

**LINGUISTIC CHOICES IN TWO RESEARCH ARTICLES IN
PHYSICS: STUDY OF AN AUTHOR'S DEVELOPMENT**
**Escolhas Lingüísticas em Dois Artigos de Pesquisa na Área de
Física: um estudo sobre o desenvolvimento de um autor**

Ann MONTEMAYOR-BORSINGER (University of Glasgow, U.K. and
Instituto Balseiro – Universidad Nacional de Cuyo / Comisión
Nacional de Energía Atómica, S.C. de Bariloche, Argentina)

Abstract

This study seeks to identify linguistic elements that could help understand the process by which scientists, who are already familiar with research article conventions, develop as authors. It focuses on the choices made in Theme, identified as elements which come in first position in the clause, and compares two articles published in Physical Review B by the American Physical Society. The articles are the first and the last of a series of five articles written by the same researcher on the same problem in physics. The study also presents interviews conducted with the author and other physicists that indicate ways in which the articles differ. The method of analysis of the texts uses a formulation of Theme that includes Subject as an obligatory element. The analysis, using taxonomies proposed by Davies (1988 & 1997) and Gosden (1993 & 1996), suggests that elements in Theme can be manipulated by writers not only to organise their message, but also, and importantly, to evaluate their research and negotiate with the reader. The results of the analysis and the interviews show a tendency for increased authorial presence as experience in communicating science increases.

Key-words: *subject; theme; authorial development; research articles.*

Resumo

Este estudo busca identificar elementos lingüísticos que poderiam auxiliar a compreensão do processo pelo qual cientistas, que já estão familiarizados com as convenções do artigo de pesquisa, desenvolvem-se como autores. O artigo focaliza as escolhas realizadas no Tema,

identificado como elementos que ocorrem em posição inicial na oração, e compara dois artigos publicados na Physical Review B pela Sociedade Americana de Física. Os artigos são o primeiro e o último de uma série de cinco artigos escritos pelo mesmo pesquisador a respeito do mesmo problema na área de física. O estudo também apresenta entrevistas realizadas com o autor e com outros físicos que indicam aspectos nos quais os artigos diferenciam-se. O método de análise dos textos utiliza uma formulação de Tema que inclui o Sujeito como um elemento obrigatório. A análise, utilizando taxonomias propostas por Davies (1988 & 1997) e Gosden (1993 & 1996), sugere que os elementos presentes no Tema podem ser manipulados pelos escritores não apenas para organizar sua mensagem, mas também, e o que é mais importante, para avaliar sua pesquisa e negociar com o leitor. Os resultados da análise e as entrevistas mostram uma tendência para uma maior presença do autor à medida que sua experiência em comunicação científica aumenta.

Palavras-chave: *sujeito; tema; desenvolvimento da autoria; artigos de pesquisa.*

1. Introduction

Growing interest in academic writing has given rise to a number of linguistic analyses of such texts. With respect to the resea

Another approach to the comparative study of academic texts – this time of non published texts rather than of research articles – is that proposed by Berkenkotter, Huckin and Ackerman (1991) who analyse three different introductions of assignments written by the same PhD student. By analysing these texts, Berkenkotter *et al* focus on how this student, who was not familiar with the conventional structure of articles, started acquiring the genre knowledge characteristic of a research community. This last study is of particular relevance for understanding how a student develops into being a researcher.

The present study, in a similar way to Berkenkotter et al, is also concerned with authorial development. However, instead of focussing

on how a student becomes a researcher, as in the Berkenkotter study, it takes as its point of departure a young physicist¹ who has already been apprenticed into his discourse community by working in research settings. Specifically it seeks to identify the different linguistic choices related to Theme the author makes as he gains experience in publishing his work. When teaching ESP, awareness of such choices can help researchers consider different ways of presenting their results. In particular, novice researchers could improve their writing and their chances of having papers accepted in international journals by looking at the choices made by more experienced researchers.

The study analyses the first and the fifth article the researcher wrote on his own, published in 1995 and 1997 respectively, in the same international journal, *Physical Review B*, of the American Physical Society. These articles, which present the results of his Ph.D. research, were written during a crucial period of development for the young scientist.

The “novice” researcher in question – “novice” in the sense that he is writing his first research papers for international journals – works in solid state physics in Argentina. He is a non-native speaker of English, but his mastery of the language is native-like especially regarding academic genres. He followed intensive courses of English for several years and used English daily during the course of his studies in physics, both as an undergraduate for four years and as a graduate student for another four years. In physics the great majority of research papers are published in English. English is also used as a lingua franca by visiting lecturers and researchers.

The study compares the different choices this novice makes as he strives to become an “expert”. “Expert” is defined here as an experienced researcher who regularly publishes in international journals, who is a referee and sometimes an editor of these same journals, and who supervises “novice” research work.

¹ The researcher is a male, and thus is referred to as “he”. The same occurs later on in the text with an “expert” researcher, also male. For the group of “novices”, where there were eight men and one woman, either the plural or “s/he” “her/his” is used. For all the other cases when I speak generically of researchers I also use either the plural or the “s/he” “her/his” forms.

In order to locate the study within the broader context in which a text is composed, several interviews were conducted in English with the author of the articles and other physicists. The author was interviewed on his writing process and on the differences he saw between the two articles. To gain additional insights into the writer's development, specialist readers were also asked for their opinions on the differences between the abstracts of the articles. Only the abstracts were discussed in detail with specialist readers because the whole articles proved to be far too long to use as a basis for interviews of about one hour each. These interviews, presented below in Section Two, give a qualitative framework indicating ways in which the articles differ and why.

Section Three discusses aspects of the Systemic Functional approach adopted for the analysis of the texts. It focuses on the first part of the sentence known as Theme, where more given and interactive meanings tend to cluster (Halliday 1994: 36-37; Berry 1995:58; Ravelli 1995:227). These meanings, in the case of highly specialized texts, are suggested as being the more discipline-independent and manageable meanings a linguist can deal with. They are classified using taxonomies of Thematic elements proposed by Davies (1988 & 1997) and Gosden (1993 & 1996).

In Section Four systems of choice within Theme for the two texts are identified. Section Five discusses the findings of the analysis within the broader context given by the interviews.

2. The Interviews

This section presents the findings of the interviews with the writer and with other physicists working in areas related to the publications. The writer was asked to compare both articles and his experience in writing them. The other physicists – one expert and nine novice researchers – were asked to compare the abstracts and talk about the differences they perceived in them. These interviews were aimed at eliciting opinions on the texts by insiders.

The interviews with the author and the expert researcher were recorded and transcribed. Pauses appear in the present text as suspension

marks. Written notes were taken of the interviews with the novice researchers. The interviews with the author are discussed below. The ones with the expert and the novices are presented in Subsections 2.2 and 2.3 respectively.

2.1. Interviews with the Author of the two Papers

Three interviews of about an hour each were conducted with the author on his writing process and on the differences between his first article, hereafter TEXT 1, and his fifth, hereafter TEXT 2. During the first interview (Interview 1) the author gave general comments on the papers, both of which present research in solid state physics, and in particular results of the same type of numerical simulations applied to superconductivity. In Interviews 2 and 3 the author gave more specific comments on the differences between the papers.

In particular the author said that although both papers presented results from the same superconductivity model using the same kind of numerical simulations, there was a qualitative difference in the results and thus a difference in their organisation. In TEXT 2 there is one central result that is presented in the most important figure of the paper. In contrast, there is no central result in TEXT 1, but several minor ones. Hence the author felt that when writing TEXT 1, the first paper, he had had to “jump around in the text from one result to the next” whereas it had been much easier to organise TEXT 2, the last paper, around the central result.

When asked to be more precise about these differences in writing up the two papers, in the author’s words,

“it’s difficult to... to distinguish between what is... what is due to the... to our understanding of the physics and my capacity to write at this or that moment ...” (Interview 2)

However, he did point out that

“at the level of sentences it was more fluid for me to write this... the last paper than the first one ...” (Interview 3).

By the time the author was writing his fifth paper, the previous four had already been successfully published and were used as citations to back up his latest work. This would explain why the author said the results in the fifth paper had a stronger scientific basis, and why he had felt more relaxed about writing up these last results.

2.2. Interview with an Expert Scientist

As was mentioned previously, because the complete texts were too long to be discussed in detail with different physicists (both texts have approximately 5,700 words each, with many equations and complex figures) the abstracts of each paper were used as a basis for the other interviews. In order to get as wide a range as possible of opinions, two different types of specialist readers were interviewed. First impressions were asked on the one hand from an expert physicist, and, on the other, from novice physicists who were themselves in the process of publishing their first articles.

The abstracts of the papers were presented to Professor A, an expert informant working in theoretical physics at Oxford University. He is the author of many research articles and textbooks within his field of research, as well as an editor and referee for publications in theoretical physics. He has supervised numerous PhD students. At the time of the interview he was a visiting professor at the author's workplace in Argentina. He only saw the texts as ASCII files, as they are presented here, without knowing if and when they had been published.

FIRST PAPER: Abstract TEXT 1

We calculate numerically the behavior of a model high-temperature superconductor described by a three-dimensional array of Josephson junctions in the presence of an external magnetic field using dynamical Langevin simulations in the extreme type-II case. In particular, the voltage generated when an external current is applied, and its dependence on the external field and thickness of the sample are discussed. We find that the $\{it\ ab\}$ -plane resistivity is well described by a thermal activation model, whereas the $\{it\ c\}$ -axis resistivity appears for higher temperatures. To make connection with recent experiments, the response to non-homogeneous applied currents is also discussed.

**LAST PAPER: Abstract TEXT 2**

We propose a phase diagram for the vortex structure of high temperature superconductors which incorporates the effects of anisotropy and disorder. It is based on numerical simulations using the three-dimensional Josephson junction array model. We support the results with an estimation of the internal energy and configurational entropy of the system. Our results give a unified picture of the behavior of the vortex lattice, covering from the very anisotropic $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8-x}$ to the less anisotropic $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, and from the first order melting occurring in clean samples to the continuous transitions observed in samples with defects.

Professor A was told they had been written at different times by a young researcher from the institution he was visiting. On the basis of what the author of the papers had said in the interviews commented above, Professor A was asked whether he felt the abstracts had been written differently, whether he thought the author was more mature and more at ease in one of the abstracts, and whether one of them “read” better than the other.

Interestingly, Professor A started his comments by focussing on “lower level issues” (Gibson, 1993), i.e. spelling, details of grammar, etc., what Professor A himself termed during the interview as “hiccoughs”:

“(…) well certainly Abstract 1 I don’t think that I would be able to guess that it wasn’t written by an English speaker... I don’t see anything wrong with it... ... whereas Abstract 2 is... you see for instance there is a word anisotropic... that should be anisotropic... you see it should be an i instead of a y... he got it correct there... maybe it’s just a slip... an understandable hiccough... but I wouldn’t... also I think... when you say... *‘Our results give a unified picture of the behaviour of the vortex lattice covering from the very anisotropic’*... that doesn’t read quite right either... you don’t say *‘covering’* you could say *‘covering examples ranging from’*... you wouldn’t just say *‘covering from’*... that’s just a minor hiccough... it could be the other way round... whereas here [pointing to Abstract 1] I don’t detect any... any hiccough at all...”

However, without my interrupting him, he then continued by discussing “higher level issues” (Gibson, 1993) of discourse, having more to do with level of formality reflected in the use of passive versus active voice:

I mean Abstract 1 is written in the... sort of professional passive sense... *‘are discussed’* ... *‘is discussed’* ... like that... this is the sort of jargon style as it were... this is more [pointing to Abstract 2] *‘this-is-what-I-did’* style... which is quite nice actually... I quite like that too... *‘we propose something’* fine good for you... *‘and this is what it is based on’* ... *‘we support results’* ... *‘we don’t say ‘a phase diagram is proposed’ ‘the results are supported’*... and so on and so on ... (laugh)... that’s what strikes me about this... there’re in different modes as it were... well who’s to say which is a better mode... I mean Abstract 1 is clearly in a more conventional impersonal mode... there is no doubt about that... but Abstract 2 is perfectly O.K.... ... and I would say that Abstract 2 reads in a very very nice friendly way... in a more chatty kind of informal way...”

Professor A was then told that Abstract TEXT 1 had been written first. He said that the scientist had done a perfect job with Abstract TEXT 1, but that although Abstract TEXT 2 had some minor flaws it was in fact more “fluid”, and that the young scientist was speaking with his own voice. He finally commented that in the case of Ph.D. students writing up their theses, they certainly knew what the usual conventions were and when starting to publish

“they might well want to be so strictly correct... and might not have the confidence... the self confidence... to write in a more personal voice”

which would explain the highly conventional and impersonal tone of Abstract TEXT 1. He concluded that Abstract TEXT 2 flowed better and was in fact more fluent because, in his words,

“he [the author] is more relaxed... now you see he has already published four papers... he feels... you know... he’s... what he’s doing is O.K he’s speaking with his own voice more...”.

2.3. Interviews with Novice Scientists

The two abstracts were also presented to nine Ph.D. students in physics from the Argentinian institution, who had either published or were in the process of publishing their first papers in English-language journals. As was pointed out in the introduction, papers in physics are mostly in English. Hence, undergraduate students of physics have to be able to read English very early on in their studies, and have to learn how to write in English during the course of their Ph.D. programme. Moreover, the Ph.D. students of the institution mentioned here are expected to have published a minimum of two or three articles in international journals by the time they get their doctoral degree.

The nine Ph.D. students were participants at a workshop on academic writing. The students were divided into three groups, and were given Abstract TEXT 1 and Abstract TEXT 2 as ASCII files (see above). They were asked, as in Professor A's case, whether they perceived differences in the way the two abstracts had been written, whether they thought the author was more mature and at ease in one of the abstracts, and whether one of them "read" better. Here again the purpose was to register participants' impressions, based on their perceptions as readers and novice writers of scientific papers.

Unlike Professor A, who had extensive experience as an editor, referee and supervisor of Ph.D. students, the participants did not attempt to approach the texts as editors, but rather just as readers. Their comments were more general and focussed on what they understood from the content of the abstracts, although they did offer comments about language features as well. Students in one group stressed that Abstract TEXT 2 was more attractive, more comprehensive and more powerful, with more far-reaching conclusions than Abstract TEXT 1. They thought Abstract TEXT 1 probably dealt with a more specific and limited research topic. The second group said that Abstract TEXT 2 seemed to be more interactive and easier to read because it had no passive verbs. Students in the third group said that in Abstract TEXT 2 it was clear who proposed the model, whereas in Abstract TEXT 1 it was not. They said they preferred the "structure" of Abstract TEXT 2 because it had clear statements that were easier to read.



In sum, Professor A felt that the author was more conventional and more constrained in Abstract TEXT 1, whereas in Abstract TEXT 2 he seemed more independent and assured. In a similar way PhD students found Abstract TEXT 1 was of a more limited nature, whereas Abstract TEXT 2 was more powerful, with clear author presence. The author himself voiced the fact that he had felt more confident and his composing process had been easier when writing up TEXT 2.

3. Method of Text Analysis

The previous section has presented the author's views on the articles, and preliminary impressions of informants based on the abstracts. This enables us to approach the linguistic analysis bearing in mind the context in which these texts were written.

Concerning the analysis per se, one way of looking at the differences in the texts is by focussing on the first part of the sentence known in Systemic Functional Linguistics as Theme. Halliday associates Theme with what is given, known, and what the sentence is about (1994:37). Moreover, as Berry (1989 & 1995) and Ravelli (1995) have shown, an important amount of interactive meanings also tends to concentrate at the beginning of the sentence.

These more interactive meanings will be examined here because once physicists have obtained results that warrant publication, i.e. new information partly under the form of figures and equations, they then have to find the appropriate linguistic expressions to pass on these results to their research community. The more interactive types of meaning, which often tend to cluster in Theme, have then to be managed by scientists to convince their peers of the importance of their results.

The remainder of the sentence, which generally contains the new information, is called Rheme. It is interesting to note that in the present corpus all the equations, which contain the new information that has to be passed on to the corresponding research community, are in Rheme position. Equations, and, in general, specialised new information are highly specific to a given discipline, and are thus

extremely difficult to classify when doing linguistic analyses. The relatively more discipline-independent aspects of the research article can be studied by focussing on an analysis of Theme.

The next two subsections discuss in more detail how Theme is handled in the present work, and are followed by another two that present taxonomies of Theme elements necessary for the linguistic analysis.

3.1. Extension of Theme

As defined by Halliday, “the Theme is the starting-point for the message; it is the ground from which the clause is taking off.” (Halliday 1994:38), but within systemics there are very different positions regarding the extension of Theme. The analysis of the present study follows Enkvist’s (1973) original proposition that Theme should include Subject. This proposition has been taken up more recently by Davies (1988 & 1997) who also includes Subject as an obligatory element in Theme. In a similar way Berry (1989 & 1995) includes in Theme everything that precedes the verb of the main clause. Mauranen, in her study of academic texts in Finnish and in English, also states that “...it seems useful to take the entire preverbal part of the sentence into consideration when comparing thematic choices...”(1996:208). Because in the present corpus there are no elements between Subject and Verb, including either Subject or all preverbal elements in Theme is equivalent.

It is suggested that these extensions to Halliday’s Theme give it more pedagogic potential and make it closer to what we feel Theme should be, perhaps because “Subject is equated with the intuitive notion of “what the clause is about” (Davies, 1988:177). Davies thus postulates two potential functions for Theme. These are “identification of [obligatory] *Topic*, realised by Subject, and provision of [optional] *Contextual Frame*, realised by elements preceding Subject” (Davies, 1997:55, italics as in the original, text in brackets added).

Here the label Subject rather than the more problematic label Topic will be used from now on, mainly because there has been

considerable discussion around what is actually the topic of a sentence². In addition, Theme is only analysed in main clauses in order to give a clearer picture of essential thematic patterns without the interference of secondary organisation. If subordinate or projecting clauses are put in front of the Subject of the main clause, these clauses are considered as performing an orienting function and are classified as Contextual Frames.

3.2. Taxonomy of Theme Components Used in this Study

For Halliday, when the Subject of a sentence is conflated with Theme it is treated as unmarked (i.e. it has no preceding elements). Both Davies (1988 & 1997) and Gosden (1996) have discussed Subject functioning as unmarked Theme in academic texts. Moreover, on the basis of work initiated by Daneš (1974) and taken up again in systemics by Fries (1981), Subject is discussed as a recurrent element in discourse. This “repeated occurrence ... of the same topical element ... as Subject is seen not only to specify Topic, but also to be the primary means by which the continuity of coherent discourse is achieved.” (Davies, 1988:177).

Optionally the Subject of a sentence can be preceded by a Contextual Frame whose function is to help “the development of Topic as the discourse proceeds” (Davies, 1997:55). When this occurs, the Theme is said to be marked. An illustration of unmarked and marked Theme is shown below with examples from the present corpus³. Table 2 shows in particular how the optional element of Contextual Frame marks Theme. We saw above that in the present corpus Rhemes will not be analysed because they are much more subject-specific: it is where all the equations have clustered, and where there is most of the “new” highly specialised message scientists want to pass on to their discourse community.

² See for instance Fries, 1995:318.

³ In what follows all the examples in italics come from the present corpus.

<i>The thermodynamical free energy F</i>	<i>is obtained by minimizing with respect to ξ_c and ξ_{ab}:</i>
SUBJECT	$F(T) = \min_{0 \leq \xi_{ab} \leq 1} \min_{0 \leq \xi_c \leq 1} \Phi(\xi_{ab}, \xi_c).$
UNMARKED THEME	RHEME

Table 1: Unmarked Theme

<i>In this paper</i>	<i>we</i>	<i>propose a qualitative H-T-η-D phase diagram of high-T_c materials that reproduces most of the available experimental results.</i>
CONTEXTUAL FRAME	SUBJECT	
MARKED THEME		RHEME

Table 2: Marked Theme

3.3. Discourse Functions of Subject

Gosden (1996) has worked extensively on unmarked Theme – conflated with obligatory Subject – within the context of scientific writing, and his taxonomy based on four domains is adopted here. The ordering of the four domains with their corresponding subdomains are presented from top to bottom reflecting the continuum from “personally visible” to “invisible” initially distinguished by Davies (1988) and developed by Gosden as a continuum from

“the **Participant** to the **Real World** Domain. Towards one end, it is typified by the increasingly overt presence of the writer as a visible participant in the research reporting process; towards the other, there is a greater focus on research-based, that is real-world physical and mental entities and activities.” (1996:98)

1. The **Participant** Domain is realised by elements such as *We* and *Our approach*, where the author blatantly appears in the text.
2. The **Discourse** Domain is realised by elements such as *This point* and *Figure 4*. These elements focus on the text and its parts and on the discourse acts of reporting and discussing.
3. The **Hypothesised and Objectivized** Domain is realised by elements such as *a unified, consistent with experiments description of the problem, even at a qualitative level (is still lacking)* [sic]⁴ representing evaluative writer comment. This domain represents “a wealth of perhaps the most subtle means by which writer’s comments on hypotheses and viewpoints can be realised” and “may therefore be seen to represent the most discreetly interactional Theme” (Gosden 1996:101). Furthermore, as Davies (1988) observes, the **Hypothesised and Objectivized** Domain enables authors to *treat* theories, hypotheses, models and categories as objective entities by putting them in Subject role, although they know such entities have a hypothetical status: “the hypotheses and categories are presented, together with evaluative comment, as objects with a greater than hypothetical status” (Davies, 1988:194). An interesting example regarding this latter potential of Subjects in the **Hypothesised and Objectivized** Domain is the following from TEXT 2:

The minimizing of F with respect to ξ_c and ξ_{ab}

SUBJECT
UNMARKED THEME

allows one to obtain the $\xi_c(T)$ and $\xi_{ab}(T)$ functions, which in turn are used to detect the superconducting transitions

RHEME

Table 3: Example of a Hypothesised and Objectivized Subject

⁴ Word order in this Subject from TEXT 2 is awkward, but this is the way it appears in *Physical Review B* of the American Physical Society.

Here the author has chosen to put in Subject role a mathematical operation, in an attempt to give it greater objective status for other scientists. Instead of writing “I have minimized F with respect to ξ_c and ξ_{ab} so that I can obtain the new functions I need to detect superconducting transitions” he presents *The minimizing of F with respect to ξ_c and ξ_{ab}* as an objective entity. It looks as if it is not the researcher that minimizes F so that he can obtain $\xi_c(T)$ and $\xi_{ab}(T)$, but rather that it is the minimization of F – presented as an objective entity independent of the researcher – which “**allows**” the researcher to do other things, when it is actually the other way round.

4. The **Real World** Domain is realised by elements such as *Impurities* and *Dissipation*, which represents the researcher’s object of study.

3.4. Discourse Functions of Contextual Frame

Davies has analyzed marked Theme and Contextual Frame and has observed that “Unlike topical elements which are the recurring elements of coherent discourse, these framing elements are typically non-recurrent and as such signal changes, shifts or stages in the progression of the discourse” (1997:55). She adopts a categorisation which

“allows for the inclusion, as examples of marked, and (multiple) theme, of elements which are not identified as such by Halliday, that is, the class of “minimal” adjuncts represented by conjunctive and modal adjuncts and conjunctions and, in addition, *a small set of thematic Subjects which are seen to be marked in their semantic role* in that they do not identify participants, ..., but instead, appear to “frame” the message by specifying discourse goals or projecting evaluation.” (1997:56, italics as in the original text).

A typical example of such frames in the present corpus is *It is thus likely that (the optimum value of h for the occurrence of the first order transition decreases with sample thickness)*, where *It is thus likely that* is a Contextual Frame which projects evaluation. Taylor Torsello offers a discussion of this type of projection noting that it is highly interpersonal

because it is “a means through which speakers limit their own commitment to the message” (1996:156).

Davies defines four categories of Discourse Functions for Contextual Frame:

1. **Logical Relations/Progression** Contextual Frames (comparison, addition, contrast, reason, consequence, condition, concession, apposition/restriction). This type of Contextual Frame is subdivided for the purpose of the present study into **Minimum** and **Maximum Logical Relations/Progression**. **Minimum Logical Relations/Progression** are short conjunctive phrases such as *however*, *in addition*, *but*, etc. **Maximum Logical Relations/Progression** are whole clauses of condition, concession, etc., such as *Although this assumption cannot be fully justified a priori...*
2. **Location** Contextual Frames (e.g. *In Section III...*, *Within each layer...*)
3. **Goal and Process** Contextual Frames (e.g. *To make this estimation...*, *In order to be able to apply a current to calculate resistivities...*).
4. **Evaluation** Contextual Frames, where the author evaluates in expressions which come before the Subject of the main clause (e.g. *It is thus likely that*, *This indicates that*, *Note that in this case*, *Fortunately*).

4. Findings of the Linguistic Analysis

The present section attempts to present a linguistic explanation for the differences perceived in the two texts by analysing Theme, where more given information and more interactive meanings tend to cluster. We saw that although both papers presented results from the same superconductivity model using the same kind of numerical simulations, the author said that in contrast to the first paper, where results had been dispersed, the last paper had been organised around one main result. This was confirmed by impressions from informants who felt the author was initially more impersonal and less assured. The two texts will now

be compared by means of the identification and categorization of Theme presented in the previous section.

4.1. Comparative Analysis of Subject

Table 4 shows how Subject is distributed following the four domains distinguished by Gosden (1996). The percentages for Subject represent all instances of a particular Subject category divided by the total number of main Themes, i.e. in the present analysis the total number of sentences. The Subjects are ordered from top to bottom, from more interactional Subjects with greater writer visibility to Subjects where the writer becomes less and less visible, that is from the Domain of **Participant** to **Real World**. One main difference between the first and second paper is the variation in the distribution of Subjects in the **Hypothesised & Objectivized** and **Real World** Domains. In particular, the frequency of Subjects in the **Hypothesised & Objectivized** Domain is much greater in the second paper than in the first. We have seen above that the **Hypothesised & Objectivized** Domain allows authors to *treat* theories, hypotheses, models and categories as objective entities by putting them in Subject role, although they know such entities have a hypothetical status. By presenting elements of their work as subjects in this domain, authors give them enhanced status within the scientific arena. This capacity of anchoring their work within an abstract world shared by the profession may help give authors, like the present physicist, a more expert tone in their writing.

Another difference between the two texts is the degree of author participation reflected in the Subjects. If we add the percentages corresponding to the **Participant** and **Hypothesised & Objectivized** Domains where the author is more visible we can see that in the first paper, TEXT 1, the writer chooses to have some degree of presence in less than half of the Subjects, whereas in TEXT 2, he is present in 77% of the Subjects. We have seen above that authorial presence can be either overt, as in the **Participant** Domain, or more covert, such as in the **Hypothesised & Objectivized** Domain, when authors are able to fashion certain types of persuasive Subjects that help convince readers of the importance and validity of their results.

SUBJECT	TEXT 1	TEXT 2
Participant	20%	24%
Discourse	5%	8%
Hypothesised & Objectivized	22%	53%
Real World	53%	15%
<i>Subtotal Participant and Hypothesised & Objectivized</i>	42%	77%
<i>Subtotal without Hypothesised & Objectivized</i>	78%	47%
TOTAL	100%	100%

Table 4: Distribution of Subject

We can also opt for a different cline, which I shall call the cline of abstractness, where **Participant** and **Real World** would be the less abstract choices, closely followed by **Discourse** and culminating with **Hypothesised & Objectivized**, the most abstract choice of all. Here, if we sum up the percentages of the three less abstract domains in the *Subtotal without Hypothesised & Objectivized*, we find that in TEXT 1 nearly 80% of the Subjects involve choices anchored in the **Participant**, **Real World** and **Discourse** Domains. Those choices might be easier to make, in the sense that they involve putting into the Subject slot pre-established linguistic elements, such as the authorial *I – we* for **Participant**, *vortex, impurities, current, superconductor* for **Real World**, all of which are commonly used elements within the discourse community concerned, and *Section 1* and *Figure 2* for **Discourse**. In contrast, **Hypothesised & Objectivized** Subjects are not commonly used elements, but rather have to be especially crafted by the author: compare for instance the difference between putting *superconductor* (**Real World** Domain) in Subject position, with putting *the superconducting coherence as deduced from simulations of the resistivity*



(**Hypothesised & Objectivized** Domain) in Subject position. In TEXT 2 more than half the choices were made in this last domain. It appears that the author in TEXT 2 is able to craft more precisely the kind of Subjects he needs to communicate his results and persuade his readers of their validity.

Finally, if we briefly compare the Subjects chosen by the author in the two abstracts shown above, in Abstract 1 only one out of four Subjects belongs to the **Participant** Domain, which is the first Subject: *We (calculate ...)*. The other three are within the **Real World** Domain, i.e. *the voltage... and its dependence (are discussed)*, *the {it ab-}plane resistivity (is well described) ...* and *the response ... to currents (is also discussed)*. In contrast, in Abstract 2 three out of four Subjects belong to the **Participant** Domain: *We (propose ...)*, *We (support ...)*, *Our results (give ...)*. The fourth Subject is the pronoun *It* that refers back to a *phase diagram for the vortex structure of high temperature superconductors*, belonging to the more abstract **Hypothesised & Objectivized** Domain.

In sum, the novice seems more reluctant to appear in TEXT 1, where Subjects tend to be more impersonal, less abstract and more commonly used terms than in TEXT 2. In TEXT 2 the analysis suggests a higher degree of authorial presence, as well as an increase in the use of abstract and complex Subjects. Such results would tend to confirm first impressions by physicists when reading the abstracts.

4.2. Comparative Analysis of Contextual Frame

Both papers have approximately 5,700 words and just over 300 Themes each, with marked Themes representing slightly less than 60% of the total Themes. The percentages shown in Table 5 represent all instances of a particular Contextual Frame category divided by the total number of Contextual Frames. Note that for Contextual Frames there is at the outset a choice: writers can choose whether to use them or not. Because they are optional elements in Theme, the results of the present comparative analysis have to be taken as being more tentative than in the case of Subjects that are obligatory elements.

Table 5 indicates some differences in the relative distribution within the four Contextual Frame categories distinguished by Davies.

CONTEXTUAL FRAMES	TEXT 1	TEXT 2
Minimum Logical Relations/Progression	44%	29%
Maximum Logical Relations/Progression	12%	13%
Location	19%	18%
Goal and Process	10%	11%
Evaluation	15%	29%
TOTAL	100%	100%

Table 5: Distribution of Contextual Frames

The main differences concern **Minimum Logical Relations/Progression** (i.e. conjunctions) and **Evaluation**. For **Minimum Logical Relations/Progression**, the difference between TEXT 1 (44%) and TEXT 2 (29%) is mainly due to a prolific use in TEXT 1 of the conjunction *and* by the novice researcher. In his later text, he reduces his use of conjunctions, in particular of *and*, and increases his use of **Evaluation** Contextual Frames.

However, as pointed out above, a comparative analysis of Contextual Frames is necessarily more tentative. In this respect it is interesting to note that when just looking at the abstracts, the Contextual Frame slot is only used in the first abstract, and not in the second. The three **Real World** Subjects of the first abstract are preceded by *In Particular* (**Minimum Logical Relations/Progression**), *We find that* (**Evaluation**) and *To make connection with recent experiments* (**Goal**). In particular, the **Evaluation** Contextual Frame used in Abstract TEXT 1 would go against the trend indicated for the articles as a whole, with nearly double the amount of **Evaluation** Contextual Frame in TEXT 2.

One possible explanation is that as in the Abstract TEXT 2 the author chose to be strongly present in three out of four Subjects; opting moreover for **Evaluation** Contextual Frames would have been excessive in this particular stretch of text.

5. Conclusions

The purpose of this work has been to show how an analysis of Theme can help visualize some of the choices an author makes with increased experience. The analysis suggests that as the author gains confidence he tends to select more interactive Subjects where his presence is either overtly manifested in the text in the **Participant** Domain, or covertly manifested in the more subtle **Hypothesised & Objectivized** Domain. Taking into consideration both domains, overall authorial presence – overt and covert – practically doubles, going from 42% to 77%. A similar trend was noticed by specialist informants when reading only the abstracts. The expert physicist contrasted the impersonal style in Abstract TEXT 1, where out of four Subjects, there is only one **Participant**, with the more personal style in Abstract TEXT 2, where, again out of four Subjects, there are three **Participants** and one **Hypothesised & Objectivized** Subject. The novices also said that Abstract TEXT 2 spoke more to the reader, and made it clear who had proposed the model.

The analysis also indicates that with experience, the author tends to shift his choice of Subjects from the more obvious and commonly used terms of the **Real World** Domain to the more abstract and especially designed terms pertaining to the **Hypothesised & Objectivized** Domain. This trend is also noticeable in the abstracts, where only the first uses **Real World** Subjects. During the interviews, the author said that his understanding of the physics involved had been much greater and that he had been able to organise more clearly TEXT 2. This might explain why in TEXT 2 he was able to compose **Hypothesised & Objectivized** Subjects that expressed his meaning more precisely. In his words “I feel that I can put on paper what I’m thinking”. Specialist informants went along with the author’s perceptions by saying the last abstract was more clearly organised and flowed better than the first.

Finally, when the author decides to mark his Subjects with Contextual Frames, he initially opts for commonplace conjunctive and circumstantial elements. As he gains experience he chooses comparatively more **Evaluation** Contextual Frames, which simultaneously also make him more visible in the text. This could again be related to the fact that the author is more confident about his findings, and has gained sufficient assurance to commit himself.

From a wider perspective the present study could help generalize findings such as that of Berkenkotter et al's which suggest that as students develop into researchers, they become increasingly capable of adapting their discourse to fit the requirements imposed by their research community. The focus here has been on a researcher's choice of Theme: as he gains experience in publishing articles, he appears as increasingly capable of composing the **Hypothesised & Objectivized** Subjects required for effectively presenting his result, of overtly appearing as **Participant** when necessary, and, if need be, of framing these Subjects with the required Contextual Frames. Devising these Thematic elements could be an important step in giving a more "expert" tone to research articles, where author choices are especially strategic as they affect the way in which findings are perceived by the research community at large.

From the pedagogical perspective of teaching writing in research settings, an outcome of the analysis could be to raise awareness of the different possibilities offered by the Subject and Contextual Frame slot to suit different communicative aims. In this respect, it is important to think about how to help novice researchers to devise appropriate Themes in order to improve their chances of being published. A step in the right direction could be the kind of text-based analysis initially used by researchers such as Swales (1981) and Myers (1985), adapted for classroom purposes. An example is selecting articles, considered as being classics within a given research field, for students to examine regarding, for instance, the type of Subjects used by expert writers, and how these writers frame their Subjects. One interesting side effect for the ESP specialist is that when devising such exercises, a dialogue can naturally be established with specialists from other research communities, whose help is often crucial in selecting relevant texts from their fields.

Recebido em: 07/2000. Aceito em: 09/2000.

References

- BERKENKOTTER, C., T.N. HUCKIN & J. ACKERMAN 1991 Social context and socially constructed texts. IN: BAZERMAN, C. & J. PARADIS *Textual Dynamics of the Professions*. University of Wisconsin Press.
- BERRY, M. 1989 Thematic Options and Success in Writing. IN: BUTLER, C.S. & R.A. CARDWELL *Language & Literature – Theory & Practice: a tribute to Walter Grauberg*. University of Nottingham Monographs in the Humanities VI. University of Nottingham.
- _____ 1995 Thematic Options and Success in Writing (revised). IN: GHADESSY, M.. *Thematic Development in English Texts*. Pinter.
- DANEŠ, F. 1974 Functional sentence perspective and the organization of the text. IN: DANEŠ, F. *Papers on Functional Sentence Perspective*. The Hague.
- DAVIES, F. 1988 Reading between the lines: thematic choice as a device for presenting written viewpoint in academic discourse *the ESPecialist*, **9.1/2**: 173-200.
- _____ 1997 Marked theme as a heuristic for analysing text-type, text and genre. IN: PIQUE, J. & D.J. VIERA *Applied Linguistics: Theory and Practice in ESP*. Valencia University Press.
- ENKVIST, N.E. 1973 Theme dynamics and style: an experiment. *Studia Anglica Posnaniensia* **5**: 127-135.
- FRIES, P. 1981 On the status of Theme in English: arguments for discourse. *Forum Linguisticum* **6**: 1-38.
- _____ 1995 Themes, methods of development, and texts. IN HASAN, R. & P. FRIES *On Subject and Theme: A Discourse Functional Perspective*. John Benjamins Publishing Company.
- GIBSON, T.R. 1993 *Towards a Discourse Theory of Abstracts and Abstracting Vol 1 & 2*. Monographs in Systemic Linguistics Number Five. Department of English Studies, University of Nottingham.
- GOSDEN, H. 1993 Discourse functions of subject in scientific research articles. *Applied Linguistics*, **14.1** Oxford University Press.
- _____ 1996 *A Genre-based Investigation of Theme: Product and Process in Scientific Research Articles Written by NNS Novice Researchers* Monographs in Systemic Linguistics Number Seven. Department of English Studies, University of Nottingham.

- HALLIDAY, M.A.K. 1994 *An Introduction to Functional Grammar*. Edward Arnold Publisher.
- HYLAND, K. 1998 Persuasion and context: the pragmatics of academic metadiscourse. *Journal of Pragmatics* **30**: 437-455.
- MAURANEN, A. 1996 Discourse competence – evidence from thematic development in native and non-native texts. IN: VENTOLA, E. & A. MAURANEN *Academic Writing: Intercultural and Textual Issues*. John Benjamins.
- MYERS, G. 1985 Texts as knowledge claims: the social construction of two biology articles. *Social Studies of Science* **15**: 593-630.
- RAVELLI, L.J. 1995 A dynamic perspective: implications for metafunctional interaction and understanding of theme. IN: HASAN, R. & P. FRIES *On Subject and Theme: A Discourse Functional Perspective*. John Benjamins Publishing Company.
- SWALES, J. 1981 *Aspects of Article Introductions*. The University of Aston, Language Studies Unit.
- _____ 1990 *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- TARONE, E., S. DWYER, S. GILLETTE & V. ICKE 1998 On the use of the passive and active voice in astrophysics journal papers: with extensions to other languages and other fields. *English for Specific Purposes* **17.1**: 133-136.
- TAYLOR-TORSELLO, C. 1996 On the logical metafunction. *Functions of Language* **3.2**: 151-183.
- THETELA, P. 1997 Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes* **16.2**: 101-118.

Acknowledgement

I would like to thank Florence Davies for many enlightening discussions. Any shortcomings are my own.

Ann Montemayor-Borsinger holds an MSc in Economics (UNAM) and an MEd in TEFL (Bristol). She is currently writing up her PhD thesis on the development of writing skills within research settings. She is a lecturer of English at a centre for physicists and engineers in Argentina, and worked as a journalist and as a teacher of English in several European and Latin American countries.