The influence of the passive on text cohesion and technical terminology
A corpus-based study of research article abstracts from the domain of electrical engineering.

Jérôme Richalot
Institut National des Sciences Appliquées de Lyon (INSA)
jerome.richalot@insa-lyon.fr

Abstract

The passive is used extensively in research article abstracts. It is often agreed that the passive helps focusing on the actual content of the article rather than on its author(s). The abundance of passive sentences is often frowned upon and regarded as unnecessarily burdening scientific discourse. The aim of this paper is to go beyond this assumption and to show that the use of the passive serves a clear purpose: underlining argumentation and therefore strengthening text cohesion.

A study of passive sentences in 83 research article abstracts from *IEEE Transactions on Ultrasonics, Feroelectrivity and Frequency Control* (UFFC) is conducted. Passive sentences are extracted and classified. The first criterion for classification is the part of the abstract the sentence belongs to (IMRaD). A study of verbs, their semantics and informative content is then carried out. Subsequently, the grammatical subjects of passive sentences are examined. The appearance of technical terms from the domain of electrical engineering and their prospective or retrospective orientation receive particular attention. An attempt is made at characterising subject noun groups.

The corpus-based approach helps underlining the blandness of verb semantics as opposed to the richness of subject noun groups in sentences that often appear unbalanced. It is thereby established that the use of the passive in the 83 abstracts studied only partially matches the description of the passive based on general English. Indeed, in the abstracts that make up the corpus, passives do not seem to either characterise the grammatical subject of the sentence or put any particular stress on the result of the process as referred to by the verb.

It is suggested that, by receiving so much emphasis, the subject noun group becomes the centre of predication for the sentence. The "castaway" verb group only carries little information and can sometimes even take on a purely metalinguistic role. This topicalisation process is seen as particularly efficient in helping authors structure their abstracts and therefore helps strengthen text cohesion.

Technical terminology seems to benefit greatly from this use of the passive. Terms appear in a co-text helping to situate them in the author's conceptual framework thus contributing to a better understanding of the specialised field. Both the corpus approach and the results yielded could be seen as valuable contributions to terminology processing and TEFL.

1. Introduction[1]

Many books have been dedicated to "English for science" aimed at specialists of other areas than English. The ones I am referring to in this introduction all deal with abstracts with more or less details.

In the instructions to contributors of IEEE Transactions on Ultrasonics Feroelectricity and Frequency Control (UFFC), it is clearly specified that "Each contribution must contain an abstract (not more than 200 words for papers and 50 words for correspondences and Letters)." This is the only instruction given to contributors on the abstract itself. However, in the confidential proof-reading form given to each author, under the heading Summary of evaluation, one of the first criteria regarding presentation is "Is the summary an adequate digest of the work reported in the paper?" This remark underlines the importance of the abstract as well as its conventional aspect. Robert A. Day (1989: 28) goes as far as considering the abstract as a "miniversion of the paper."

The aim of the abstract is to help the distribution of the article it refers to hence this statement by Day (1989: 28) "The abstract should provide a brief summary of each of the main sections of the paper." Lobban et Schefter (1992: 47) add another characteristic to the abstract which they see as a "self-contained synopsis of the report". Sites (1992: 113) also endorses this remark stating that "abstracts are self-sufficient". These elements of definition however concise they may seem underline the secondary aspect of the abstract. Day (1989: 28) uses the term "secondary publication" with reference to secondary services in which an abstract can be published independently from the article it sums up (Biochemical Abstracts, Chemical Abstract, etc.) It therefore seems appropriate to consider the abstract as a metatext. It accompanies research articles but usually has its own layout. In the same way

[1] Quotations from French references have been translated and marked *[translation]*. The French is given in footnotes when necessary.

as literary criticism it is based on another text. The subject of the abstract is the article. According to Lobban and Schefter (1992: 47) “The Emphasis in an abstract is on the results and conclusions. It should have only the objectives from the Introduction, and only a brief reference to the Materials and Methods (unless the experiments focused on methods).” They explicitly refer to the Introduction, Methods and Materials, Results and Conclusion parts of the article. In many cases the structure of the abstract does indeed reflect the IMRaD structure of the article. Day (1989: 28) clearly lays out the following principles: “The Abstract should (i) state the principal objectives and scope of the investigation, (ii) describe the methodology employed, (iii) summarize the results, and (iv) state the principal conclusion.”

Sites (1992: 113) makes a difference between informative abstracts and descriptive abstracts qualifying the latter of “little more than prose table of contents”. In the study of passive conducted here I shall try to show that but for the negative connotation Sites’ remark is very appropriate.

The very functional elements of definition given above are usually followed by a series of rules and regulations for a “good abstract”. Here are just a few examples. For Day (1989: 159) “Most of the abstract should be in the past tense because you are referring to your own present results.” For Vernon Booth (1993: 13), “The passive voice, commonly used to describe results, sometimes makes clumsy constructions. Turn a passive voice to direct style when you can.” These instructions seem difficult to apply and sometimes even to understand. More pragmatic advice such as Lobban and Schefter’s (1992: 57) is certainly preferable: “However, there are divergent and often strong opinions about whether use of the passive is good or bad. [Ask what your professor wants !]. One of our colleagues aggressively promotes the passive. (“The report is about the experiment, not about you ! she exclaims.) Another colleague equally hotly denounces the use of the passive as “Victorian prudery” that leads to “committee writing [style], ponderous discussions, and avoidance of responsibility.” ”

In the literature for specialists of other areas than English there seem to be as many “recipes” for a good abstract as books published on the topic. Still however awkward some of the instructions might sound, they show that abstracts do have their own style. If the authors quoted above seem to find it difficult to clearly define this style — Lobban and Schefter (1992: 57) even resort to a so-called “scientific etiquette” — they at least fully recognise its existence.

2. A linguistic approach

Since linguistic studies are one of the only research areas in which the object being studied is also the one used to report on those studies, a clear metalanguage should be defined. I shall therefore first set out as clearly as possible the definitions for the terms used throughout this study. In the case of the passive, this seems all the more important as the metalanguage gives us a first hindsight into the actual operations at work when building a passive sentence.

As Henri Adamczewski (1993: 180) points out the very terms of *passive verb* or *passive form* do not seem appropriate. They tend to restrict the passive to a phenomenon mainly concerned with the verb group. The verb group is indeed on the forefront of morphological and syntactic changes and no research on the passive can possibly be conducted without a thorough study of the verbs being used in the passive, however this study focuses on sentences (or *utterances*[2]) and I shall therefore, depending on the context, use either *passive sentences* or *passive voice* in which *voice* refers to subject-verb-object ordering. Although it seems awkward to do without the very term *passive*, particularly in language teaching, linguistic grammars have made it difficult to actually justify its use. Jean-Rémi Lapaire and Wilfried Rotgé (1991: 364) propose the following etymology. *Passive* as an adjective comes from the Latin *pati* for “suffer” or “undergo”. *Passive* would therefore be motivated by the appearance, in a passive sentence, of a subject undergoing the process described by the verb. It seems obvious that this feature greatly depends on the semantics of the verb used. In (1), *John* does indeed undergo a process.

(1) John was hit by Paul.

But the situation is altogether different in (2).

(2) John was greeted by Paul on the doorstep.

Furthermore, the emergence of a subject undergoing the process described by the verb is not, in itself, a prerequisite for a passive sentence as in (3).

(3) John suffered the same fate as his brother.

[2] Marie-Line Groussier and Claude Rivière define the *utterance* as a basic unit of study. “An utterance is produced throughout a unique *uttering act* by an *utterer* for whom it makes up a whole, which has consequences on prosody.” [translation] Utterances are therefore not exclusively oral. Groussier and Rivière point out that utterances and sentences often correspond. But featuring a verb is not a prerequisite and they consider the French “*Moi, un voleur !*” as an utterance.

The inaccuracy these three basic examples underline most probably arises from the confusion between linguistic and extralinguistic, or in other words between a universe of experience and one of representation. André Joly's and Dairine O'Kelly's remarks (1990: 152 [translation]) offer significant help: "[...] subject and object are grammatical functions, they refer to imaginary behaviours existing as thoughts and more particularly thoughts related to language. On the contrary, the impressions which define the notions of agent or patient do not correspond to functions – that is to roles played by a noun or its substitute in a sentence –, but to situations, states originating in the world of sensations."

This distinction enables me to clearly disconnect the active voice from the passive voice. The active/passive couple is an inheritance from descriptive grammars in which a sentence with a transitive verb "in the active voice" had its equivalent "in the passive voice". Whenever "John was eating an ice-cream" that very ice-cream was systematically "being eaten by John." The trend was most certainly further accentuated by generative grammars. In *Syntactic Structures* (1957: 80), Chomsky does wonder about the possibility of considering passive sentences as part of the kernel of English, but rejects the idea faced with the complexity of the transformations then required: "When we actually try to set up for English, the simplest grammar that contains a phrase structure and transformational part, we find that the kernel consists of simple, declarative, active sentences (in fact, probably a finite number of these), and that all other sentences can be described more simply as transforms."

Lucien Tesnière (1959: 242) opened new perspectives considering the passive in the wider scheme of transitivity. He defined four *diatheses*[3]:

- the active diathesis. "Alfred hit Bernard."
- the passive diathesis. "Bernard was hit by Alfred."
- the reflective diathesis. "Alfred killed himself."
- the reciprocal diathesis. "Alfred and Bernard killed each other."

Each diathesis is a virtual realisation of transitivity. Adamczewski (1993: 185) as well as Joly and O'Kelly (1990: 151) praised the coherence of Tesnière's model but also underlined its limits. In Tesnière's model, the passive diathesis remains closely linked to the active diathesis.

Nowadays, the so-called subordination of the passive to the active is widely recognised as irrelevant and a few simple examples effectively stress the point. Adamczewski (1993: 186 [translation]) sees the passive as "a phenomenon inscribed in the workings of the human language." He presents the following verb classification[4]:

- Verbs with an active or a passive orientation (also known as notional passives)
  (4) They grow tomatoes.
  (4') The tomatoes won't grow this year.
  (5) She reads the play well.
  (5') The play reads well.
- Exclusively passive verbs
  (6) He was addicted to heavy drinking.
  (7) She is bound to find it if you're not careful.
- Non-passivable verbs
  - subject oriented and non reversible (middle verbs)
  (8) The young soldier weighed six stones.
  (8') * Six stones were weighed by the young soldier.
  - with a cognate object
  (9) She shrugged her shoulders.
  (9') * Her shoulders were shrugged by her.
  - reflective
  (10) John praised himself.
  (10') Himself was praised by John.

Still, I do not intend to reject as a whole previous studies on the passive and will not deny that the active voice often sheds light on the operations being carried out when a passive sentence is produced. The metalanguage used in the studies I quote as references is often reminiscent of the traditions of descriptive grammars. When Adamczewski (1991: 187 [translation]) proposes the following model for the passive voice:

N_2 {*be*} + past participle {*by*} N_1

the very labels N_2 and N_1 are reminiscent of the N_1-Vt-N_2 model for the active voice.

[3] The examples are a translation of Tesnière's.

[4] The examples are Adamczewski's.

Joly and O'Kelly (1990: 155) do propose the term *resultative voice*[5] which is however first introduced in a parallel with their *operative voice*.

It therefore seems appropriate for Lapaire and Rotgé (1991: 365) to point out an ambiguous dependence of the passive on the active. And it should come as no surprise that Adamczewski (1993: 181 [translation]) himself refers to the active voice as a structure "perceived as primary though with a certain unease."

However, Adamczewski goes on explaining the reasons for this uneasiness. He sets out the minimal ternary model N_1 Vt N_2 as corresponding to the traditional SVO model of transitivity. He argues the SVO model of transitivity can be considered as primary being "as close as possible to the extralinguistic universe: in the SVO model, the autonomous subject uses the semantics of the verb to somehow modify or affect the object of the verb. Transitivity can be nothing else than this movement or dynamics, which in itself is the language parallel of man's action on the material world." (1993: 186-187 [translation])

Joly and O'Kelly (1990: 154 [translation]) quote Moignet (1974 and 1981) adding that "any event expressed by a verb has a *causation*: an initiating reference. It also has an *effection*[6]: upon completion of the operation represented by an event, a result is necessarily attained." The model they propose (figure 1) features the same "movement" or "dynamics" referred to by Adamczewski:

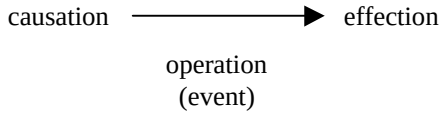

Figure 1

The passive therefore does appear as a secondary construction, the reversing not of the constituents of an hypothetical active sentence but of a predicative movement. The "affected" – for which we might venture the term "effected" - becomes the origin of the sentence, the second term of the equation being the verbal predicate. The newly founded relation between the *effected* (N2) and the verbal predicate (Vt) is the very *raison d'être* of the passive. Adamczewski (1993: 187 [translation]) puts it as straightforwardly as possible "given an abstract clause: N1 Vt N2, N2 has been chosen as the origin of the utterance and what the utterer is aiming at is the N2-Vt relation, N1 being absent." The passive can clearly be considered as an operation pertaining to *thematisation*.

The syntactic components of a passive sentence are most interesting with regards to the very nature of this phenomenon. As seen above (3) having a subject undergoing the process described by the verb is not enough for a sentence to be considered passive. Lapaire and Rotgé (1991: 348 [translation]) consider the use of *be* as a "capital phenomenon" through which "the predicate [is] formally "separated" from the grammatical subject [and] is placed under the control of the utterer." They propose (1991: 347-348 [translation]), a very interesting brief diachronic study of *be* which should certainly be considered here. According to them the original meaning of *be* was "to occupy a place" but it soon became "to have a place in the universe, thus to exist." They therefore define *be* as "a marker or instigator of existence, either absolutely (I think therefore I AM), or relatively enabling the utterer to assign the grammatical subject a permanent or temporary characteristic (e.g. I AM tall, she IS angry.)" They do see in *be* "an actuating power – in the philosophical meaning of the term "give a reality to...", "state to the world the reality of..." – and a characterising power: the utterer takes a defining point of view with regards to the grammatical subject by assigning it one or several traits he deems essential or at least relevant (at least in a given situation)."

This hints at the double effect of *be*. In terms of position in the sentence, between the grammatical subject and the verb, it breaks up the direct predicative link and freezes the grammatical subject. Lapaire and Rotgé (1991: 368 [translation]) quite efficiently use the term "freeze frame". They remind us that this impression is further accentuated by the fact that "in terms of semantics, the subject of a passive construction does not control the lexical verb." Let us add that most passive sentences are in fact agentless, which might have a role to play in freezing the grammatical subject.

The second effect of *be* is a highlighting one. As can be understood through the previous paragraphs, *be* is an attributive auxiliary, it looks towards the grammatical subject of the sentence and actualises or characterises it. The model Joly and O'Kelly (1990: 161) propose (figure 2) underlines this:

(11) The door is locked.

[5] "voix résultative" in French as used by Joly and O'Kelly.

[6] *Causation* and *effection* in French as used by Joly and O'Kelly.

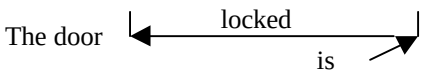

Figure 2

Trying to replace *be* with another auxiliary offers interesting contrasts. Statistically *get* is the most likely candidate but also possible are *become, remain, seem, appear, feel, look* – with an effect close to that of *get*. Let us consider the following examples:

(12) He got run over by the whole pack.
(13) He became irritated and soon lost his temper.
(14) He seemed affected by the sad news.
(15) He appeared relieved when she left.
(16) He feels overwhelmed.
(17) He looked disappointed.

Although these sentences remain subject-oriented they tend to let us see the process described by the verb (Vt-EN) as unfolding. Joly and O'Kelly (1990: 161 [translation]) do remind us of the existence of the auxiliary *weorðan* in old English meaning *become* the use of which, in passive sentences, was close to the one I describe here. Adamczewski (1993: 189 [translation]) points out that "in German *werden* (a metalinguistic operator derived from werden = become) is used and not *sein* (*be*)."

The semantics, or semic programme[7], applied to the grammatical subject is of course that of the transitive verb (Vt-en). The very form of the verb, the past participle, is revealing. We might instinctively recognise Vt-EN and V-ED (preterit) as quite close, however the two forms are quite different. According to Lapaire and Rotgé (1991: 446 [translation]) "thanks to -EN, the predicate is linked to a past period of time (cf. *past* participle)." Joly and O'Kelly (1990: 159 [translation]) offer a slightly different perspective. "[...] The form of the past participle, e.g. *broken*, marks the verb as having reached the end of its tension: it is its *effection* form. The verb is then reduced to the result of the operation called operatively by *break*." The past participle form of the verb therefore signals the operation of the verb as "effective"; from this point on, what is being considered is the result of this operation.

The syntax of a passive sentence enables us to look back on the operations carried out by the speaker. Off course, it supports the potential of the passive as described above. With the help of *be* the speaker enters the predicative relation and makes it his own. From this point on, he handles discourse objects and gets further away from the extralinguistic universe.

Lexis as defined by Groussier and Rivière (1996: 112-113 [translation]) could help summing up what I have developed so far. Groussier and Rivière make a difference between *predicable lexis* and *predicated lexis*[8]. The predicable lexis is the notional model of a process in terms of actors and it is up to the speaker to assign a value to each of the actors described in the model. The label given to the first actor "ξ0" which becomes "a" or "C0" once assigned a value, symbolises its position as origin of the message. The predicated lexis is therefore the result of predication applied to the predicable lexis." Passive predication is "the choice as first argument, not of the actor in the ξ0 position of the lexis model (the source of the primitive relation), but of the actor in the ξ1 position (aim of the primitive relation)."

So far I have tried to show that using the passive is in fact setting up a predicative model the aim of which is to promote the secondary notion in the predicable lexis to a primary status in the predicated lexis. The operation means defining the starting point of the sentence, its theme. In this study of highly specialised but also very controlled texts (see definition and characterisation of the abstract above), aimed at a well defined discourse community, thematisation takes on an even greater importance. For the authors of the abstracts which make up the corpus I am studying choosing the theme of a sentence means deciding which information should be considered as given, as determined enough to be shared with their discourse community, with no further explanation. I therefore expect these theme segments to reveal some of the characteristics of the discourse community the abstracts are aimed at. This could be particularly interesting with regards to technical terminology shared by the discourse community. The theme segments should feature a high density of technical terms related to the specialised field being considered.

[7] *Semic programme* refers to the semantic features or *semes* of the verbs. These traits represent the full semantic potential of the verb only part of which realises at predication.
[8] "Lexis prédicable " and "lexis prédiquée" as used in French by Groussier and Rivière.

3. A corpus-based study

The purpose of the following corpus-based analysis is to test my hypothesis on the passive. No in-depth study of a corpus of research articles is in fact necessary to underline the frequency of passives. A mere count of the occurrences of *be* (be/being/is/are/was/were) over several articles will actually yield very little noise. I decided to conduct a study of passive sentences in research article abstracts. I selected 83 abstracts from *IEEE Transactions on Ultrasonics, Feroelectrivity and Frequency Control* (UFFC)[9] that is the abstracts from the three issues of the first semester of 1999. Thanks to Mike Scott's Wordsmith tools suite I extracted the passive sentences from the corpus. Using the Concord tool, passive sentences can be extracted automatically concordancing for *be* or any of its derived forms *being*, *is*, *are*, *was*, *were* and **ed* or any of the infamous "irregular past participles" to its right. The concordance should exclude past participles left of *be* and look for the past participle two to three words right of *be*. Minor post editing is necessary.

Using this set up, I extracted 358 passive sentences from ABCORP. Throughout the study of these 358 sentences I shall use and adapt the terminology first used by Adamczewski. I shall therefore refer to the grammatical subject of the sentence and to the verbal predicate (a transitive verb) using respectively N_2[10] and Vt-en[11].

I shall first try to classify the Vt-en's used in the passive sentences in ABCORP. Then I shall study their N2's to try to get a better understanding of the Vt/N2 relation which I defined above as being the *raison d'être* of the passive.

4. Categorisation of transitive verbs in passive sentences

The first data that can be used upon processing of the corpus are the raw results given by *Concord*. From the 358 passive sentences I have extracted 368 occurrences of Vt-en's. The difference is due to sentences in which two or more Vt-en's are co-ordinated by *and* or *or* without *be* being repeated, as in the two examples below.

(18) Frequency spectra and modes are computed and examined.

(19) These effects can be reduced or eliminated by using narrow-band experiments.

The 368 occurrences correspond to 145 different verbs. Their frequency is interesting. On the one hand, 30% of those Vt-en's appear only either once (81)[12] or twice (23). On the other hand the four most frequent Vt-en's represent 20% of the 368 occurrences. These four verbs are in fact somehow emblematic of the IMRaD structure. They are *use* (29) for the Method section, *present* (20) for the Introduction section, *obtain* (16) for the Result section, and *propose* (13) for the Discussion section. However this mere frequency count should not lure us into hasty conclusion.

Table 1 below maps those verbs over the IMRaD part they appear in. It also adds other less frequent verbs with a semantic content close to them. The methodology I am using here is somewhat similar to the one suggested by Biber (1998: 123). Biber describes the register of research articles in experimental science as "one of the few English registers that clearly distinguishes among internal purpose-shifts. […] Because each of [the IMRaD sections] is overtly marked in the text and has distinct communicative functions […]"

This table leaves out some of the verbs likely to be classified at this stage but which only occurred once throughout the corpus. I consider these verbs of no statistical relevance.

The verbs in Table 1 all have a very large semantic content, i.e. a low number of semes or semantic features. All are widely used outside the scientific register and their meaning does not vary significantly whether within or without this register. None of them can be said to belong exclusively to a specialised scientific domain. I shall refer to these verbs as category 1 verbs.

[9] The abstracts are available under HTML format from IEEE's web site at http:\\uffc.brl.uiuc.edu\Trans\ I shall henceforth refer to this collection of abstracts as ABCORP.

[10] N2 can refer to a simple noun as well as a complex noun phrase.

[11] In metaoperational linguistic grammars such as Adamczewski's the use of Vt-en to refer to the past participle of a transitive verb underlines the regularity of the system. In the system, the verb (V) can be V-ing, V-ed or V-en.

[12] The figure in brackets indicates the number of occurrence of form being referred to.

Part of the abstract	Vt-en	
Introduction	achieve (1/4)	apply (2/7)
	base (2/7)	consider (3/5)
	describe (4/7)	determine (1/4)
	discuss (3/7)	find (2/7)
	improve (1/5)	investigate (3/6)
	know (2/3)	obtain (2/16)
	present (6/20)	propose (7/13)
	study (2/5)	use (6/29)
Method and Materials	achieve (1/4)	apply (4/11)
	base (5/6)	consider (1/5)
	demonstrate (1/3)	describe (3/7)
	determine (3/4)	discuss (2/7)
	give (1/6)	identify (3/3)
	investigate (2/6)	observe (1/5)
	obtain (3/16)	perform (6)
	present (5/20)	produce (1/4)
	propose (5/13)	show (1/10)
	study (3/5)	suggest (1/2)
	use (20/29)	
Results	achieve (2/4)	apply (1/11)
	consider (1/5)	demonstrate (2/3)
	find (5/7)	give (5/6)
	investigate (1/6)	observe (2/5)
	obtain (11/16)	present (8/20)
	produce (2/4)	show (9/10)
	suggest (1/2)	use (1/29)
Discussion	discuss (2/7)	observe (2/5)
	present (1/20)	produce (1/4)
	propose (1/13)	use (2/29)

Table 1[13]

The second category of verbs extracted from the passive sentences of the corpus gathers verbs with a more restricted semantic content. Those verbs describe processes central to research in science or technology. Table 2 below shows they are used more frequently in a scientific context than in "general English". To compare the frequency of these verbs in ABCORP and in "general English", I have chosen the British National Corpus[14] as an element of comparison. The size of the BNC, 100.1 million words, and its diversity, some 4124 different files, make it the corpus most representative of "general English" available today. It could be argued that the BNC represents British English whereas ABCORP is composed of abstract published in an American Journal. However, apart from minor spelling differences (i.e. analyse vs. analyze) I do not think that the abstracts are overtly marked by their "Americanness" and I am convinced that the comparison made with the BNC yields reliable results.

Because I do not think they can be of any statistical relevance for this study, I have excluded from the comparison and the resulting categorisation the verbs that could have been taken into consideration at this point but which only occurred once throughout the corpus even once lemmatised (e.g. *adjust* in (20)). I have kept the verbs describing processes that I considered central to research in science or technology even though they occurred only once if they somehow participated in a specialised term of the domain of ferroelectricity, ultrasonics and frequency control (e.g. *enlarge* in (21) was kept because *asymmetrical bandwidth enlargement* belongs to the terminology of UFFC and occurs in ABCORP). I have also kept the verbs likely to be categorised here but which only appeared once in the corpus if a word of the same family appeared in at least one file other than the one the passive had been isolated in (e.g. *tune* in (22) because *cavity-tuned hydrogen masers* appears in a file different from the one (22) appears in).

[13] The first figure in brackets refers to the number of occurrences of the verb in the part of the abstract, the second figure refers to the total number of occurrences of the verb in the corpus.

[14] Data cited herein has been extracted from the British National Corpus Online service, managed by Oxford University Computing Services on behalf of the BNC Consortium. All rights in the texts cited are reserved.

(20) The time interval between two consecutive samples can be continuously *adjusted* to avoid undesirable sample volumes.

(21) At higher pressure levels, it even may happen that, in a steady, unidirectional flow (which should generate only positive Doppler frequencies), the Doppler spectrum is *enlarged* up to the point that negative Doppler shifts also are produced.

(22) As an example of this new class of voltage-tunable chip SAW devices, a voltage-controlled oscillator (VCO) is presented in which the output frequency can be *tuned* by an applied gate voltage.

For the purpose of this comparison, I have used one of the frequency lists provided by Kilgariff[15] (RF2). The list is a lemmatised frequency list. It features the words occurring more than 800 times in the BNC and gives their part of speech. I have lemmatised the verbs describing processes which can be considered as scientific or technological and which appear in passive sentences in ABCORP and calculated their raw frequency (RF1). I have then calculated a mean frequency (MF1) over 1,000 words for each of them. In a similar way, I have calculated a mean frequency (MF2) for the same verbs in the BNC. I have subtracted each verb's MF1 to its MF2. A positive result means that on average the verb occurs more frequently in ABCORP than in the BNC; a negative result means that on average the verb appears more frequently in the BNC, or in "general English" than in ABCORP.

Table 2 shows the ten most and least frequent verbs from this second category. The top ten undoubtedly describe a process central to research in science or technology. I have decided to keep the verbs with an MF1-MF2 close to zero or negative arguing that those verbs have a rather large semantic content which realises somewhat differently in "general English" and in ABCORP. The category gathers 78 verbs which I shall refer to as category 2 verbs.

Verb infinitive	**ABCORP**		**BNC**		
	RF	MF1	RF2	MF2	MF1-MF2
measure	16	0,8676	6683	0,0668	0,8009
compare	16	0,8676	12591	0,1258	0,7418
detect	13	0,7050	3231	0,0323	0,6727
calculate	12	0,6507	3922	0,0392	0,6115
provide	19	1,0303	47923	0,4788	0,5516
analyze	9	0,4880	4106	0,0410	0,4470
simulate	8	0,4338	<=800[16]	0,0080	0,4258
operate	8	0,4338	<=800	0,0080	0,4258
design	10	0,5423	11810	0,1180	0,4243
develop	12	0,6507	24205	0,2418	0,4089
…					
realize	2	0,1085	9726	0,0972	0,0113
fix	1	0,0542	5282	0,0528	0,0015
drive	3	0,1627	16477	0,1646	-0,0019
train	2	0,1085	11907	0,1190	-0,0105
prefer	1	0,0542	6854	0,0685	-0,0142
point	2	0,1085	12844	0,1283	-0,0199
receive	4	0,2169	24111	0,2409	-0,0240
expect	1	0,0542	27221	0,2719	-0,2177
carry	1	0,0542	31258	0,3123	-0,2580
set	1	0,0542	40381	0,4034	-0,3492

Table 2

Table 3 maps these verbs over the IMRaD part of the abstracts they appear in.

Part of the abstract	Vt-en		
Introduction	align (1/2)	analyze (1/4)	compare (1/7)

[15] Kilgariff's frequency lists are available from *http://www.itri.bton.ac.uk/~Adam.Kilgarriff/bnc-readme.html* or via ftp from *ftp.itri.bton.ac.uk/bnc*.

[16] Kilgariff does not list lemmatised forms with a frequency smaller than 800. For the purpose of my calculations, I have consider the RF2 of verbs not listed by Kilgariff as 800.

	construct (1/2)	design (1/2)	detect (1/5)
	develop (5/8)	establish (1/2)	expect (1/1)
	fabricate (1/4)	fix (1/1)	monitor (1/1)
	operate (1/2)	relate (1/3)	
Method and Materials	align (1/2)	analyze (3/4)	assess (1/3)
	average (1/2)	calculate (6/6)	calibrate (2/2)
	choose (1/1)	compare (5/7)	compute (3/5)
	control (1/1)	convert (1/1)	correct (1/1)
	decrease (1/1)	degrade (1/1)	derive (3/3)
	detect (2/5)	develop (3/8)	dominate (1/1)
	drive (1/1)	establish (1/2)	evaluate (1/2)
	fabricate (3/4)	formulate (1/1)	insert (2/2)
	introduce (1/1)	locate (1/1)	manufacture (1/1)
	map (1/1)	measure (2/2)	model (1/1)
	modify (2/2)	observe (1/5)	operate (1/2)
	position (1/1)	predict (1/1)	propagate (1/1)
	provide (1/1)	realize (1/1)	record (1/1)
	scan (3/3)	select (1/1)	separate (2/2)
	set (1/1)	substitute (2/2)	test (1/1)
	train (1/1)	transform (3/3)	utilize (1/1)
	validate (1/1)		
Results	affect (1/1)	assess (2/3)	attenuate (1/1)
	attribute (1/1)	average (1/2)	carry (1/1)
	compare (1/7)	compute (2/5)	confirm (1/1)
	construct (1/2)	detect (2/5)	enlarge (1/1)
	evaluate (1/2)	examine (1/1)	illustrate (1/1)
	increase (1/1)	indicate (1/1)	model (1/1)
	observe (2/5)	point (1/1)	rate (1/1)
	receive (1/1)	reconstruct (1/1)	reduce (1/1)
	relate (2/3)	report (1/1)	tune (1/1)
	verify (1/1)		
Discussion	characterize (2/2)	design (1/2)	implement (1/1)
	observe (2/5)	prefer (1/1)	quantify (1/1)
	reconstruct (1/1)	require (1/1)	simulate (1/1)
	solve (2/2)		

Table 3

The third category of verbs extracted from the passive sentences of ABCORP gathers verbs that would only occur in a scientific register[17] and quite specifically in the domain of electrical engineering. This category is made up of only two verbs: *pole* (2) and *repole* (1). The two verbs appear in the same abstract; *pole* occurs twice in the Discussion part whereas *repole* occurs in the Method and Materials part.

As stated earlier in this paper, I have decided to study the passive as a sentence-level operation. I therefore have not calculated a mean frequency count[18] for passives by section as proposed by Biber (1998: 124-125) since his count is word-based. The example given by Biber, a study conducted over 19 research articles from either the *New England Journal of Medicine* or the *Scottish Medical journal* published in 1985, points out to Methodology sections being "marked by their extremely frequent use of agentless passives". Unlike Biber's, this study is based on a corpus of research article abstracts and not full-length articles and so far, it has made no difference between agentless passives and passives with an agent. However, the three categories that I have defined hint at similar results in terms of the frequency of passives over the IMRaD sections of abstracts. Table 4 below shows the breakdown of passive sentences by IMRaD section.

[17] *Pole* as a verb only occurs three times in the BNC with a meaning similar to the one it has in the following example. "This is how we pole a raft and just because a white man is watching through his funny machine we aren't going to do it any differently." (Barnes J 1990 *A history of the world in 10½ Chapters* London, Picador). *Repole* does not occur in the BNC.

[18] For an explanation of the mean frequency count, see Biber's methodology box 8, the unit of analysis in corpus-based studies (1998: 269).

Part of the abstract	First category	Second category	Third category	Total
Introduction	26% (53)	13% (18)		20% (71)
Method and Materials	42% (84)	55% (80)	33% (1)	47% (165)
Results	28% (56)	23% (33)		26% (89)
Discussion	4% (9)	9% (13)	66% (2)	7% (24)
	(202)	(144)	(3)	(349)[19]

Table 4

Biber's brief study is given as an example for the study of discourse characteristics and it concludes that agentless passives contribute to "presenting events impersonally, with no acknowledged agent." The following aims at explaining that this effect, although sensed by most readers, is probably a secondary effect for the targeted discourse community.

5. Categorisation of subject noun-groups in passive sentences

Building on the above categorisation, a study of the verbs' grammatical subjects (N2's) is particularly revealing. I shall focus on noun determination.

N2's first appear as very diverse as shown by the examples below:

(23) *The* effects of the array parameters on the array performance, such as the selectivity of Lamb modes and effectiveness of Lamb wave generation are investigated.

(24) *A* new procedure for preparing lead zirconate titanate (PZT)/poly(vinylidene fluoride-trifluoroethylene) (P(VDF-TrFE)) 1-3 composites with both phases piezoelectrically active is described.

(25) *Æ* Electromechanical coupling mechanisms in piezoelectric bending actuators are discussed.

(26) These findings are shown to be in excellent agreement with previously reported theoretical predictions by the authors.

(27) *Three* specimens are used in the study: a block of Plexiglas that has a linear attenuation, a layer of a special rubber compound with an attenuation proportional to f1.38, and a phantom made of castor oil that has an attenuation proportional to f1.67.

(28) Some important factors are studied for the bimodal ultrasonic motor design.

(29) The composite disks have been fabricated into transducers with air-backing and with no front face matching layer, and *their* performance characteristics have been evaluated in water.

(30) *It* was found that the ~680 μm spot size of the experimental zone-plate did not vary appreciably with changing frequency, whereas the focal length increased markedly with increasing frequency (from ~5 mm at 450kHz up to ~15 mm at 900kHz).

(31) Tissue-characterization parameters, *which* have been used successfully by other authors, were calculated for each segment.

Examples (23) to (31) illustrate the nine different criteria I used for N_2 classification. Sentences (23), (24) and (25) are examples of the use of *the*, *a*, and *Æ*. (26) features a deictic (*these*). The N2's in (27) and (28) are both determined by quantifiers either numerical or not. In (29), the N2 is determined by a possessive adjective. (30) and (31) both have a pronoun as an N_2.

Beyond this simple labelling, a further classification is necessary. Each of the above determiners (and for the purpose of this study, I will consider pronouns as having an built-in determination) contributes to a thematic or rhematic environment. N_2 determination therefore participates in text cohesion and in the use of technical terminology.

In this part I shall study N_2 determination with relation to the abstract's IMRaD parts but also to Vt-en categorisation as shown above.

Table5 provides an overview of determination patterns in the introductions of abstracts in ABCORP.

	Introduction
∅	20% (14)
a/an	35% (25)

[19] The restrictions in the categorisation of verbs resulted in the exclusion of 33 verbs from the study.

The	38%	(27)
Others	7%	(5)
		(71)

Table 5

Zero-determination (marked ∅) is seen by Adamczewski (1993: 210) as a direct reference to the extralinguistic universe[20] with no thematisation. It occurs in 14 of the 71 passive sentences in the Introduction sentences in ABCORP. Eleven sentences feature verbs from the first category defined above (i.e. with a broad semantic content). Example (32) below is typical of zero-determination.

> (32) Quasi-monolthical integration of thin GaAs/InGaAs/AlGaAs-quantum well structures on LiNbO3 SAW devices is achieved using the epitaxial lift-off (ELO) technique.

By giving the chemical composition of the components, the speaker can hardly be any closer to the extralinguistic universe. Other N_2's in this category are technical terms which are absolutely central to UFFC such as *deformation, acoustic fields, electromechanical coupling mechanisms, piezoelectric bending actuators, piezo helical springs, curved unimorphs,* or *piezo springs* or concepts derived from such terms such as *Improvement of sensitivity in ultrasonic fields of piezocomposite transducers*.

The information is apparently not proposed as given but as new which may seem incompatible with a passive sentence in which the grammatical subject is proposed as given. I shall address this apparent conflict of interests further below.

Although only three sentences in the Introduction section of abstracts in ABCORP feature a "∅ N_2" sequence and a category 2 verb (i.e. a verb describing a scientific or technical process) it is interesting to notice that the observation made above does not seem to apply to these three sentences. Their N_2's are *small cracks, known defects, previous studies* which are indeed technical terms but not in the domain of UFFC.

In ABCORP, 25 sentences featuring a passive as part of the introduction of an abstract also feature *a/an* as a determiner for the N_2 of the sentence. Using *a/an* enables the speaker to refer to a specific occurrence of a notion therefore limiting its scope. *A/an* also conveys minimal classification. Adamczewski (1993: 213) gives the following example:

> - Let's not talk about all this. It was only *a* dream.

The speaker recognises the reference of *this* as being "an occurrence of a dream" and classifies it as such. Using *a/an* is a twofold operation which first entails acknowledging the existence of a reference and then classifying it albeit minimally. The operation sets the noun group in Adamczewski's phase 1[21] i.e. a rhematic phase. The classifying aspect of the operation is of course particularly interesting for anyone concerned with the technical terminology of a given area of specialisation. *A/an* therefore introduces the speaker's conception of the noun or noun group it determines. He refers to the prototypical occurrence of a notion to acknowledge the existence of his own occurrence of the notion. This appears quite clearly in this subsection of ABCORP in which six of the 25 N_2's feature the adjectives *new or novel* as does example (33) below.

> (33) A new numerical model of a short-term stability measuring system of quartz crystal resonators is presented.

However, the main characteristic of the N_2's extracted from this subsection of ABCORP is their extreme categorisation as shown in the following three examples

> (34) A resonant liquid capillary wave theory which extends Taylor's dispersion relation to include the sheltering effect of liquid surface inclination caused by air flow is presented.
>
> (35) A condition monitoring nondestructive evaluation (NDE) system, combining the generation of ultrasonic Lamb waves in thin composite plates and their subsequent detection using an embedded optical fiber system is described.
>
> (36) In this paper a novel ultrasound tomography imaging system is presented.

[20] The absence of any visible or audible morpheme for this determination is symbolic of the lack of intervention of the speaker on the extralinguistic universe.

[21] Adamczewski has defined a system of 2 phases for sentences. Phase 1 is rhematic or non-presupposing whereas phase 2 is considered thematic or presupposing. Typically, a sentence in the simple present or simple past will be in phase 1 whereas one featuring the be + V-ing aspect will be in phase 2.

(34) and (35) show several levels of embedded clauses and (36) exemplifies heavy noun premodification. All three N_2's thereby reach a level of classification which seems hardly compatible with the rhematic operation described above. I shall try to reconcile these two aspects further below.

Eight sentences from this subset feature verbs from the second category however, only four different verbs are used (*develop*, *fabricate*, *construct*, and *design*). Their N_2's seem to have characteristics similar to those of category 1 verbs.

Having said that the grammatical subject of a passive sentence is chosen as the starting point or theme of the passive sentence and is therefore considered as given information as opposed to new information, it should come as no surprise that a majority of the N_2's in the passive sentences in ABCORP are determined by *the*. Adamczewski's views on determination using *the* are particularly helpful. "*The* indicates that the noun group is in phase 2. The relation is presented as thematic by the utterer." (1993: 215 [translation]) He gives the following example:

- Mother, did I ever tell you? I am lucky !
- " No, you never did ", said the mother.

The first occurrence of "mother" is a direct reference to the person the sentence is directed to. However, the second occurrence , "does not refer to the partner in the conversation but to the mother already mentioned." This shows the discourse has moved a step further away from the extralinguistic universe. It underlines the role of the speaker who filters the extralinguistic universe. Adamczewski refers to "distanciation from phase 1". He further defines the notion when he deals with the ± generic effect of the *the* operator (1993: 218 [translation]):"*The* signals that the utterer somehow works on the notion's semic programme. We no longer have a truthful image of reality but a filtered one; there is a discrepancy with the extralinguistic universe. The relative alteration of the notion's semantics is modulated by the context and the situation." These few lines are of an even greater relevance when studying scientific discourse. The specific operation of determining with *the*, and therefore of filtering the extralinguistic universe, is in total agreement with passivisation as described earlier. The speaker distanciates himself from the extralinguistic universe, he handles discourse objects to distribute information. Adamczewski (1993: 218 [translation]) offers the following summary: "*The* is related to phase 2 and enables the utterer to play on notion extension depending on his intentions and the requirements of the context."

Just under 40% of N_2's in the Introduction section of abstracts in ABCORP feature *the* as their determiner. This represents 27 sentences out of 71 of which 24 feature category 1 verbs. Here are two examples.

> (37) The time-frequency distribution (TFD) of Doppler blood flow signals is usually obtained using the spectrogram, which requires signal stationarity and is known to produce large estimation variance.

> (38) The optically pumped cesium beam clock named Cs~IV is operated with a new short Ramsey cavity satisfying strict requirements on the microwave leakage level.

The starting point of those sentences has been selected and is presented as such by the speaker. Through the use of the passive, he has chosen as N_2 a noun group which he reckons is understood by the discourse community he is targeting. He has made of the understanding of the N_2 a defining characteristic of his discourse community. In (37), the speaker's intervention is further underlined by the use of the abbreviated form in brackets stressing his endorsing of the terminology. In (38), the same intervention is made even more explicit through the embedded "*named Cs~IV*" thereby fixing *Cs~IV* as a term in a highly definitory context. In the 27 sentences, abbreviated forms appear seven times.

In ABCORP, this set-up highlights such terms as *longitudinal leaky surface waves* (*LLSW*), *cross-correlation method* (*CCM*), *coherent population trapping* (*CPT*), *Doppler angle*, *direct digital frequency synthesizers* (*DDFS*), *digital-to-analog (DAC) output*, *surface acoustic waves* (*SAW*), *piezoelectric leaky surface waves*, among others.

The operation is further emphasised by the category the verb belongs to. By having a very broad semantic content, the Vt-en actually reveals very little of his N_2.

In ABCORP *The*, *a/an* and *Æ* account for more than 90% of N_2 determination in passive sentences from Introduction sections. The remaining ten percent (five sentences) are occurrences of deictics, quantifiers, relative pronouns or possessive adjectives. They do not seem to have any statistical relevance at this point of the study.

Although I have clearly stated above that the passive contributes to promote the grammatical subject of the sentence to the position of theme, of the three operators studied so far, *the* is the only one to apparently follow suit.

In fact, *A/an* and ∅ seem to do quite the opposite — that is put forward new information. Still they are heavily used throughout the introduction of abstracts in ABCORP. Example (39) may offer clues as to the actual value of this combination.

> (39) A new numerical model of a short-term stability measuring system of quartz crystal resonators is presented.

This is the first sentence of an abstract entitled "Modeling of a short-term stability measuring system of quartz crystal resonators." In itself, this sentence does not bring any more information than already present in the title of the abstract. Its purely informative content is null. I believe the role of such a sentence is therefore not linguistic per se but metalinguistic. What the passive promotes is the operation carried out on the N_2 and not the N_2 itself. In (39) the author of the abstract acknowledges the existence of an already extremely complex notion, that of *numerical model of a short-term stability measuring system of quartz crystal resonators* and of an occurrence of this notion. His acknowledgement of the existence of such a notion is a trace of the categorisation operated by the speaker i.e. one of the building blocks of the experience he is trying to recreate through his text. He is setting the conceptual boundaries of his discourse.

This metalinguistic feature can be implemented by other means than the passive. Examples (40) and (41) are also the first sentences of abstracts. They define the conceptual boundaries of the speaker's experience more explicitly.

> (40) Recent research has shown that, for a rotating phantom, the speckle pattern may not replicate the phantom motion, rather it may show a large lateral translation component in addition to rotation.

> (41) Recent papers have shown that focused ultrasound therapy may be feasible in brain through an intact human skull by using phased arrays to correct the phase distortion induced by the skull bone.

I am assuming that the same operation occurs when the N_2 of a passive sentence in the Introduction section of an abstract is determined by *Æ* or *the*. What is promoted is the determination of the N_2 not the N_2 itself. Because of the very nature of *Æ*, the operation leads to a definition of wider conceptual boundaries hence, in ABCORP, the occurrence of such central terms to UFFC as *deformation*, *acoustic fields*, *Electromechanical coupling mechanisms* etc. (see list above). Inversely, the conceptual boundaries set through the use of *the* are only central to the research article itself. By being promoted N_2's they endorse the role of key concept hence the high density of technical terms occurring as N_2 of a Vt-en in the introduction section of abstracts in ABCORP.

However, the above description of determination is very close to what it would be in a context other than passive. I have emphasised that the passive is a well-thought reorganisation of an initial pattern for a purpose. It clearly underlines the intent of the speaker that is to relate the scientific process he has been through, to recreate it for the benefit his reader. It seems that the passive contributes to bringing forward the two-way dimension of communication. It does not just enable the speaker to reorganise and, as shown above, reconceptualise his own experience of the process he wants to relate but it also calls for the reader's (co-utterrer's) acknowledgement and acceptance of the delimitation/definitions operated over the extralinguistic universe.

As already mentioned above, the Method sections of abstracts in ABCORP contain more passives than any other parts. Whereas introductions contain 71 passive sentences, Method sections contain 165. The breakdown by verb category is also quite different: 84 of the Vt-en's are from category 1, 79 from category 2 and 1 from category 3. This breakdown confirms that the "quantum of information", as it is referred to by Haliday (1994: 34), is not evenly spread over the clause. It shows that, even in a section in which one expects accounts of procedures and therefore a sizeable amount of material processes, in passive sentences, a majority of verbs, though a short one, still convey very little of those processes most of the information being conveyed by their N_2's.

A comparison of N_2 determination in this subsection of ABCORP and in the one previously studied also shows interesting results. Those results are summed up in table 6.

	Introduction	Method
∅	20% (14)	24% (40)
a/an	35% (25)	15% (25)

The	38% (27)	46% (75)
Others	7% (5)	15% (25)
	(71)	(165)

Table 6.

When considering zero-determination in passive sentences, the comparison between the Introduction section and the Method section of abstracts in ABCORP highlights a similar frequency. Of the 40 N_2's determined by *Æ* in the Method section of abstracts in ABCORP 21 are subjects of category 1 verbs and 19 of category 2 verbs. However, the N_2's determined by *Æ* are quite different from the ones undergoing similar determination in introductions. In the Method section, when the verb belongs to the first category, a majority of N_2's refers to such concepts as *computations*, *method*, *technique*, *measurements*, *simulations*, or *experiments*.

When premodified, those N_2's take on a more restrictive semantic content than in introductions as can be seen in examples (42) and (43)

> (42) *Optical interferometric measurements of pellicle displacement at discrete frequencies in tone-burst fields* are converted to acoustic pressure, and the hydrophone for calibration is substituted at the same point, allowing sensitivity in volts per pascal to be obtained directly.

> (43) Moreover, to guarantee the convergence of identification and tracking errors, *analytical methods based on a discrete-type Lyapunov function* are proposed to determine the varied learning rates of the FNNI and the optimal learning rate of the adaptive controller.

As in introductions, the other N_2's are also essential concepts to UFFC or more precisely to the subject of the study being carried out. In (44) the concepts of *location* and *extent* are central to the study describing an application to detect prostatic carcinoma.

> (44) For these patients, *location and extent of the carcinoma* were known from histological findings after radical prostatectomy.

Quite expectedly, there are fewer occurrences of determination using *a/an* in the Method sections of abstracts in ABCORP than in Introduction sections. In general, categorisation seems to have already been operated. The boundaries the author has set for his research are well in place once he gets down to explaining his methodology. This is emblematically highlighted by the one and only occurrence of the sequence *a/an new N* whereas the same sequence was concordanced in 24% of the a/an determined N_2's of passive sentences introductions.

Still, categorisation does occur in 15% of all passive sentences in the Method sections of abstracts in ABCORP. It seems however less marked than in introductions. Examples (45) and (46) are particularly relevant at this point

> (45) A line-of-sight optical projection *through a test object* is identified from an amplitude null and a sharp phase transition produced by diffusive waves originating from two in-phase (initial phase 0o) and out-of-phase (initial phase 180o) light emitting diode sources.

> (46) A second air-coupled capacitance detector (apertured to 200 µm) was scanned in the field of the zone-plate source in order to image the generated ultrasonic field at various frequencies of operation.

In (45), the author of the abstract only makes use of one prepositional group for postmodification whereas in examples (34) and (35) three whole embedded clauses are used. In (46) extra information is provided in brackets whereas it could perfectly have been embedded as a clause.
When in need of detailed categorisation, authors seem to resort to premodification. Example (47) takes this to great length.

> (47) *A 0.91Pb(Zn1/3Nb2/3)O3-00.09PbTiO3(PZN-PT) single crystal* with high electromechanical coupling factor (k33)>90% has been used to fabricate a 40-channel phased array ultrasonic probe with greater sensitivity and broader bandwidth than concentional probes.

Premodification offers a better integration of the modifier but it also blurs the relation between the modifier and the head noun making it implicit. The resulting noun group is therefore "denser" and more is demanded of the reader to understand it correctly. The same is true of abbreviation as used in example (48).

(48) Then, because the dynamic characteristics of the USM are complicated and the motor parameters are time varying, *an AFNNC* is proposed to control the rotor position of the USM.

As already hinted at in the study of *a/an* determination of N_2's in introductions, a "semantically rich" N_2 seems to call for a "semantically poor" Vt-en (category 1 verb). In keeping with the observations made when studying N_2 determination in introductions, it come as no surprise that 15 out of the 25 Vt-en's extracted in this subsection of ABCORP belong to category 1. Once again, the quantum of information is very unevenly spread over the sentence. On the one hand the speaker sets out and acknowledges the existence of highly specialised notions, on the other hand he tells us very little about them and the actual role they might play in his research. The fact that the sentence is in the passive emphasises the effect. It makes is clear for the reader that he is expected to fully apprehend the whole semantic content of the N_2 with very little co-textual help. His understanding requires appealing to his very own experience of the specialised field. By using the passive, the speaker gives the impression of setting out the N_2's he has selected as the "compulsory checkpoints" in his research. Whoever would not go through those checkpoints would not understand his research. As mentioned in the introduction of this paper when dealing with the actual purpose of the research paper abstract, it appears quite clearly at this stage that through the abstract, the speaker carefully signposts[22] the way for his reader. In this respect, the role of the abstract is very close to that of a table of contents. The function of the passive in this case seems to be a metalinguistic one pertaining to emphasising discourse[23] cohesion.

This metalinguistic function of the passive is even clearer when N_2's are determined by *the*. In this case the modifiers and the head noun are sealed into a concept and labelled as a term as in *the Lagrange multiplier method*, *the dynamic range (DR) of the new TIAOC*, *the backpropagation algorithm*, which all occur as N_2's in the Method section of abstracts in ABCORP. As could be expected, N_2's are less premodified than when determined by a/an and off course by *Æ*.

In the Method section of abstracts, the author takes his readers through the successive steps of the method he implemented. The fact that this section features the highest frequency of passive sentences but also of N_2 determination by *the* shows that a particular emphasis is put on signposting this section more than any other. This is confirmed by the very high frequency of category 1 Vt-en's. Although some of the research described in ABCORP involves extremely precise processes, category 1 Vt-en's still account for almost half of the Vt-en's in the Method section. For example, such a process as "sintering", which is central to piezoelectric ceramics production, does occur in ABCORP but never as a verb and although "poling" and "prepoling" are essential they only rarely occur as verbs.

The occurrence of N_2's as intrinsically "weak" as *the model* or *the experimental data* with category 2 Vt-en's is marginal and can hardly be considered a trend. Example (49), which features a rather weak N_2, would prove the opposite but it being the only occurrence of a category 3 verb in this subsection of ABCORP, drawing any conclusion would be hazardous to say the least.

(49) Sintered PZT rods are inserted into a prepoled copolymer matrix, and *the composite* is repoled under a lower electric field.

A closer observation of the category labelled "Others" in table 6 proves an efficient conclusion to this part in which I have tried to bring forward the role of the passive as a means for discourse cohesion. "Others" represent 15% (25 occurrences) of N_2 determination in the Method section of abstracts in ABCORP. Among these 15%, a majority of determiners are deictics (9 occurrences which represents just above 5% of the total for Method sections) all of which are occurrences of *this* or *these*. With *this* or *these* as a determiner, the N_2 is introduced as rhematic with a reference to the extralinguistic universe. Example (50) features the only case of a premodified head noun in an N_2 from this subsection of ABCORP.

(50) By using an extension of the slowly varying functions method, *this differential equation* is transformed into a nonlinear differential system with perturbation terms as the right-hand side.

The other N_2's are *this method*, *these methods*, *these parameters*, *this analysis*, *these data*, *these devices*, *this error*, *this line-of-sight.* All trigger sentences developing right of the verb group and strengthen discourse cohesion.

---

[22] "signposting" is a term used by Salager-Meyer (1990) and quoted by Gledhill (2000: 42)

[23] At this stage in the study, through the use of *discourse* rather than *text* as first announced in the title of the paper, I intend to emphasise the role of the speaker in the production of the abstract. This goes against the idea that the abstract would result from the mere selection of key points (usually on methods and results) by the author.

In the above paragraphs, I have tried to show that the operation carried out by the passive in the Method section of abstracts in ABCORP is intrinsically identical to the operation carried out by the passive in the Introduction section of abstracts in ABCORP. However, there being a major "purpose-shift" between the introduction of an abstract and its Method section and the two having distinct "communicative functions" the effect of passivisation can be quite different. In the Method sections most passives seem to highlight the key points of the research conducted. These points have been selected by the author to signpost his abstract.

It is usually made more or less explicit, in the instructions for contributors, that abstracts should state at least the method used for the study and the results achieved. Still the Result section of abstracts in ABCORP is not as overtly marked by passives as the Method section as table 7 shows.

	Introduction	Method	Results
∅	20% (14)	24% (40)	31% (28)
a/an	35% (25)	15% (25)	17% (14)
The	38% (27)	46% (75)	31% (28)
Others	7% (5)	15% (25)	21% (19)
	(71)	(165)	(89)

Table 7

Because of yet another "purpose-shift" from the Method section to the Results section, operations carried out by the speaker, although similar, take on a different interpretation.

Zero-determination occurs for 31% of N_2's in the Results section of abstracts in ABCORP, which is the highest frequency for this type of determination. However, the N_2's determined by *Æ* are quite different from the ones in the previous subsections. Indeed, this subsection of ABCORP features ten occurrences of *results*, *data* to which can be added occurrences of such N_2's as *measurements*, *comparison*, or *performance*. (51) and (52) are two examples:

> (51) Data are given showing the results of using the linear quadratic Gaussian (LQG) technique to steer remote hydrogen masers to Coordinated Universal Time (UTC) as given by the United States Naval Observatory (USNO) via two-way satellite time transfer and the Global Positioning System (GPS).

> (52) Results are presented demonstrating that the new method has both satisfactory tracking performance and the potential for practical real-time implementation.

These two examples show the cataphoric function of the N_2 and the sentence developing right of the verb group. The N_2's play a role identical to the one played by deictics in the previous part. They do take on a metalinguistic function and strengthen discourse cohesion. It is no coincidence to find nine cases of determination by deictics and four occurrences of a cataphoric *it* as an N_2, also in this subsection of ABCORP. In this perspective, the role of the N_2 and hence of the passive in (51), (52) and (53) is certainly similar.

> (53) It was found that the ~680 µm spot size of the experimental zone-plate did not vary appreciably with changing frequency, whereas the focal length increased markedly with increasing frequency (from ~5 mm at 450kHz up to ~15 mm at 900kHz).

The structure seems even more efficient with category 1 verbs, which is the case in 21 of the 28 passive sentences of this subsection.

The technical terms determined by *Æ* in the Results section of abstracts in ABCORP are usually more specific to the study than to UFFC in general.

The proportion of N_2's determined by *a/an* in this section is identical to the previous section and quite similarly the Results section does not seem to be the place to either categorise or introduce new notions.

There are 28 N_2's determined by *the* in this subsection of ABCORP ; 15 are grammatical subjects of category 1 verbs and 13 of category 2 verbs. What differs at this point of the study is the nature of the N_2's. Only 13 of them are actual specialised terms and most of the N_2's refer directly to their co-text. As examples (54) and (55) show, specialised terms can take part in the N_2 but are no longer the head nouns of the N_2.

> (54) In addition, *the effectiveness of the adaptive fuzzy-neural-network (AFNN) controlled USM drive* system is demonstrated by some experimental results.

> (55) *The resulting improvement in radial and lateral blind deconvolution* is demonstrated on six short ultrasound image sequences recorded in vitro or in-vivo.

From the Method section of abstracts in ABCORP and the Results section the "purpose-shift" does not seem to be as marked as it is between the Introduction section and the Method section. Therefore the effects of the passive in Method sections and in Results sections bear similarities and emphasis is put on text cohesion. Few specialised terms need to be introduced at this late stage of the study. They tend to be put in relation with another noun as part of the N_2 rather than head the N_2.

As Table 8 shows, only 24 passive sentences occur in the Discussion section of abstracts in ABCORP, which barely represents 7% of passive sentences in the whole corpus. Very little conclusive evidence can actually be drawn from such a small sample I shall therefore only underline that by definition, the Discussion section of an abstract is the part of the study the reader cannot anticipate on. The bulk of the informative content is provided by the author. The effects conveyed by the passive do not seem to match the communicative function of this section.

	Introduction	Method	Results	Discussion
∅	20% (14)	24% (40)	31% (28)	21% (5)
a/an	35% (25)	15% (25)	17% (14)	12% (3)
The	38% (27)	46% (75)	31% (28)	50% (12)
Others	7% (5)	15% (25)	21% (19)	17% (4)
	(71)	(165)	(89)	(24)

Table 8

6. Conclusion

As underlined in the introduction to this paper, the abstract is often considered as a "miniversion" of the research paper and should be "self-contained." The IMRaD sections of the research paper also characterise the abstract and each section participates in a rhetorical development. Throughout this study I have tried to show that far from betraying stylistic shortcomings, the use of the passive in research article abstracts participates in this rhetoric. I have first established that the passive is not a mere mechanical transformation of a hypothetical primary active sentence but a clear and well-thought reversing of an original pattern (*predicable lexis*). I have then underlined two effects of the use of the passive. The passive first enables the author to set out the terminology he intends to use. In this respect and with regards to the consistency of use of the passive, I have tried to bring forward the passive as another example of "the way phraseology helps to shape a specific view of transitivity at the same time as framing terms stereotypically" as underlined by Gledhill (2000, 167). The passive as staging technical terminology is particularly present in the Introduction section of abstracts.

In abstracts, the passive also contributes to signposting. It clearly marks the compulsory "checkpoints" the reader has to go through in order to understand the abstract. In this role, the passive strengthens discourse cohesion., which is particularly noticeable in the Method and Results sections of abstracts in ABCORP. Salager-Meyer (1990: 378[24]) has underlined the lack of cohesive devices in "unsuccessful abstracts" and the remark certainly applies to ABCORP. However, when lexical cohesive devices occur, they tend to do so in active sentences or to co-occur with "weak" N_2's. Quite logically, *Also* is the link-word occurring most frequently in passive sentences; it introduces a rather neutral link unlike more marked link-words such as *thus* or *therefore*. Table 9 shows the results of a concordance of *also* over ABCORP.

Also	In active sentences	In passive sentences	
		With a semantically	With a semantically

[24] As quoted in Gledhill (2000: 42)

		weak subject	strong subject
Number of occurrences	13	7	5
Type of N2		▪ Experimental results also are presented ▪ […] the same data also can be averaged ▪ Data also are shown […] ▪ It is also substantiated […] ▪ This improvement also can be observed […] ▪ The numerical results also are compared to […] ▪ Also, it was observed that […]	▪ A dispersive perfectly matched layer (DPML) boundary condition, which is suitable for boundary matching to such a dispersive media whole space, is also proposed […] ▪ Modern beam forming techniques such as apodization, dynamic aperture, elevational focusing, multiple transmit focusing, and dynamic receiving focusing also can be simulated. ▪ […] the free and soft planar baffle also can be considered. ▪ Good image reconstructions based on simulations and real objects also are provided […] ▪ […] negative Doppler shifts also are produced. A general overview of the theory behind the LQG technique also is given.

Table 9

In this study, I have categorised verbs occurring in passive sentences. The categories I have proposed are based on the semantic content of the verbs. They somehow differ from the categories set out by Gledhill in his study of the phraseology of cancer research papers. Gledhill (2000: 213) defines four main process types: research, empirical, clinical and biomedical. He points out that "these four dimensions form a continuum in which they represent the relative involvement of the author in scientific activity (either in experiment or writing up)." I have tried to show that the passive is one of the clues of this involvement, my category 1 largely overlapping Gledhill's research and empirical ones.

The categorisation of N_2's I have proposed, based on their determination, has underlined the lexical cohesion of discourse through labelling (Francis' term 1994). N_2's become grammatical metaphors contributing to, in Gledhill's words (2000: 204) "the distribution of thematic roles within the clause and at the same time [being] a key mechanism in the construction of new meanings."

In keeping with Gledhill's (2000: 166) and Nwogu and Bloor's (1991[25]) findings, the metalinguistic role of the passive contributing to strengthen discourse cohesion confirms that "abstracts tend to employ simple thematic progression, linearly converting rheme to theme."

Authors are often unaware of whether they "should" use the passive or not. This study has tried to show that the passive is used with great consistency in research paper abstracts through which "grammatical collocations" (Gledhill's term) have emerged. This would favour the existence of what Gledhill terms (2000: 167) a scientific "voice".

Acknowledgements

The author wishes to thank Professor M. Clay from Lyon III University for his suggestions. This study was supported by INSA-Lyon's ESCHIL, Centre for Humanities, and Department of Electrical Engineering.

References

Adamczewski H 1993 *Grammaire Linguistique de l'Anglais* Paris: Armand Colin

[25] As quoted in Gledhill (2000: 166)

Biber D, Conrad S, Reppen R 1998 *Corpus linguistics*. Cambridge, Cambridge University Press.
Booth V 1993 *Communicating in science.* Cambridge, Cambridge University Press.
Brown G, Yule G 1983 *Discoursed Analysis*. Cambridge, Cambridge University Press.
Couper-Kuhlen E 1979 *The prepositional passive in English.* Tübingen, Niemeyer.
Day R 1989 *How to write and publish a scientific paper.* Cambridge, Cambridge University Press.
Francis G 1994 Labelling discourse: An aspect of Nominal Group Cohesion in Coulthard (ed), *Advances in Written Text Analysis*. London, Routledge, pp 83-101.
Gledhill C 2000 *Collocations in Science Writing.* Tübingen, Gunter Narr Verlag.
Granger S 1983 *The be + past participle construction in spoken English*. Amsterdam, North Holland.
Groussier ML & Rivière C 1996 *Les mots de la linguistique, lexique de linguistique énonciative.* Paris, Ophrys.
Halliday MAK 1994 *An Introduction to Functional Grammar*. London, Edward Arnold.
Joly A & O'Kelly D 1990 *Grammaire Systématique de l'Anglais* Paris: Nathan
Lapaire R. & Rotgé W. 1991 *Linguistique et Grammaire de l'Anglais.* Toulouse, Presses Universitaires du Mirail.
Lobban C & Schefter M 1992 *Successful lab reports.* Cambridge, Cambridge University Press.
Nwogu KN & Bloor T 1991 Thematic progression in professional and popular medical texts. In Ventola (ed), *Functional and systemic linguistics: approaches and uses.* Den Haag, Mouton de Gruyter, pp 369-384.
Salager-Meyer F 1990 Discoursal Flaws in medical English abstracts. *Text* 10(4): 365-384.
Sides C 1991 *How to write and present technical information.* Cambridge, Cambridge University Press.
Svartvik J 1966 *On voice in the English verb*. Den Haag, Mouton de Gruyter.
Tesnière L 1959 *Eléments de syntaxe structurale*. Paris, Klincksieck.