Syntactic differences in oral and written scientific discourse: the role of information structure

Shirley Carter-Thomas, Elizabeth Rowley-Jolivet

HAL Id: halshs-00199788

https://shs.hal.science/halshs-00199788v2

Submitted on 3 May 2008

Prepublication version

Syntactic differences in oral and written scientific discourse: the role of information structure

Shirley CARTER-THOMAS
Institut National des Télécommunications, Evry

Elizabeth ROWLEY-JOLIVET
ESPEO, Université d'Orléans

This article analyses some of the differences in syntactic structure that emerge from a comparison between a series of oral presentations, given at an international conference in physics by English native speaker scientists, and the corresponding articles, published in the proceedings of the same conference. Our motivation was not only to provide a statistical account of these differences but also to propose some explanations for the differences observed. In our view, a study of the syntactic differences between these two genres of scientific discourse can be usefully linked to the linguistic notion of information structure. By studying information packaging arrangements we can gain added insight into why, in a specific context, certain syntactic forms are preferred. The syntactic forms that will be particularly focused upon in this article are extraposition, existential constructions, inversion and pseudo-clefts.

Nous examinerons dans cet article certaines différences syntaxiques qui se profilent à partir d'une comparaison entre une série de communications scientifiques orales faites par des chercheurs scientifiques anglophones lors d'un congrès de physique et les articles de recherche correspondant à ces communications, tels que publiés dans les Actes du même congrès. Notre objectif n'est pas seulement de fournir un inventaire de ces différences mais également d'en proposer quelques explications. A notre avis, une analyse des différences syntaxiques relevées dans ces deux genres de discours scientifiques peut bénéficier d'une réflexion sur la notion linguistique de structuration informationnelle. Une telle approche permet en effet de cerner de plus près l'influence du contexte énonciatif sur les formes syntaxiques employées. Nous examinerons en particulier l'utilisation de quatre structures syntaxiques : l'extraposition, les énoncés existentiels, les inversions et les pseudo-clivées.

Introduction

Different types of scientific discourse make use of a variety of syntactic structures in differing proportions. We have chosen to examine some differences in syntactic structure that emerge from two distinct discursive situations: the paper delivered orally at an international physics conference and the corresponding article as published in the conference proceedings[1].

The motivation for this study is twofold. First, while there exists a considerable body of work on written scientific discourse, particularly the research article, we know much less about the syntactic features of the oral equivalent of the research article, namely the conference presentation[2]. It is important to know more about the differences between how scientists write about their work and how they speak about it. Many of us in the GERAS community are involved in the design and teaching of academic presentation skills courses and the research article cannot be used as a model for giving a successful oral presentation - the two, we would contend, are different genres. A first step was therefore to evaluate certain syntactic differences between the two genres in quantitative terms. The issues centre on preferred structures, and a basic distinction is drawn between those common to the two genres and those which occur in only one of the two genres examined. Frequency of occurrence is therefore a key consideration.

Our approach however is not restricted to the statistical level. The second objective for our study was to find explanations for the differences observed in the two genres and evaluate them in qualitative terms. What are the practical imperatives that lead to one structure being preferred to another? What is different about how the message is conveyed to a live audience in one case as opposed to the reader of the article in the other? Can the same type and quantity of information be transmitted? Why is certain information omitted or, on the contrary, highlighted? At the centre of such preoccupations is the linguistic concept of information structure. It is our contention that many of the syntactic differences between the oral and written conference papers can be usefully explained by looking at how information structure functions in these two very different communicative contexts.

1. Information structure

Information structure deals with the way the information conveyed by the discourse in question is packaged into informational units within and between clauses by the enunciator (writer or speaker), in order to satisfy two communicative goals: i) making the information conveyed by the discourse easier for the reader/hearer to understand; ii) indicating what the enunciator considers to be the most important or salient items of information.

---

[1] See Appendix 1 for details of the conference and a list of the presentations and articles in the corpus. Full information about the criteria used to select the conference and the presentations can be found in (Rowley-Jolivet 1998, Chap. 4 : 195-223).
[2] There is some interesting work on university lectures, but this is didactic discourse and therefore rather a different genre from research presentations to peers (Flowerdew, ed. 1994).

Generally speaking the writer/speaker will place at the beginning of the clause what he is speaking about, the *theme*, which usually contains given (or recoverable) elements that the listener can easily identify. New elements of information concerning this theme will figure afterwards in the rheme. As a rule the clause will gradually gain in communicative importance with the most communicative dynamism, in the words of Firbas (1992), falling on elements which carry a substantial new information load in the rheme. In oral discourse in English this gain in communicative dynamism usually culminates in the focal nuclear stress placed on an item in clause final position. The rise in information load also often corresponds to the actual size and complexity of clause items, with long complex elements tending to occur towards the end of the clause, thus easing the processing burden of the receiver, a phenomenon frequently referred to as 'end-weight'.

This type of organisation with *given* information first and *new* information second, reinforced by the principle of end-weight and final focal stress, is considered to be the unmarked option for organising the information. On the syntactic level unmarked information packaging arrangements are often rendered in English by an SVO structure, where the grammatical subject of an active verb provides the thematic anchorage, with the verb complex and object constituting the *new* or rhematic part of the clause[3]. However this SVO order is not always sufficient or suitable to create a coherent text. In order to respect unmarked information structure the enunciator may be obliged to use a different syntactic arrangement. He may also choose to deliberately reverse the usual information packaging arrangements in order to give certain elements of information particular salience or focus. For these reasons, the enunciator needs at his disposal a battery of linguistic resources so as to indicate unambiguously, either through prosody or syntax, the information packaging arrangement desired.

We have chosen to focus on the syntactic resources available for indicating information status and focus. At issue here are specialised syntactic structures, such as the passive, cleft constructions, extraposition, inversion and existential *there*, structures that specifically enable the enunciator to manipulate different clause constituents and thus explicitly denote a particular information packaging arrangement. Whilst these structures are theoretically available at all times, their deployment depends to a large degree on the choice of a particular enunciator, within a particular mode (oral or written) and communicative context.

2. Research Design

In order to investigate the precise impact of the communicative context and the mode on information packaging strategies and the consequent choice of syntactic structures employed, a corpus of strictly parallel texts was collected. This corpus comprises 9 conference presentations given by native speakers of English at an international conference in physics, and the 9 corresponding articles published in the proceedings of the conference. The oral presentations were recorded on video and then transcribed. The parameters held constant between the oral and written sub-sets are therefore as follows: the scientific research topic (identical for each presentation-article pair); the author (the

---

[3] See (Carter-Thomas 2000, Chap. 2) for a more detailed account of information structure and the interaction between syntax and information structure.

speaker is in all cases one of the authors of the written paper); the audience targeted (other researchers in the same physics speciality); the occasion (the physics conference). The only variable parameter is the mode, oral or written.

The first step in the analysis consisted in a quantitative comparison of the articles and presentations to establish the relative frequency of occurrence of the specialised syntactic structures in each sub-set. These comparative data are presented in the next section.

3. Quantitative Data

Table 1 gives the occurrences of each structure in the articles and presentations (columns 2 and 3 respectively). For comparison, columns 4 and 5 provide information on the frequency of occurrence of the same structures in two registers – written academic prose (ACAD) and spontaneous conversation (CONV) - taken from the corpus-based Longman Grammar of Spoken and Written English (Biber *et al* 1999, hereafter referred to as LGSWE). The signs +, – and ≅ signify, respectively, that the register in question contains typically more, less, or approximately the same number of occurrences of a given structure.

It should be noted that while the two sub–sets of articles and presentations are almost identical in length (19,475 and 19,502 words respectively), there is a much lower number of clauses in the proceedings articles (1,489 compared to 2,063 in the presentations), due to a lower verbal density : the verbal densities, expressed as the number of verbs per 1000 words are, respectively, 76.5°/oo (articles) and 103°/oo (presentations).

	Articles	Presentations	ACAD	CONV
Passives • mean percentage	480 32%	172 9%	+ 25%	– 2%
Extraposition	54	9	+	–
Existential *THERE*	13	51	≅	≅
Inversion	0	63	–	+
IT-Clefts	2	2	+	–
WH-clefts	0	31	–	+

Table 1 : Quantitative comparison of conference proceedings articles and oral presentations

(Mean percentage is given for the passive only; it is not indicated for the other specialised structures due to the small numbers involved.)

It can be clearly seen from Table 1 that there are marked differences in the frequency of occurrence of the structures in the articles and oral presentations. Structures that are heavily used in one mode are absent, or used only to a minor degree, in the other mode, indicating that the strategies used to manipulate information structure are strongly mode-dependent.

The **passive** will not be examined in detail in this paper, as it has been extensively studied elsewhere (Barber 1962; Hanania & Akhtar 1985; Banks 1993 : 13-59; Tarone

et al 1998), due to its high frequency of usage in scientific research articles. Note that the percentage of passives in the proceedings articles (32%) agrees with the figure of over 30% found by other studies of scientific journal articles, while the somewhat lower figure (25%) in LGSWE can be explained by the fact that the register ACAD includes not only scientific articles but also articles from other academic fields, where the passive is not used to the same extent. It should be stressed, however, that the passive plays an important role in information packaging, or more precisely in enabling the writer to match unmarked information structure with syntax (Carter-Thomas 1998 : 76-78). By using a passive the writer can thematise and present as 'Given' what would normally have been a syntactic object (and thus part of the rheme) in an active clause. The use of a passive construction thus allows the writer to maintain thematic continuity between clauses, as an inanimate element already mentioned in a previous clause (either in the theme or rheme) can be thematised thanks to a passive in a subsequent clause.

The oral presentations, in contrast, contain a relatively low percentage of passives (9%), even though the subject-matter is the same, indicative of the strong influence of the oral mode on the syntactically preferred structures. This difference is highly significant ($p = 0.001$). The low proportion of passives is correlated with a correspondingly high proportion of personal pronoun (animate) syntactic subjects used as themes in the oral presentations. A more detailed comparison of the use of the passive in conference proceedings and oral presentations can be found in (Rowley-Jolivet 2001).

Extraposition, another feature generally associated with scientific discourse, is, like the passive, typical only of the written mode and is rarely used in the oral presentations. **Existential** *there,* in contrast, shows exactly the opposite modal distribution, with four times the frequency in the oral presentations compared to the articles. The remaining two constructions that will be examined, namely **inversions** and **wh-clefts**, show the most marked mode-dependent distribution, with no occurrences in the articles, but an appreciable number in the presentations. Our figures here confirm both the data in LGSWE, where inversion and wh-clefts are found to be characteristically more common in conversation than in written academic English, and those given by Weinert and Miller (1996) on the use of wh-clefts in speech. **It-cleft** constructions are negligible in the data analysed here, with only 2 occurrences (1 per 10,000 words) in each mode, and will therefore not be discussed further. This figure is considerably lower than the figure of 12 per 10,000 words in academic prose given by LGSWE but corroborates the data in (Banks 1995 : 5).

A clear overall trend emerges when the data for the oral presentations are compared with those from LGSWE: whatever the structure considered, the conference presentations bear more resemblance to the features of conversation than to those of academic prose.

4. Qualitative analysis of specialised syntactic structures

4.1 Extraposition

In extraposition, the theme position of the matrix clause is filled by an anticipatory 'it', which refers to a clausal item occurring in the postponed extraposed clause. In our

corpora, extraposition is basically of two kinds: finite 'that' clauses account for 70% of occurrences and non-finite infinitive clauses for 30%[4].

(1)   Based on these results, **it may be possible** to design a uniform BWO where the forward traveling volume harmonics can serve the role of a prebuncher. (A14)[5]
(2)   **It is well-known that** this behaviour arises from the inability of the dipoles within in the materials to respond (A4)

While, however, finite extraposition dominates in our two data sets, extraposition in general is considerably more frequent in our written corpus of articles than in the oral presentations (see Table 1).

There are several possible explanations for the frequent recourse to extraposition in the proceedings articles. Firstly, from the point of view of information structure, extraposition allows the writer to distribute the information load more evenly over the sentence. The use of a non-extraposed structure, as below, would have resulted in a very heavy and complex theme:

?   Based on these results, designing (to design) a uniform BWO where the forward travelling volume harmonics can serve the role of a prebuncher may be possible.

Extraposition makes it possible for the enunciator to avoid such heavy themes and thus respect the principle of end-weight.

Linked to this first point is the fact that a heavy, complex theme as in the non-extraposed version above (?) is very likely to contain elements of new information. By using extraposition the writer ensures their rhematisation. Communicative dynamism thus coincides with the object of the 'that' or 'to' clause placed in the rheme.

Another explanation for the frequency of extraposition in the scientific article concerns the phenomenon of hedging. Extraposition can be used as a hedging tactic. It allows the writer to thematise his own comment on the value or validity of the information in the extraposed clause; his point of view is objectivised by being placed in a separate 'it' clause:

(3)   It is clear from these experiments that ... (A4)
(4)   It is expected that technological innovation will lead to advances in... (A14)

Extraposition is in fact called 'Thematised Comment' in systemic linguistics for this reason (Thompson, 1996 : 129-130). The scientist can thus make a subtle commentary without appearing overtly in his text.

Given the relatively high frequency with which extraposition is used in the articles, and the extremely low frequency in the oral presentations, we decided to compare all the extrapositions in the articles with the passages in the oral papers where the same information is presented, in order to discover what sort of structures are used orally to transmit the same information.

---

[4] These figures confirm those of Banks (1995 : 6-9) as far as the written corpus is concerned. Note, however that finite extraposition also dominates in our oral corpus.
[5] Examples taken from the proceedings articles are referred to by the letter A plus the number of the article, and examples from the oral presentations by the letter P.

However, in two-fifths of the cases, no comparison can be made, because the information in the extraposed clause in the article is omitted in the oral presentations, either because it is part of the literature review, which is very short or sometimes totally absent in oral presentations:

(5)   **It has been reported that** the number of spark gaps can be reduced somewhat by using a high-voltage transformer pulsed from a capacitor bank [2,3]   (A3)

or because it is part of the 'Materials & Methods section' in the article, which is much less detailed in oral presentations:

(6)   **It is important that** the silver tracks are intimately bonded onto the ceramic surface and there are no air gaps   (A4)

Generally speaking, in the oral presentations the speaker only has a limited amount of time, and therefore aims to convey a fairly simplified version of his results, without going into the finer distinctions.

When these omissions are accounted for, we are still left, however, with a considerable number of cases where extraposition is used in the article, but not in the oral presentation, to communicate the same information. In no case, in fact, is the extraposition in the article paralleled by extraposition in the oral presentations.

An initial point that needs making is that one of the main motivations for using extraposition is to avoid employing heavily pre- or postmodified nouns as theme. This consideration does not apply in the presentations, where in half of the clauses the role of theme is filled by a pronoun - which by definition cannot be pre- or postmodified - and in other cases noun modification is much lighter, due to constraints on real-time processing for both the speaker and the audience.

This is illustrated in the following examples, from the oral presentation and article respectively. The scientist is making approximately the same point in each case:

(7)   ... you **want to stress it as far as possible** and **that**'s limited by the breakdown field (P4)

(8)   ... it is advantageous to stress the blocks to the maximum electric field that can be safely tolerated without breakdown. (A4)

However, whereas in the article the writer makes use of extraposition, thus avoiding placing a complex clausal theme in sentence initial position, in the oral presentation the need for such a response or tactic is obviated. The themes of each of the two clauses in the oral presentation are brief ('you' and 'that') and the linear progression operated[6] ('that' refers back to the rheme of the preceding clause) enables the speaker to provide the required explanations progressively to the audience.

Another reason for the frequent use of extraposition in the article is its function as a hedging tactic. Likewise this does not apply in the presentations, where speakers tend to express their evaluation and comments far more openly.

ARTICLE	PRESENTATION
It is shown that	we've demonstrated
It seems that	we think that probably

---

[6] See (Thomas 1999) for the different types of thematic progression used in scientific writing.

It is possible to define	we can define
It seemed that	we've developed an opinion
It has been found that	we find in practice
It can be implied that	we know
It is believed that	we believe
It is shown that	we've shown
It is hoped that X	X was a thought

As the above table makes clear, there is a strong association between the use of the passive and the use of extraposition in the articles: indeed, 26 (i.e. half of the 54) occurrences of extraposition in the articles use a passive verb. In the oral presentations, in contrast, the hedging typical of the research article (avoidance of first person pronouns) does not apply, and researchers freely take responsibility for their own opinions.

The fact that speakers can manage without extraposition for hedging raises the possibility that some of the cases of extraposition found in the articles and that are considered to be hedges, may be, in fact, simply stylistic conventions of the article genre. We find some support for this in the fact that these types of expression - *it should be noted that, it has been shown that* - called lexical bundles by LGSWE are very stable collocations that are found only in academic prose.

A final minor reason for the absence of extraposition in the oral presentations is that – whereas extraposition is used to refer to the Figures in the articles– in the oral presentations, visuals are referred to either with deictics *(here's..., this is...)* or, when a verb is used, it is in the active form:

ARTICLE	PRESENTATION
it can be seen in fig. 3	the top trace here shows
it can be seen that	here's a picture of
it has been found that	the top graph shows

Two points stand out from this analysis of extraposition. Firstly generic conventions would seem to have a strong bearing on the information packaging strategies used. This is both the case concerning conventions of style, as displayed by the use of hedging tactics, and conventions of content, certain sections of the research article being practically absent in the oral presentations. Secondly, contextual constraints, and in particular the real-time processing of information in the presentations also impact on the information packaging strategies that are available to speakers.

4.2 Existential *there*

The existential *there* structure is used four times more frequently in the oral presentations than in the proceedings articles (see Table 1). This result differs from that of LGSWE, which reports a similar frequency for written academic prose and for conversation, namely 50 and 60 occurrences per 20,000 words respectively (as against 13 and 51 per 20,000 words in the data analysed here). The low count observed in our proceedings articles may possibly be an artefact of the small data set, for, as Banks (1995 : 9) points out, there is considerable inter-author variation in the use of the existential *there* construction in scientific research articles. Nevertheless, the marked

difference observed in the data raises the question as to whether the structure plays the same role in the two modes.

We will first examine those functions of existential *there* as an information packaging device which are common to both the written texts and to the oral presentations, before going on to investigate what additional functions it may fulfil in the oral presentations.

In existential *there* constructions, the theme and Given position in the clause is filled by the semantically empty *there*. Any new elements of information can then be placed in rhematic position, in accordance with the traditional Given-New division. The basic function of this structure is therefore not to predicate a property of a referent but to introduce a brand-new referent into the discourse. It is for this reason that Lambrecht prefers to call this structure 'presentational' rather than 'existential' :

> From the point of view of information-structure analysis, the label 'existential' is somewhat misleading. Mere assertion of the existence of some entity is a rather special kind of speech act which is of limited use in everyday communication (...) it is therefore preferable to interpret the function of such sentences as that of presenting or introducing a referent into the 'place' or 'scene' of the discourse and thereby of raising it into the addressee's consciousness, rather than of asserting its mere existence. (Lambrecht 1994 : 179)

An example of this presentational function from the oral data is the following :

(9)   here are some of the factors that influence the behaviour of dielectric, there's obviously chemical composition (...) and then there's external factors like temperature and the applied electric field ...(P4)

In the articles, considerations of information load and end-weight are also relevant. As mentioned in the previous section on extraposition, noun groups in the scientific research article are often heavily pre- or postmodified. Placing the noun group in rhematic position, thanks to the existential *there* construction, avoids overloading the clausal theme, thus respecting the principle of end-weight, as in the following example :

(10)   There is essentially no difference in spectral content between sharpening the leading or trailing edge of the pulse. (A2)

There is no obvious canonical counterpart to existential *there* clauses, but any reformulation of this structure would normally entail presenting certain elements of information as Given :

*?   Essentially no difference in spectral content occurs between sharpening the leading or trailing edge of the pulse.

While considerations of end-weight concern more particularly the articles, a second function of the existential *there* structure that is common to both articles and presentations is its role in ensuring a smooth information flow between clauses. The new information introduced in the rheme of one clause is frequently taken up as theme of the subsequent clause, thus creating a linear progression. Example (11) below is an instance of simple linear progression, and example (12) of split linear progression, in which only one part of the preceding rheme becomes the subsequent theme. This second type of progression, as Dubois (1987) points out, is common in scientific discourse due to the complexity of the noun groups.

(11) In addition to the resonant charging network there is a high impedance power supply at the switched voltage. This maintains the drain voltage at low repetition rates. (A9)

(12) Since the characteristic impedance is $Z_0 = (L/C)^{\frac{1}{2}}$, there is a trade-off between high voltage capability and low impedance. If the oscillator impedance is too high it will be severely loaded by the antenna... (A2)

A final function of existential *there* shared by the articles and the presentations is that of asserting or denying the existence of a physical phenomenon. As Lambrecht, in the previously quoted passage (ibid. 179) remarks, this discourse function is rarely called upon in everyday conversation. In science, in contrast, it is essential to the establishment of scientific facts and theories. Some examples from the data are :

(13) it tapers out slowly so that we know there cannot be a hard arc there (P2)

(14) the Curie point is the point where there is a physical phase change in the material (P4)

(15) ...the paraelectric phase in which there is no spontaneous dipole moment (A4)

In this function, the label of 'existential' structure is fully warranted.

In addition to these three roles, existential *there* appears to have other specific uses in the oral presentations which could explain its greater frequency in this mode. The first of these is its role in enumeration. In the presentations, *there is / are* frequently serves to initiate, to continue, or to conclude an enumeration of various features. In this function, it is generally preceded by a sequencer such as *then* ('and then there's a balanced Blumlein': P2), or followed by a quantifier or a numeral:

(16) there are essentially three types of operation you can get ... there's a stable operation ... and there's a third one (P11)

(17) there are several techniques that we use (P9)

(18) there are two other factors (P13)

By organising the listing, it helps the audience to process sequential information. In a written text, the organisation of an enumeration can be expressed by punctuation (colons, semi-colons, bullets, etc.) and by page layout.

A second function of existential *there* in the oral presentations is that of discourse organiser. In this role, the scope of *there* would seem to be considerably wider than in an enumeration, in that it serves to open or to close a segment of discourse dealing with a given topic. When preceded by *now*, *there* serves to initiate a whole new topic. This is closely related to the function traditionally attributed to existential *there* of introducing a new element into the rheme, but its association with *now* seems to further underline its role as a discourse organiser at the beginning of a section or sequence :

(19) now there's lots of different ferroelectric materials (P4)

(20) now there's a lot of data to show (P13)

In the same way, *there* is also used as a discourse organiser to close a topic, by summing it up. In this case, *there* is regularly preceded by *so / so that* and has an anaphoric role of pointing back to a previous segment in the discourse :

(21) so there are practical applications in that area. (P13)

In the written article, in contrast, topic opening and closing can be expressed by page layout (segmentation into paragraphs, section headings, etc.).

Finally, in some cases, existential *there* is used to refer to what is shown on the visual, and is preceded either by an invitation to the audience to look at the visual or by a locative adverbial :

(22)   as you can see, there's almost exact correspondence of the DC part and even of the RF part (P11)

The greater frequency of existential *there* in the oral presentations appears therefore to be related to the specific communicative context of this genre: in order to facilitate information processing for the audience, the speaker needs to indicate clearly, by using segmentation devices, how the discourse is organised; he also needs to refer constantly to the visual channel. The existential *there* construction is one of the specialised syntactic structures which answer these information packaging needs.

4.3 Inversion

In a language such as English in which the SV(O) sequence is the canonical (or unmarked) word order, the reversal of this sequence has traditionally been referred to by the name of inversion. This syntactic manipulation of the clause can be of two kinds:
- Subject/Operator inversion (or partial inversion), as in "Only then did she realise that..."
- Subject/Verb inversion (or full inversion), as in "Here comes the sun".

Both are triggered by some element other than the subject being placed in clause-initial position.

Subject/Operator inversion will not be examined here, as only 3 occurrences were found in the data. The proceedings papers contain one occurrence of grammatically constrained S/O inversion after *as* used for comparison:

(23)   In addition, the degree of nonlinearity induced will be lessened, as will the amount of pulse sharpening.   (A4)

In the oral presentations, there is one occurrence of the same structure :

(24)   laser diodes have been driven as have X-ray diodes   (P3)

and one occurrence of a dependent interrogative clause with S/O inversion :

(25)   so at least you're getting a qualitative picture and you can identify what are the mode competition effects and so forth   (P11)

Inversion of the verb and syntactic subject, in contrast, is a frequently used structure, with 61 occurrences in the oral presentations. There are however no occurrences whatever of S/V inversion in the articles. The fronted element in the oral presentations is, in all cases, a locative adverbial expression, and the verb is the copula BE (see Table 2). This agrees with studies of various registers of English (Quirk *et al* 1985; LGSWE) which have shown that S/V inversion with the locative adverbials *here / there* + BE in the simple present is common in conversation but rare in writing. The pre-subject positioning of the verb is syntactically constrained since the adverbial expression constitutes part of the predicate.

Fronted element	Occ.
here (here's / here are)	44
there (there's / there are)	7
other locative adverbials • on this axis is... • out to the left is... • in the centre is...etc.	10

Table 2 . Subject/Verb Inversion in oral presentations : Clause-initial elements
Oral : 61 occ.   Articles : 0 occ.

The role of this structure in information packaging is, like the existential *there* construction discussed in the previous section, presentational in nature. Its communicative function is not to predicate a property of the referent but to introduce a brand-new referent into the discourse. Once it has been introduced, the referent is cognitively accessible and can then be taken up as the theme of the subsequent clause, as in the following example:

(26)   here's a picture of one of the devices; this one has plastic insulation, it runs from 2 to 300 kV reliably (P2)

The high frequency of this syntactic structure in scientific conference presentations appears to be motivated by three features of the communicative context. The first one concerns the concurrent presence of two channels of communication, the verbal and the visual. As can be seen from Table 2, in all the cases of inversion observed in the data, the introductory term has a deictic function, signalling either the projection of a new visual on the screen, or pointing to a particular place on the visual already displayed. In 80% of the inversions with *here*, for example, the deictic is the first word pronounced by the speaker as he places a new transparency on the OHP, as the following extracts from one presentation (P2) illustrate :

Transparency #	Opening words
8	**here's** a picture of one of the devices
9	**here's** a photograph of the line output
10	**here's** a – I'll try to do this real quickly
11	**here's** a picture of the EMBL source
12	**here's** the output waveform from the EMBL
13	**here again is** received electric field signal

The frequent use of this type of inversion would appear to be directly related to the discourse context of the conference presentation, in which the speaker has constantly to integrate information from the visual and the verbal channels. He has first to focus the audience's attention on the transparency before proceeding to explain what they can see there. We would follow Enkvist (1991) here in calling this a **text strategy**, which in this case overrules the canonical SV(O) order. It is the requirements of the discourse which determine what the appropriate information packaging arrangements are and the consequent use of syntactic structure.

> ... the syntactic arrangement of the sentence is not independent but subservient to a text strategy. The discourse is the father of the text, not vice versa, and each sentence must serve its discourse and text by confirming [sic] to its general strategic principles and thus linking up, deictically and stylistically, with the discoursal situation (Enkvist 1991 : 16)

A second function of this type of inversion is that the initial adverbials serve the purpose of segmenting the text into sections and sub-sections. Each visual projected can be seen as constituting a section or text unit, with *here*-inversion signalling the onset of a fresh unit. In several cases, there are sub-units, or hyponomies, within the unit. These are frequently signalled by the fronting of a more specific locative adverbial, as in example 27. The speaker, after having signalled the start of a new unit with *here's* when he first shows the visual, then goes on, when commenting on the same visual, to signal sub-units with other inversions (*on the left..., on the right...*):

(27) here's a typical confrontation with the experimental and computer simulation data, on the left are the [experimental] measurements of the DC and RF current for each of the cavities (...) and on the right are the corresponding curves one gets from simulations (P11)

This allows the speaker to break up the complex visual information on the transparency into more digestible chunks for the audience.

The third feature of the communicative context which motivates speakers' recourse to inversion concerns the principle of end-weight. In all the cases where inversion is used, the noun group is much weightier than the initial adverbial element, and contains the new information :

(28) Here is two measurements of power density as a function of angle away from the central line of the horn antenna (P14)

Inversion allows the speaker to place the heavy elements towards the end of the clause, following the principle of end-weight, and thus to ease the mental processing effort for the listener. This structure thereby contributes to creating an information flow that is adapted to the constraints of real-time speech.

It is clear that inversion, in scientific conference presentations, is a pragmatically motivated structure, which answers specific discoursal requirements. Note moreover that in several cases (cf. example 28), *here's / there's* functions as a deictic block, with no agreement in number between the subject and copula. Indeed, the term 'inversion' itself – implying that the normal word order has been violated – seems to be a misnomer, in that it is the non-inverted word order which would appear 'abnormal' in this context. The inverted structure cannot therefore be considered as simply a variant on the canonical one.

(29) Here's a typical confrontation with the experimental and computer simulation data

?* a typical confrontation with the experimental and computer simulation data is here

* Here a typical confrontation with the experimental and computer simulation data is

Indeed, in the above example, only the inverted structure would seem feasible.

4.4 Pseudo-clefts

There are practically no occurrences of IT-clefts in our two data sets, but there are, however, a certain number of pseudo-clefts, all in the oral presentations (see Table 1). Pseudo-clefts can be of 2 types (Weinert & Miller 1996, LGSWE 1999): ordinary WH clefts (WH) and reverse WH clefts (RWH).

Ordinary WH-clefts (WH) have the following form :
WH-P' is X
(30)   What we did was....

In reverse WH-clefts (RWH), on the other hand, the clefted constituent occurs in sentence initial position:
X is WH-P'
(31)   it's really important how fat the dipole is. That's **what this slide shows**.

We have 12 occurrences of WH-clefts, and 16 of RWH in our oral data. We will first examine WH-clefts, to study their role in managing information flow.

4.4.1. WH-clefts

Pseudo clefts do not change the canonical information structure pattern theme/rheme. However the division between 'given' and 'new' is highlighted, and as a result the listener is encouraged to pay attention to the 'new' elements, which are given more focus.

(32) So what do I mean by pulse sharpening using non-linear ferroelectric dielectrics?
Well, **what we're talking about here**   is pulses in the voltage range of 10 to 40 kV with
                                                 a peak current in the range of (P4)....
            Given   **New**

This example comes right at the beginning of the talk. The information in the WH part of the cleft is presupposed. The speaker has just raised the issue, which is in fact the subject of the talk, in the preceding rhetorical question. The cleft enables him to draw attention to the forthcoming explanation. The audience is prepared and forewarned that an important explanation is imminent. Without the cleft the information would have received less impact.

The pseudo-cleft would therefore seem to have two main functions:
* it slows down the discourse flow, chunking the information conveyed into two distinct parts, thus helping the listener to process it more easily.
* it enables the speaker to set the stage for the new forthcoming information, thus increasing its salience.
In other words, it clearly identifies the important information, thereby increasing its salience, and at the same time it gives the listener strict processing instructions by making it quite clear against what background a proposition has to be understood. When we compared the pseudo-clefts in the oral papers with the passages in the articles that contain the same information, we found very little explicit highlighting or focusing of this kind. Writers just give the bare facts.

The WH-cleft could also be considered as an interactive strategy that the speaker can use to prepare his audience for the upcoming information, thus helping them to assimilate the most important points of the talk. Indeed, it seems that the presence of the WH- element implies some sort of questioning process. In the above example (32) the speaker actually asks this question himself but in many WH-clefts there would seem to be an underlying presupposed question that either the audience is likely to ask at that stage, or that the researchers asked themselves at that point.

Q   [what did you do then? / What shall we do now?]

(33)   so **what we did here was** took a lot of data with different gas pressures...(P2)

This might help explain why WH-clefts are only found in oral data - either conversation or our conference presentations - as the interpersonal level of question and answer is much more present, even if the speaker is delivering a monologue.

4.4.2. Reverse WH-clefts
Reverse WH-clefts, as already mentioned, are slightly more common then WH-clefts, in our oral data. What is immediately striking is that practically all have a demonstrative[7] as the cleft constituent:

THAT   (34) it's really important how fat the dipole is, **that's** what this slide shows
                                                                                   (P2)
               (35) You can see the larger one in this example in cavity two compared to one and three....So **that's** what the experimentalists can see (P11)

THIS   (36) **This** is kind of like what it all looks like in the lab   (P3)
               (37) **This** is what EMBL looks like   (P2)

However, as can be seen from the examples given here (and confirmed by the other examples in the data), THIS and THAT fulfil different functions. 'THAT' is anaphoric and has a resumptive function; it slows down the discourse by summing up a segment and emphasising its importance before moving on to the next new piece of information. THIS on the other hand, has a deictic function in our corpus, pointing forward to a new element of information displayed visually on the slide.

     (verbal) ← that   this →   (visual)

The fact that demonstratives are frequently used in the oral presentations for these two functions seems particularly significant: the speaker is constantly engaged in pointing, or showing, the way so that the participants can orient themselves in the discourse. THAT relates more to the verbal channel, and is used for summarising ideas expressed verbally, while THIS refers more concretely to the visual channel. Used together the judicious use of THIS and THAT reverse WH-clefts serves to integrate the two channels of communication, the visual and the verbal, and contributes significantly to the clarity of the discourse.

Like the ordinary WH-cleft, the reverse WH-cleft helps to chunk the discourse into manageable units for the audience, and has a framing function. However, while THAT concludes a segment of discourse, THIS opens up a new segment. As with inversions, this text strategy of creating text units, or fairly short portions of texts which are clearly bounded (marked for onset and close) is dictated by the discourse context, and entails the use of certain specialised structures which are very little used in scientific writing. Written discourse is not ephemeral, and perhaps does not require these attention markers and small, intermediary summaries as the discourse progresses.

5. Conclusion

---

[7] Biber *et al* 1999, also refer to these constructions as demonstrative WH-clefts.

This study has compared two scientific genres – written proceedings articles and oral conference presentations in physics – from the point of view of information structure. It clearly emerges from the study that the enunciative context, and in particular the mode, written or oral, has a strong influence on the information packaging resources used by writers and speakers. By holding constant other variables such as subject matter, author, and audience, we have shown that the mode of presentation makes a considerable difference to the way in which the subject matter is packaged and presented. The processing constraints in each mode are radically different, and as a result different information packaging strategies are required. This in turn means that different specialised structures, or the same structures used in different proportions, are needed in order to package information effectively.

Of the four structures analysed in detail here, extraposition is the only one to occur more frequently in the written articles than in the oral presentations. It is widely employed in the articles not only as a hedging device but also as a means of facilitating information processing for the reader by enabling the given and new elements to be distributed more evenly over the sentence. In the oral presentations, in contrast, such a device is used far less frequently because it is not needed: authorial comment is expressed congruently, without recourse to extraposition for hedging, and clausal themes are considerably less complex, due to the constraints of live communication, thus obviating the need for this syntactic structure.

The second specialised structure that is found in both modes, namely existential *there*, shows the opposite distribution, being far more frequent in the oral presentations than in the articles. While its traditional function of introducing a new referent into the discourse is apparent in both modes, its role as an enumeration device appears to be particularly marked in the oral presentations, where it helps the audience to follow the discourse organisation and to process sequential information. We surmised that the same role was fulfilled by typographical devices in the written article. The use of existential *there* in spoken and written science indicates that the same specialised structure may, therefore, be used in both oral and written modes but not necessarily in response to exactly the same information packaging needs.

It is in the use of the remaining two structures, pseudo-clefts and inversion, that the differences in information packaging requirements between spoken and written science appear most clearly. Neither occur in the proceedings articles, but both are frequently used by speakers in their presentations. Pseudo-clefts serve to slow down the discourse flow in the oral presentations, highlighting the division between given and new information and helping the listener to focus on a particular item of information. They thereby function to segment the discourse into smaller, more manageable chunks and to indicate salience. We would therefore interpret the use of this syntactic resource as being linked to the constraints of the real-time processing of information in speech. Live communication is ephemeral in nature; as a result, the speaker needs structures which enable him to refocus the audience's attention whenever necessary in order to ensure that they do not lose the thread of his argument. Writing, on the other hand, is not ephemeral, and so does not require such attention markers to the same extent.

The use of inversion in the oral presentations appears to have a similar function of assigning focus to new information. In addition, this structure plays an important role in ensuring that the principle of end-weight is respected. The main reason for its frequent

occurrence in scientific conference presentations, however, is the communicative context of this genre, namely the need for the speaker to integrate two channels of communication, the verbal and the visual. Inversion is the preferred structure for drawing the audience's attention to the visual display, prior to commenting verbally on it, and for guiding the audience's reading of the visual itself. Indeed, inversion with an initial place adverbial is such a natural packaging strategy in this enunciative context that the very term 'inversion' would seem to be a misnomer.

We believe that by studying information packaging arrangements we can gain greater insight into why, in a specific context, certain structures are preferred or dispreferred. The implications of this in pedagogical terms are numerous. The present study is in fact part of a wider project in which we intend to go on to study the differences between how native and non-native speakers package information. An initial hypothesis is that specialised structures are not always used efficiently by non-native speakers in science. Non-native speakers seem to differentiate less between written and oral modes of scientific communication, tending to overuse, in oral presentations, structures that are typical of writing such as the passive, and not to use structures that are typical of speech such as pseudo-clefts and inversion. In other words, they lack pragmatic competence. As a result, this makes their discourse much more difficult for the listener to comprehend and process. As teachers we have to show our students that scientists do not write and speak about their work in the same way. In a language of fairly fixed word order such as English, the specialised syntactic structures examined here afford a certain latitude on the discursive level, enabling the enunciator to orient or reorient the informational weight of clauses according to the effect required for a given communicative situation. Different enunciative situations require different solutions.

APPENDIX I

1. The data used in this study comprise oral presentations given at the following physics conference; the articles are from the Book of Proceedings of the same conference :

> *Euro Electromagnetics. International symposium on electromagnetic environments and consequences.* Bordeaux, May 30 – June 4, 1994.

2. The titles of the 9 proceedings articles (A) and the corresponding 9 oral presentations (P) are as follows :

A2 – P2   Sandia National Laboratories' high power electromagnetic impulse sources
A3 – P3   Stacked Blumlein pulse generators : versatile sources of high power repetitive waveforms
A4 – P4   High voltage pulse sharpening using nonlinear ferrolectric ceramic dielectrics
A5 – P5   Lightweight, tactical prime power for mobile pulsed power applications
A9 – P9   The development of high peak power solid state pulse generators
A11 – P11   Milo experiments and computer simulations
A13 – P13   Low-voltage, explosive whisker emission cathode studies
A14 – P14   Efficiency enhancement of high power vacuum backward-wave oscillators driven by short pulse and long pulse electron beams
A15 – P15   Nonlinear space-charge waves in an axially uniform waveguiding structure

REFERENCES

Banks, David. *Writ in water. Aspects of the scientific journal article*. E.R.L.A., Université de Bretagne Occidentale, 1993.

Banks, David. « There is a cleft in your sentence : Less common clause structures in scientific writing ». *ASp, 7/10*, 3-11, 1995.

Barber, C. L. [1962] «Some measurable characteristics of modern scientific prose». reprinted in J. **Swales**, ed. *Episodes in ESP*. Pergamon Press, 1985.

Biber, Douglas, S. **Johansson**, G. **Leech**, S. **Conrad**, E. **Finegan**. *Longman grammar of spoken and written English*. Harlow: Pearson Education Ltd., 1999.

Carter-Thomas, Shirley. *Organisation thématique et qualité textuelle : Une analyse des difficultés rencontrées par des élèves ingénieurs francophones lorsqu'ils rédigent en anglais*. Thèse de doctorat, Université de Paris V, 1998.

Carter-Thomas, Shirley. *La cohérence textuelle – pour une nouvelle pédagogie de l'écrit*. Paris : L'Harmattan, coll. Langue et parole, 2000.

Dubois, Betty Lou. "A reformulation of thematic progression typology". *Text 7 (2)*, 89-116, 1987.

Enkvist, Nils Erik. «Discourse strategies and discourse types». In E. **Ventola**, ed. *Functional and systemic linguistics. Approaches and uses*. New York: Mouton de Gruyter, 1991.

Firbas, Jan. *Functional sentence perspective in written and spoken communication*. Cambridge : Cambridge University Press, 1992.

Flowerdew, John, ed. *Academic Listening*. Cambridge : Cambridge University Press, 1994.

Hanania, E. & K. **Akhtar**.. "Verb form and rhetorical function in science writing : a study of MS theses in biology, chemistry, and physics". *The ESP Journal 4*, 49-58, 1985.

Lambrecht, Knud. *Information structure and sentence form*. Cambridge : Cambridge University Press, 1994.

Quirk, R., S. **Greenbaum**, G. **Leech**, J. **Svartvik**. *A comprehensive grammar of English*. London: Longman, 1985.

Rowley-Jolivet, Elizabeth. *La communication scientifique orale. Etude des caractéristiques linguistiques et discursives d'un genre*. Thèse de doctorat, Université de Bordeaux 2, 1998.

Rowley-Jolivet, Elizabeth. «Activating the passive. A comparative study of the passive in scientific conference presentations and research articles». *Les Cahiers de l'APLIUT*, Vol. XX, n°4, 38-52, juin 2001.

Tarone, Elaine, S. **Dwyer**, S. **Gillette**, V. **Icke**. «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*, 113-132, 1998.

Thomas, Shirley. «Thematic networks and text type». *ASp, 23/26*, 139-148, 1999.

Thompson, Geoff. *Introducing Functional Grammar*. London: Arnold, 1996.

Weinert, Regina & J. **Miller**. «Cleft constructions in spoken language». *Journal of Pragmatics, 25*, 173-206, 1996.

Shirley Carter-Thomas est Maître de Conférences à l'Institut National des Télécommunications (INT). Ses domaines d'intérêt sont l'analyse du discours spécialisé écrit et le passage de l'oral à l'écrit, la linguistique fonctionnelle et la syntaxe.
Elle est membre du groupe de recherche OSTERLITS (FRE 2203) du CNRS.

Shirley.Thomas@int-evry.fr

Elizabeth Rowley-Jolivet est Maître de conférences à l'Ecole Supérieure des Procédés Electroniques et Optiques, Université d'Orléans. Ses axes de recherche sont l'analyse du discours scientifique oral et le passage de l'oral à l'écrit, la communication visuelle en science et la sociologie de la science.
Elle est membre du groupe de recherche OSTERLITS (FRE 2203) du CNRS.

Elizabeth.Jolivet@univ-orleans.fr