.6
 Psychologist says maths can predict chances of divorce / Home News
 Tim Radford Science editor
 A psychologist claims that a newly devised mathematical model can predict with 94% accuracy which couples will divorce - entirely on the basis of the first few minutes of a discussion about some disputed issue.

The second group of verbs identified in the corpus is the one related to ‘mental processes’, including verbs that express opinion, beliefs and thoughts (typically *believe* and *think*) that the attributees manifest. There have been identified only three occurrences of this group of verbs in the whole *Pilot Corpus*, but none of them has been located in the initial sections of the articles.

The third group of verbs analyzed in the corpus can be classified according to the taxonomy put forward by Hyland (2000) and also followed by Calsamiglia and López Ferrero (2003), who divide reporting verbs in two different groups: content-oriented processes and evaluation verbs. Belonging to the first group, we have identified *findings* (*show, discover*) and *popularizing acts* (*present, publish, explain*). Belonging to the group of evaluative verbs, we have identified positive reporting verbs (*demonstrate, find, identify, confirm, conclude*), neutral reporting verbs (*study, call, raise questions*) and also tentative reporting verbs, such as *suggest*.

A total of 29 occurrences have been identified in the whole corpus, 6 of which are located in the initial sections of the articles (either in the headline, the lead or the first paragraph).

Finally, in reference to the structural patterns through which explicit attribution is constructed in the corpus, we have found that when attribution sources are located in thematic position, they typically appear conflated with the subject of the sentence and that, when occupying a rhematic position, they are typically introduced by means of ‘*according to*’. A summary of this can be found in Table 9 below, where it can be observed that there is also a marginal presence of other constructions, which includes cases of explicit attribution through passive constructions, by means of disjuncts such as ‘*researchers reported yesterday*’ or constructions with the function of object complement (OC).

	Thematic	Medial	Rhematic	Others	TOTAL
Conflated theme	31.58 % (6)				31.58 % (6)
‘<i>according to</i>’		5.26 % (1)	36.84 % (7)		42.10 % (8)
Passive structure			5.26 % (1)		5.26 % (1)
Disjunct			5.26 % (1)		5.26 % (1)
Other structures				5.26 % (1): OC 10.53 % (2): Not conflated subject	15.79 % (3)
TOTAL	31.58 % (6)	5.26 % (1)	47.37 % (9)	15.79 % (3)	(100 %) (19)

Table 5. *Relative position and type of structural patterns of explicit attribution in the first paragraph.*

In first sections of science popularization articles, the majority of explicit attribution sources are expressed by means of ‘*according to*’ and located in rhematic position, followed in frequency by those cases where the attribution is thematised and conflated with the subject of the sentence. In this respect, the data obtained here are consistent with the previous data found by García Rianza (2010a, 2010b) on the particle ‘*according to*’ in first paragraphs of science popularizations and its role as an introductory element of explicit attribution. Single instances of a passive structure (‘*...were being made by NASA officials yesterday*’) and a disjunct (‘*...researchers reported yesterday*’), (the only ones located in the whole corpus) have also been observed in the initial sections of these texts. Finally, some texts show different ways of expressing attribution, such as Texts 13 and 15, where the explicit attribution source occurs in the middle of a sentence with the function of subject, or also Text 9, where the attribution source has the function of object complement.

7. CONCLUSIONS

Through the development of this study we have been able to gain deeper insight in the anatomy of science popularization articles, especially in how first paragraphs behave in this genre as realized in *The Guardian* newspaper. If popularizations as the ones analyzed

can be said to be a relatively short type of text (both when compared to their corresponding source articles and to other genres within newspaper discourse), first paragraphs can be said to reflect this by presenting a pattern as 'single-sentence' paragraphs.

In relation to the patterns which are most typically found in science popularizations as represented in our corpus, we have observed that the initial sections vary in distribution across texts. In the corpus studied, 24 % of the texts have a lead after the headline and preceding the first paragraph; a section that acts as an introduction that summarizes the content and point of view from which the article will develop. A 60 % of the texts studied have no lead. This tendency should be further explored in order to see whether this is only typical in science popularization articles in English or rather a feature of this genre as realized in *The Guardian* newspaper.

When the articles have a *bulleted lead*, explicit attribution is to be found both in the headline and in the first paragraph, but never in the lead. The results of our analysis show that there are no cases in our corpus where explicit attribution in the headline and the lead occur simultaneously.

Concerning the average length of first paragraphs, in terms of number of sentences, it is also relevant to comment that there is a strong tendency of first paragraphs to consist of a single sentence (76.19 %) in opposition to two-sentence paragraphs that constitute 23.81 % of the total. No instances of paragraphs longer than two sentences have been found in the corpus.

In the light of the data gathered, we can state that explicit attribution sources tend to be found in first paragraphs for the first time (57.89 %). In our view, the fact that the authoritative source tends to appear in a textual prominent position such as the first paragraph comes to reinforce Bell's conception that news is indeed what an authoritative source transmits a journalist (Bell 1991: 191), even if it tends to appear in rhematic position (47.37 % of the cases found).

As expected, most of the attribution sources found in texts are actors (62.96 % of them refer to *actors*), while in 37.04 % of texts the information is explicitly attributed to *entities*. We can conclude that, although the presence of people is more frequent than the presence of entities as authoritative sources for scientific information in popularizations in English, as represented by *The Guardian* newspaper, the tendency of some entities in particular (studies, research processes or publications) to fulfil the role of attribution sources is not marginal at all.

As concerns the position where these explicit attribution sources can be found in text, we have also observed a relevant tendency for them to be located in the first sentence of the paragraph (89.47 %) against the few times in which explicit attribution is found in the second sentence of the first paragraph of the articles (10.53 %).

In respect of the position that explicit attribution sources typically occupy in sentence and paragraph structure, we have to highlight 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.

Attributional meanings are predominantly conveyed in the corpus studied through verbal processes of 'saying', which are expressed by means of both quotes and reports. A stronger tendency has been identified towards quoting rather than reporting in the corpus but, contrary to expectation, we do not find that same pattern in first paragraphs, where the total of the verbal processes of 'saying' encountered are reports.

In respect to the structures that most frequently convey attributional meaning, the most relevant pattern encountered in thematic position is that of the explicit source being conflated with the theme of the sentence. In rhematic position, the most repeated structure to convey attribution is '*according to*', data which are consistent with previous findings (García Riaza 2010a, 2010b).

In a second stage of the research, an extended corpus will be used in order to test these preliminary findings. We expect to carry out a deeper analysis both on the nature of attribution and, more specifically, into the different strategies used in English (and made use of in this type of text) to include other voices in the text; what Thompson (1996) calls 'language events' reports', as well as on the distribution of the presence of attribution sources in the text. In this way, we will be able to characterize in further detail the relationship held between first paragraphs and the rest of the text in science popularizations in *The Guardian* newspaper, in order to gain a deeper insight of this genre as realized in British newspaper discourse from a dynamic perspective on how the narration is constructed and justified by means of external sources of attribution.

NOTES

* Corresponding autor: Blanca García Riaza. Dpto de Filología Inglesa. C/Placentinos nº 18. Salamanca, 37008. Teléfono: 923 294400 (Ext.1752), bgr@usal.es

REFERENCES

- 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.
- Bell, A. 1991. *The Language of News Media*. Oxford: Blackwell.
- Calsamiglia, H. and C. López Ferrero. 2003. "Role and position in scientific voices: Reported speech in the media". *Discourse Studies* 5 (2): 147-173.
- Conrad, S. and D. Biber. 1999. "Adverbial marking of stance in speech and writing". *Evaluation in Text*. Eds. S. Hunston and G. Thompson. Oxford: Oxford University Press. 56-73.
- Dor, D. 2003. "On newspaper headlines as relevance optimizers". *Journal of Pragmatics* 35: 695-721.
- Elorza, I. 2010. "A cross-cultural analysis of popularizations in English and Spanish newspaper discourse". *English Studies* 31: 85-109.

- Fagan, A. and P. Martín Martín. 2004. "The use of critical speech acts in Psychology and Chemistry research papers". *Ibérica* 8: 125-137.
- García Rianza, B. 2010a. "Explicit attribution sources expressed by 'according to' in a corpus of science popularization articles from *The Guardian*". *Interlingüística* 21: 157-168.
- García Rianza, 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.
- 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.
- Halliday, M. A. K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday M. A. K. and C. M. I. M. Matthiessen. 2004. *An Introduction to Functional Grammar*. 3rd revised edition. London: Edward Arnold.
- Ho-Dac, L.-M. 2008. "Discourse organisation though Theme position". *Proceedings of the Nineteenth European Systemic Functional Linguistics Conference and Workshop*. Saarbrücken, Germany. Eds. E. Steiner and S. Neumann. [Available at <http://scidok.sulb.uni-saarland.de/volltexte/2008/1691/>].
- Hoey, M. 2005. *Lexical Priming*. London: Routledge.
- Hunston, S. 1999. "Evaluation and the planes of discourse: Status and value in persuasive texts". *Evaluation in Text*. Eds. S. Hunston and G. Thompson. Oxford: Oxford University Press. 176-205.
- Hyde, J.R. 1999. "Lo demás es literatura': Predicción y adversatividad en textos argumentativos en inglés y español". *Lingüística para el siglo XXI. Vol. II*. Eds. J. Fernández González, C. Fernández Juncal, M. Marcos Sánchez, E. Prieto de los Mozos and L. Santos del Río. Salamanca: Ediciones Universidad de Salamanca. 935-941.
- Hyland, K. 2000. "Academic attribution: Interaction through citation". *Disciplinary Discourses: Social Interaction in Academic Writing*. London: Longman. 20-40.
- Hyland K. 2010. "Constructing proximity: Relating to readers in popular and professional science". *Journal of English for Academic Purposes* 9 (2): 116-127.
- 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.
- Mahlberg, M. 2007. "Lexical items in discourse: identifying local textual functions of *sustainable development*". *Text, Discourse and Corpora: Theory and Analysis*. Eds. M. Hoey, M. Mahlberg, M. Stubbs and W. Teubert. London: Continuum.
- Mahlberg, M and M. O'Donnell. 2008. "A fresh view of the structure of hard news stories". *Proceedings of the Nineteenth European Systemic Functional Linguistics*

- Conference and Workshop*. Saarbrücken, Germany. Eds. E. Steiner and S. Neumann. [Available at: <http://scidok.sulb.uni-saarland.de/volltexte/2008/1700/>].
- Marks, L.A., N. Kalaitzandonakes, L. Wilkins and L. Zakharova. 2007. "Mass media framing of biotechnology news". *Public Understanding of Science* 16 (2): 183-203.
- Myers G. 1990. "Making a discovery: narratives of split genes". *Narrative in Culture: The Uses of Storytelling in the Sciences, Philosophy, and Literature*. Ed. C. Nash. London: Routledge. 102-126.
- Myers, G. 1994. "Narratives of science and nature in popularizing molecular genetics". *Advances in Written Text Analysis*. Ed. M. Coulthard. London: Routledge. 179-190.
- O’Keeffe K. J. 2002. "The quality press in Ireland: A linguistic analysis of two text-types". *Windows on the World: Media Discourse in English*. Ed. A. Sánchez Macarro Valencia: Universitat de València. 239-263.
- 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.
- Reah, D. 2002. *The Language of Newspapers*. Abingdon: Routledge.
- Scott M. 1997. "PC Analysis of key words – and key key words". *System* 25 (2): 233-245.
- Scott, M. 2004-2005. *WordSmith Tools, Version 4.0*. Oxford: Oxford University Press [Software].
- Skorczynska, H. 2001. "Metaphor in scientific business journals and business periodicals: an example of the scientific discourse popularisation". *Ibérica* 3: 43-60.
- Thompson, G. 1996. "Voices in the text: Discourse perspectives on language reports". *Applied Linguistics* 17 (4): 501-530.
- Thompson, G. 2004. *Introducing Functional Grammar*. 2nd edition. London: Edward Arnold.
- Williams Camus, J. T. 2009. "Metaphors of cancer in scientific popularization articles in the British press". *Discourse Studies* 11 (4): 465-495.
- Zapaterra, Y. 2008. *Diseño Editorial: Periódicos y revistas*. Barcelona: Gustavo Gili. [Translation from Zapaterra, Y. 2007. *Editorial Design: For Print and Electronic Media*. London: Lawrence King].

