*Studia Anglica Posnaniensia 42, 2006*

# A MULTI-DIMENSIONAL DESCRIPTION OF SUBJECT ASSIGNMENT IN ENGLISH: A CORPUS-BASED STUDY[1]

CAROLINA RODRÍGUEZ JUÁREZ

*University of Las Palmas de Gran Canaria, Spain*

ABSTRACT

The accessibility of terms to the grammatical operation of Subject assignment seems to be constrained by properties which can predict their level of accessibility to this function and which are organised in a hierarchical fashion. The relevance of such feature hierarchies has been stressed in the theory of Functional (Discourse) Grammar, and it is within this framework that the present research has been conducted. Thus, it has been my main concern to test the validity of each of these priority hierarchies in the process of Subject assignment and to provide a descriptive analysis of the different factors determining Subject selection with regard to a particular language, namely English, by analysing a corpus sample of written English and by observing whether different levels of dominance could be determined among the relevant priority hierarchies both in active and passive constructions. On the basis of the results obtained, a new level of hierarchical organization has been suggested as regards these constructions, by presenting a hierarchy of hierarchies (the Prioritising Hierarchy) which describes the different degrees of fulfilment of these hierarchies in the accessibility of terms to Subject assignment in the English language.

## 1. Introduction

The theoretical model of Functional Grammar (henceforth FG) initially developed by Dik (1997a, b) and recently improved and turned into Functional Discourse Grammar (FDG) by Hengeveld (2004a, b, 2005: 54-72) and Hengeveld and Mackenzie (forthcoming a, b), seeks to explain the reflection of the structure of natural languages as regards their main purpose, communication. This functional approach claims that the different linguistic constructions which have

[1] I would like to thank my colleague María Jesús Pérez Quintero for her encouragement and useful feedback in the elaboration of this paper. Likewise, I am deeply indebted to José S. Gómez Soliño for his continuous interest in my work and for the time invested in the elaboration of my research.

been registered in natural languages are the result of the application of different operations to various term positions within a predication. One of these grammatical operations is the assignment of syntactic functions to a constituent of a predication. Thus, many languages, among which English is found, give the speaker the possibility of describing the same *state-of-affairs*[2] (SoA hereafter) from different viewpoints depending on the constituent within the same predication to which Subject function has been assigned (resulting in active or passive constructions).

Subject assignment in English seems to be determined by different factors which constrain the degree of accessibility of constituents to Subject function. These factors represent hierarchical, intrinsic and functional properties which are presented in the form of implicational hierarchies predicting the priority of some term positions over others in having access to Subject (e.g. Definiteness Hierarchy, Semantic Function Hierarchy, etc.).

In the light of the alleged relevance of priority hierarchies in the grammatical domain of Subject assignment, it has been my intention to test their validity in relation to the phenomenon of Subject selection on a corpus of written English, by following the pluridimensional approach to the study of Subject assignment presented in FG. Thus, semantic, functional and pragmatic parameters which describe the properties attributed to terms and explain the restrictions and priorities which condition the perspective adopted by the speaker when presenting a particular SoA have been taken into account. The analysis of the data has shown solid evidence which has made me wonder about the empirical question of whether the different priority hierarchies could be grouped together showing different levels of dominance and dependence among them as far as passive and active constructions are concerned. Consequently, I can claim that a higher level of hierarchical descriptive organization could be presented in relation to Subject selection for both passive and active constructions in the English language in the form of what I have come to call *The Prioritising Hierarchy*.

The paper is organised as follows. In Section 2, a brief account of the operation of Subject assignment and of the notion of hierarchy as conceived within a functional approach is presented. In Section 3, the data and methodology used in the elaboration of this research are described. Section 4 deals with the kinds of constraints which influence Subject assignment, namely hierarchical (section 4.1.), functional (section 4.2.) and intrinsic restrictions (section 4.3.). The latter are organised in the form of feature hierarchies which are described individually in sections 4.3.1. to 4.3.6. The last section introduces the conclusions of this

[2] Within the functional approach, a state-of-affairs is conceived as something that might obtain in some world, be it real or imaginary, and that "can be located in space and time and ... be evaluated in terms of its reality" (Hengeveld – Mackenzie (forthcoming a); Dik 1997a: 51-52, 124-126).

research with the presentation of a meta-hierarchy (The Prioritising Hierarchy) which is the result of observable interrelations and dependencies among these hierarchies, and which shows a multi-dimensional description of Subject selection in English for active and passive constructions.

## 2. Subject assignment and hierarchies: From FG to FDG

The theoretical framework of Classical Dikkean Functional Grammar has stressed the relevance of implicational universals in different grammatical operations. These implicational universals are organised into hierarchies which can be described in terms of priorities which seem to have both intralinguistic and interlinguistic validity and which have been claimed to impinge on grammatical operations such as Subject assignment (Dik 1997a: 279, 1997b: 359-361).

In FG, Subject assignment is studied as one of the components belonging to the *Theory of Perspective* as postulated by Dik (1997a: 254): "modulations of perspective effected by Subj/Obj assignment". Thus, the function Subject (and Object) is conceived as a *pointer* (Dik 1997a: 251), i.e. as a perspectival function which indicates the viewpoint adopted by a particular speaker when presenting a particular SoA: in languages with Subject, there is a possible choice between Subject assignment to a first argument (A1), i.e. to the most central term required by the semantics of a predicate, resulting in an *active* construction, or to a non-first argument (A2 or A3), resulting in what has traditionally been described as *passive* constructions.

Thus, and following the priorities established by the different hierarchies, a predication frame like the one exemplified in (1) which indicates that a predicate *hit* establishes a relation between two entities represented by two arguments which carry the semantic functions of Agent and Goal respectively, could be expressed by two different linguistic expressions (examples (2a) and (2b)) (Dik 1997a: 252):

1) Past $e_i$ : [hit [V] (i1$x_1$: man [N])$_{Ag}$ (d1$x_2$: dog [N])$_{Go}$]
2) a. A man hit the dog.
   Past ei : [hit [V] (i1x1: man [N])AgSubj (d1x2: dog [N])Go]
   b. The dog was hit by a man.
   Past ei : [hit [V] (i1x1: man [N])Ag (d1x2: dog [N])GoSubj]

In example (2a), the SoA has been described from the standpoint of the A1, to which Subject function has been assigned, and as a consequence, results in an active sentence. In example (2b), on the contrary, Subject function has been assigned to a non-first argument carrying the semantic function Goal, which means that there has been a change in the perspective adopted to describe the

predication, and the resulting linguistic expression is a passive sentence.

The new version of FG, Functional Discourse Grammar, attempts to devise a grammar which apart from being pragmatically and typologically adequate is also psychologically adequate, thus proposing a radical shift from sentence to discourse in the object of study, and describing the language production process as a top-down rather than bottom-up process in which the grammatical component, which is included in a wider theory of verbal communication where a conceptual, a contextual and output component are also envisaged, is made up of different levels of linguistic organization. Within the modular reorganization of the different levels of organization, the interpersonal, representational and structural levels work simultaneously and are organised into a hierarchical layering. In this attempt to improve Classican Dikkean Functional Grammar, Mackenzie (2000, 2004) has also proposed his own model, Functional Incremental Grammar, whose main contribution is to achieve psychological adequacy by seeing "discourse production as a dynamic process occurring in real time and the expression of the clause as a similarly real-time process" (Mackenzie 2004: 182).

Within FDG, syntactic functions are located at the structural (morphosyntactic) level, and are regarded as grammatical notions which become operative once the pragmatic (interpersonal level) and semantic (representational level) functions have been assigned. Expression rules will finally determine the term which should be assigned Subject or Object function. Thus, there has been a change from FG to FDG in the sense that syntactic functions are no longer defined as purely perspectival notions which show the viewpoint adopted by a speaker when presenting a particular SoA (Dik 1997a: 251), but are rather regarded as grammatical notions which are the result of pragmatic and semantic choices at higher levels (Hengeveld 2005: 72):

> In FG functions play an important role: semantic functions are part of FG predicate frames, pragmatic and syntactic functions are assigned to constituents. In FDG ... pragmatic functions are part of interpersonal frames, semantic functions are included in representational frames, and syntactic functions, in languages in which they are relevant, are part of the morphosyntactic clause templates. Syntactic functions are thus no longer considered to be prespectivizing in nature, as they are in FG. Rather, they are matched to pragmatic and semantic units as part of the encoding operation. The pivotal nature of syntactic functions can thus be attributed to the semantic and pragmatic factors that trigger their occurrence.

As was briefly mentioned in the introduction, the degree of accessibility of term positions to grammatical operations is conditioned and restricted by hierarchical, functional and intrinsic properties which reflect semantic, pragmatic and cognitive priorities which can be collected in linear sequences of the form "x > y > z", where the items placed at the left of the scale will me more accessi-

ble to a particular grammatical operation than the items located at the right end of the scale. These sequences are called *hierarchies* and are conceived as sequences "of properties, claimed to be of absolute or statistical validity, such that a preceding property can occur without the following properties but not the other way around" (Dik 1997a: 31). The relevance of hierarchies for the study of natural languages lies in the fact that they reflect both cognitive aspects, which are determined culturally as well as psychologically, and pragmatic aspects, which are associated with the deictic centre of the speaker, that is, with what is more familiar and closer to the speaker's pragmatic information. Thus, and according to the predictions established by these hierarchies, that information which is closer to the speaker will be placed first in the linear order of the constituents of a predication. Besides, these hierarchies give information about the frequency of use of certain grammatical constructions in natural languages, rather than information about the possibility or impossibility of using such constructions (Dik 1997a: 36).

Hierarchies, in addition, can be interpreted in two ways. On the one hand, they can be understood as implicational universals which describe priorities that are typologically relevant. Implicational universals not only typify the types of linguistic patterns which may be found across languages but also point out those aspects which differentiate them as regards the linguistic subdomain to which the hierarchy has been applied, by characterising, for instance, where the cut-off point is (i.e. the point up to which a language proceeds in the hierarchy) for a particular language. Thus, in the case of Subject assignment, implicational hierarchies characterise which constituent or constituents can possibly be assigned Subject function in natural languages. The second aspect to hierarchies is that they may be applied to the description of an individual language with regard to a particular grammatical operation, indicating the different degrees of accessibility of the constituents of a predication and showing language-internal frequency distributions.

It is within this second, more specific, descriptive interpretation that hierarchies have been studied in this paper, by analysing a group of priority hierarchies which have been proposed as relevant to Subject selection in English: the Definiteness Hierarchy (definite > other specific > non-specific), the Person Hierarchy (first person / second person > third person), the Number Hierarchy (singular number > plural number), the Animacy Hierarchy (human > other animate > inanimate force[3] > other inanimate), the Concreteness Hierarchy (concrete entities > abstract entities), and the Entity Hierarchy (first-order entities > higher-order entities) (Dik 1997a: 279). I have proposed a further hierarchy to the study of Subject assignment, the Term Hierarchy (Section 4.3.6.),

[3] Inanimate forces refer to entities such as "wind, storm, rain", etc. (Dik 1997a: 35) and are classified in the "EuroWordNet top-ontology" as concrete first-order entities (Vossen 2001).

which predicts the accessibility of term positions taking into account the internal structural complexity of the term.

## 3. Selection and analysis of the corpus: The data

The written data used for the analysis of Subject assignment in English has been gathered from the LOB corpus[4] and the tool chosen for the execution of such a corpus has been Wordsmith Tools. A previous step to the actual selection of the examples from the corpus was to decide which types of constructions were needed in order to carry out this research. Thus, the study has been limited to those constructions which allow the possibility of alternative Subject function assignment in the English language, i.e. to predications with more than one argument which allow both the presentation of the SoA from the perspective of the first argument as well as from the standpoint of a non-first argument, which means that the data is exclusively made up of transitive active (Subject selection: first argument) and passive (Subject selection: non-first argument) constructions.

Both the active and passive groups which make up the corpus are made up of main clauses as well as adverbial, relative and nominal subordinate clauses, although no specific classification has been attempted since the grammatical operation of Subject assignment seems to have no direct correlation with the status of the clause.

### 3.1. Gathering of the data

In order to gather the passive and active constructions needed for our analysis, different verbal *search-words* have been used, which show a list of *concordances* in which the relevant verb in its context (two or, if required, more lines before and after the search-word) can be found. The context includes the Subject and the implicit or explicit *by-phrase* of passive verbal forms, as well as the Subject and Object of active constructions, which are the arguments which in the underlying representation of the clause could potentially be selected as Subject. The different types of search-words which had to be used were mainly selected in terms of factors such as polarity, mood and finiteness (specified in Table 1), although the search has always been conditioned by the various possibilities given by the retrieval tool used to extract the relevant structures.

The examples initially collected included cases of ambiguity both in the gathering of active and passive constructions. In the cases of passive construc-

[4] I decided to limit the data to examples of written language with a view to later testing the results of this corpus-based study on a corpus of oral data.

tions, examples of ambiguity between participle forms and adjectives in *-ed*, which the tagging programme does not distinguish, were found. As a result, a disambiguation process had to be carried out manually following the criteria presented by Johansson *et al*. (1986: 4), with examples such as *be accustomed to* and *be acquainted with* being excluded from the data. In active constructions, I searched for all those constructions with a verb followed by a nominal group (with and without articles and determiners), a pronoun and a nominal clause which from the syntactic point of view would function as the Object of a transitive construction. The gathering of these data also included examples of ambiguous sentences such as the case of finite Object nominal clauses introduced by the conjunctions *that, if* (interrogative) or *whether*, in which the disambiguation process consisted in eliminating all the instances (693 examples) of other adverbial conjunctions which were retrieved by using the tag for conjunctions *CS* (*before, although, after, as, until, if* (conditional), *as though, because*, etc). Similarly, with regard to the examples of interrogative nominal clauses introduced by wh-determiners [*_VB* _WDT: *what, whatever, which*], 26 examples in which the word *which* functioned as a relative pronoun were eliminated.

### 3.2. Determining sample size

The first stage in the selection of the sampling presented all the examples included in the LOB corpus extracted from the different search-strings, which constitutes the so called *population*. As has already been indicated, some disambiguation processes were carried out over the whole population so that the sample would only include relevant examples of active and passive constructions. After the disambiguation processes, the global sample of passive and active examples was made up by 9,603 and 37,355 examples respectively. The following step in the selection of the relevant sample consisted in calculating the size which the sample should theoretically have by applying statistical methods which calculate the number of examples which should be considered as homogeneous and representative of the whole population (Blecua *et al*. 1999: 63). In order to obtain the sample, it was necessary to calculate the minimum number of examples which would be required so that the sampling results could then be generalized to the whole population with a margin of error of 0.05 (5%) (Neuber 1980: 48-49).

In order to obtain the representative number of examples for the passive and active populations, and due to the fact that the two populations are divided into subgroups which correspond to each of the search-strings used in the initial selection of the global population, it was necessary to calculate the margins of error for each subgroup of the sample to be considered so that we would obtain representative subsamples. The formula used to calculate the sample size of the

different subgroups is the following (García-Ferrando 1985: 142):

$$n = \frac{z^2 \cdot N \cdot p \cdot q}{N \cdot E^2 + z^2 \cdot p \cdot q}$$

[n = sample size; N = population size; $z^2$ = level of confidence [z = 2 (95.5%)]; E = margin of error (0.05); p = individual probability of the phenomenon (0.5); q = complementary probability (0.5)]

The different values which the formula takes in each of the specific calculations will determine the theoretical sample size (n) which the sample should have. Let's illustrate this with a practical example. In the gathering of passive structures with *be* in positive declarative sentences with or without modals, a global sample of 8,659 examples was found. Thus, in this particular case the formula adopts the following values, indicating that the sample size for this structure should include 383 sentences:

z = 2; E=0.05; N= 8,659; p=0.5; q=0.5

$$n = \frac{4 \cdot 8{,}659 \cdot 0.5 \cdot 0.5}{8{,}659 \cdot (0.05)^2 + 4 \cdot 0.5 \cdot 0.5} = 382.34$$

These statistical methods were applied to each of the subgroups which constitute the population of passive constructions. Further calculations, however, had to be carried out due to the fact that the examples collected in each of the subgroups are subdivided depending on the type of genre they belong to, which means that the examples which had to be selected should be representative of the different genres which make up the LOB corpus.[5] For example, for the structure "auxiliary be in positive declaratives with or without modal verbs", 8,659 examples were found, of which, and according to the statistical methods applied, 383 examples should constitute this subsample. For genre A (press: reportage), 830 examples were found, that is, 9.6% of the totality (8,659), which means that the proportional number of examples corresponding to the subsam-

[5] The text categories included in the LOB corpus are classified by assigning a capital letter to each genre: A: press: reportage; B: press: editorial; C: press: reviews; D: religion; E: skills, trades, and hobbies; F: popular lore; G: belles lettres, biography, essays; H: miscellaneous: government documents, foundation reports, industry reports, college catalogue, industry house organ; J: learned and scientific writings; K: general fiction; L: mystery and detective fiction; M: science fiction; N: adventure and western fiction; P: romance and love story; R: humour.

ple (383 examples) should be 37 examples (the first 37 examples shown by the programme were selected). For genre B (press: editorial), 468 examples were gathered in the subgroup of 8,659 examples, which means that 21 would be the number of examples which should proportionally correspond to the representative subsample of 383 examples. The same procedure was followed in each of the subgroups of passive constructions and in each of the genres as can be seen in Appendix 1. The global number of passive sentences which according to the statistical methods should be analysed in order to work with a representative sample was finally made up of 797 examples.

The sampling process used for active constructions indicated that the total number of the sample to be analysed should include 1,516 examples of active constructions. Due to the large number of examples found (31,217), it was impossible to determine the exact number of cases included in each of the genres, as was done with the passive examples, and the statistical calculations were applied to each of the three subgroups making up the sample (verb + noun / verb + pronoun / verb + determiner + noun), and within each subgroup to each of the search-strings which had to be used in the selection process. Thus, for the construction "verb + pronoun", 7,010 examples were collected, from which 379 examples constituted a sufficient and representative number of the global sampling. In order to select the 379 examples, it was necessary to obtain the number of examples found from each of the searches conducted (possessive pronoun, personal pronoun, etc.), and by using these figures it was possible to calculate the proportional number of examples which should be selected for each of the searches. For example, of the 1,992 examples found with the structure "verb + third personal singular pronoun in accusative form (him/her)", only 107 examples were necessary in our analysis. Appendix 2 breaks down these results.

Thus, the global sample of active and passive constructions is made up of 2,313 examples, of which 797 tokens are passive examples and 1,516 active ones. Although at first sight it might seem that the two subsamples are unbalanced, the global figures obtained are the result of applying statistical calculations and correspond to the minimum number of examples of passive and active constructions with a margin of error of 0.05% which should be analysed for the results being representative of the global population. Besides, it should be remembered that the importance of determining sample size in corpus selection derives form the fact that samples that are too large may waste time, resources and even money, whereas samples that, on the contrary, are too small may lead to inaccurate results (http://www.isixsigma.com/ library/ content/ c000709.asp). Thus, Table 1 specifies the size of the sample and shows the individual figures which correspond to each of the search-words used in order to gather the relevant examples:

Table 1. Size and organization of the sample

| PASSIVE CONSTRUCTIONS | | EXAMPLES |
|---|---|---|
| PASSIVE WITH *BE*: | | |
| positive declaratives | | 383 |
| negative declaratives | | 104 |
| *be* + adverb + participle | | 262 |
| *be* + pronoun+ participle | | 11 |
| *be* + (det) noun+ participle | | 8 |
| PASSIVE WITH *GET*: | | |
| *get*s + participle | | 3 |
| other positive and negative declarative/interrogative present forms and interrogative/negative past forms (*get* + participle) | | 21 |
| positive declarative past forms (*got*+ participle) | | 3 |
| participle (*getting* + participle) | | 2 |
| TOTAL | | 797 |
| ACTIVE CONSTRUCTIONS | | EXAMPLES |
| verb + noun phrase | Verb + noun | 370 |
| | Verb + pronoun | 379 |
| | Verb + determiner + noun | 392 |
| | Verb + *that* , *if* , *whether* | 182 |
| verb + noun clause | Verb + *wh*-word | 48 |
| | Verb + *to*-infinitive | 128 |
| | Verb + *-ing* | 17 |
| TOTAL | 1,516 | |
| TOTAL SAMPLE | 2,313 | |

3.3. Analysis of the sample

As regards the analysis of the sample in terms of the priority constraints associated with Subject assignment, both the Subjects and by-phrases in passive constructions, and the Subjects and Objects in active sentences had to be studied and conveniently analysed with respect to the Semantic Function Hierarchy and the other priority hierarchies presented by Dik[6] so as to be able to observe the dominance of some priorities over others with respect to the accessibility of such terms to Subject.

The criteria which have been followed in judging whether a hierarchy is respected or not are the following. In the first place, the properties of the relevant Subject and Object/by-phrase of a particular example were noted down with reference to each of the hierarchies being studied in this research and were represented by means of the symbol >, which indicates that the property preceding it has won out the property following it. Thus, for instance, the notations h > in (Animacy hierarchy) and $1^{st}$ > $3^{rd}$ (Entity Hierarchy) would indicate that the Subject presents the characteristic of being a first-order human entity, whereas the Object (of an active construction) or by-phrase (of a passive construction) is an inanimate third-order entity. In this particular example, the two hierarchies have been fulfilled since they respectively predict that human entities are more accessible than inanimate terms and that first-order entities are the most accessible of all the other order entities.[7]

On the contrary, a hierarchy is judged as not being respected when the Subject presents some of the properties being located at the right of the scales and the Object or by-phrase has a property placed towards the left extreme, thus violating the prediction established by the priority hierarchies. For example, the notation an > h shows that the Animacy Hierarchy (human (h) > other animate (an) > inanimate force (f) > other inanimate (in)) has not being fulfilled since a term with the property non-human animate (an) has been more accessible to Subject than another constituent in the same predication which is human (h).

---

[6] Although in this research I have mainly followed Dik in the discussion of the different types of hierarchies, other works on the notions of hierarchies should also be pointed out. For example, for one of the earliest discussions of hierarchies including frequencies in English, see Greenberg (1966); for a discussion of hierarchies in general, but especially of the hierarchies of person and animacy (including Subject accessibility hierarchies such as Subject > Object), cf. Artstein (1998).

[7] In the analysis of the sample in terms of the different priority hierarchies, the examples in which both terms, i.e. the Subject and Object in active constructions and the Subject and by-phrase in passive sentences, present the same property (for example, both terms being definite, or plural, etc.) have been excluded from the analysis since they do not show any preference of a particular feature over another, and as a result are irrelevant for the study of the priorities predicted by the feature hierarchies.

Finally, those examples in which the two terms to be analysed present two features (b > c) which do not correspond to the first feature (a > b > c) in the hierarchy, have been registered as fulfilled instances of the hierarchy if the two features mirror the ordering predicted by the hierarchy. Thus for instance, examples were found in which a non-human animate entity (an) was assigned Subject function over an inanimate one (in), and this instance was classified as an example of fulfilled hierarchy because although there was no human term involved, the sequence established by the hierarchy as more frequent and less marked is still respected.

The results of the descriptive analysis of the data were noted down in tables created in the Microsoft Excel programme which, by means of specific formulae and filtering devices, yielded the relevant quantitative results.

## 4. Constraints on the accessibility of term positions to Subject assignment

The various types of linguistic constructions which can be found in natural languages are the result of different operations which can be applied to the different term positions of a predication. In the case concerning us, the grammatical operation of Subject assignment determines that some terms are more accessible than others to occupy the Subject position. The concept of *accessibility* which Dik borrows from Keenan (1976, 1987) and Keenan and Comrie (1977) is thus central in the study of Subject assignment and is conceived as "the capacity of a term position to be the target of some grammatical operation. A term position T to which an operation O can be applied is accessible to O; otherwise it is inaccessible to O" (Dik 1997b: 357).

The constraints which govern term accessibility to different grammatical processes seem to be related to cognitive aspects associated with the degree of closeness of the constituents with respect to the deictic centre of the speaker and which imply that such properties appear in the first positions in the implicational hierarchies: "there are connections between grammatical and cognitive accessibility in the sense that, to a certain extent, those constituents which are most accessible to grammatical processes are at the same time most accessible in a cognitive sense" (Dik 1997a: 41).

In the case of the Subject assignment operation, the degree of accessibility of a term position is conditioned by hierarchical, intrinsic and functional constraints which, although they will now be presented independently, may interact.

### 4.1. Hierarchical constraints in Subject assignment

Hierarchical constraints on the accessibility to Subject function involve the po-

sition of the term which could be eligible for Subject assignment in the overall organization of the hierarchical structure of the clause (Dik 1997b: 363). In Classical Dikkean Functional Grammar, it is claimed that the levels where syntactic functions are assigned is the level of the core predication (Dik 1989: 232, 235) since it is arguments and level-1 satellites ($\sigma_1$), which specify, define and identify the quality of the SoA designated by the predication, that can generally be assigned syntactic functions (Dik 1997a: 64, 275, 278).[8] In FDG there has been a reorganisation of the different levels of representation which constitute the grammatical component, and it is at the structural (or morphosyntactic) level that syntactic functions are operative.

Thus, the hierarchical restrictions which condition the accessibility of terms may be represented by means of formulae such as O [$_x$... (T) ...] which should be read as follows: "the operation (O) may only be applied to terms (T) which form part of a constituent of type X" (Dik 1997b: 363). In the case at hand, the hierarchical constraint to Subject assignment for English could be represented as SubjAssig [$_{e/structural}$ ... (T) ...] which indicates that the constituent which is accessible to Subject belongs to the SoA (symbolised by means of the variable (e)) and is located at the structural level.

### 4.2. Functional constraints in Subject assignment

Semantic, pragmatic and syntactic functional constraints directly condition the accessibility of term positions to partake in grammatical operations in natural languages (O $(T)_F$): "the operation (O) may only be applied to terms (T) with the function F" (Dik 1997b: 365). In the case of Subject assignment, it is semantic functional constraints that clearly restrict the accessibility of term positions to be assigned such a function. These constraints are presented in the Semantic Function Hierarchy (SFH)[9] which predicts that Subject assignment may only be

---

[8] In a limited number of languages such as the Bantu and the Philippine languages, the assignment of syntactic functions takes place at the level of the extended predication after the insertion of level-2 operators and satellites carrying the semantic functions of Temporality or Location. These cases of extraordinary Subject assignment are restricted by specific grammatical requirements which fall outside the scope of the assignment of syntactic functions. Thus, Philippine languages such as Cebuano and Kalagan (Dik 1997a: 272) illustrate this fact by allowing Subject assignment to a $\sigma_2$ (Temporality) which is grammatically restricted to obtain in relative constructions (in these languages the relative pronoun can only function as Subject). Examples of Subject assignment to level-2 satellites with the semantic function Location have been found in many Bantu and Philippine languages where once again this assignment is restricted to grammatical operations such as relativisation, question formation or clause linkage among other conditions (Dik 1997a: 272; Siewierska 1991: 105).

[9] In FDG, the Semantic Function Hierarchy has been reformulated and expanded with the missing category of Experiencers: Agent > Patient > Experiencer > Recipient > Beneficiary > Instrument (Hengeveld – Heesakkers 2004). However, in the present research, the data were ana-

applied to terms with one of the semantic functions included in the hierarchy, establishing the priority of some semantic functions over others: SubjAssig $(T)_{SFH}$. These semantic functions are assigned to three types of arguments, A1 {Agent, Positioner, Force, Processed, Zero}, A2 (Goal, Recipient) and A3 (Recipient), and to a limited number of level-1 (Beneficiary, Instrumental) and level-2 satellites (Location, Temporality):

| *A* > | *Go* > | *Rec* > | *Ben* > | *Instr* > | *Loc* > | *Temp* |
|---|---|---|---|---|---|---|
| A1 | A2 | A2/3 | | $\sigma_1$ | | $\sigma_2$ |

In order to analyse the A2/3 (Subject) and the internal-A1 (by-phrase)[10] of passive structures and the A1 (Subject) and A2/3 (Object) of active constructions respectively according to the SFH, it was necessary to draw a distinction between bivalent and trivalent predicates so that the different argumental positions could be analysed in terms of their semantic function (Table 2):

Table 2. Distribution of bivalent and trivalent predicates in the passive data

| BIVALENT PREDICATES 775 (97.2%) | TRIVALENT PREDICATES 22 (2.8%) |
|---|---|
| Subj Goal 9 (1.1%) | Subj Recipient 13 (1.7%) |

As regards trivalent predicates, it could rightly be objected that on the basis of such a small sample (only 22 examples) it is difficult to draw any kind of conclusions as regards the competition for Subject assignment between Goals and Recipients. However, I still consider that the results obtained from the analysis of the trivalent constructions may be revealing and a reflection of what really happens in most of the cases. Thus, against the prediction established by the SFH, it is noted that terms with the semantic function Recipient are more frequently assigned Subject function than those with the semantic function Goal,

lysed in terms of the original SFH.

[10] I have coined a new term for the concept *agent complement* since I consider that the traditional expression may lead to false conclusions as regards the semantic function assigned to this term: agent. In fact, in the analysis of the passive data I found different examples of *by-phrases* (mapped onto first arguments) with semantic functions other than Agent grouped under the A1: Positioner, Force, Processed [Exp], Zero [Exp]. Thus, I consider that on the one hand the term *internal first argument* (*internal-A1*) does not exclusively limit the different semantic possibilities to the function Agent but expands them to the whole group of semantic functions which constitute the first argument (A1), and on the other hand reflects the fact that this term is not explicitly encoded in the sentence in most of the cases, hence the expression *internal*.

which means that the validity of the SFH as regards trivalent predicates should be reconsidered and perhaps even reformulated. In fact, different examples may be found in the literature which reflect the priority of Recipient over Goals. For instance, Ruiz-Yamuza (1996: 203; 1999: 347) presents the following semantic ordering for classic Greek: Agent >> Recipient/Beneficiary >> Patient >> Others. Similarly Givón (1984: 139, vol. 1) describes his "topic (accession) hierarchy" as "Agent > Dative/Benefactive > Patient > Locative > Instrument > Associative > Manner adverbs" where the Recipient (Dative) and even the Beneficiary (Benefactive) precede the Goal (Patient). Wolvengrey, likewise, states that in Algonquian languages the animate dative (recipient, benefactive) consistently outranks the non-necessarily-animate theme, and thus the Semantic Function Hierarchy for such languages establishes that "Agent → Recipient/Benefactive → Theme" (2005: 426). Sentences (3) and (4) taken from our data illustrate examples of Subject assignment to Recipient (over Goals) in trivalent predicates:

3) He (Rec) was also given strips of papers, sticks, rulers, etc (Goal) … (556/E11-36)[11]
4) All those denying birth control practice (Rec) were shown a numbered list of non-appliance methods (Goal) … (272/J31-110).

As for the absence or presence of the internal-A1 in passive sentences, only 14.3% of the examples explicitly manifest it. In the rest of the cases (85.7%) it is only implicitly present but recoverable from the linguistic and extra linguistic context in most of the cases. The intrinsic properties of the internal-A1 have been analysed with respect to the different priority hierarchies, and show that a typical internal-A1 is always a simple term, especially non-specific indefinite (75.5%), plural (84.7%) and third person (91.2%), which in most of the cases refers to a concrete first-order entity (96.0%) and is preferably human (91.7%). When the internal-A1 presents a different animacy feature (non-human animate or inanimate) (8.3%), this term will be explicitly manifested in the construction in most of the cases in part due to the fact that they are less frequent and as a consequence less expected.

The semantic function which is more often mapped onto the A1 is the Agent both in active and passive constructions. Table 3 breaks down the specific percentages for all the relevant semantic functions and shows that the types of

[11] The references which come at the end of each example indicate (i) the number that I have assigned to that example in the total corpus (556), (ii) the type of text (genre) from which the example was taken represented by means of a capital letter (E: skills, trades, and hobbies), and (iii) the number and line assigned to the text in the LOB corpus, in this example, 11 and 36 respectively.

SoAs which admit alternative assignment of the Subject function are those which describe Actions, Positions and Processes (no examples of States (Subject-Zero[Exp]) are found) in this order. The linear order of the semantic functions which belong to the A1 group reflects the frequency of use of these types of SoAs, which does not mean that these semantic functions reproduce a hierarchical ordering. Siewierska (1991: 112) has the same intuition: "We can thus observe a weakening of the $A^1$ semantic functions for subject … as we move from the Agent to Positioner, then Force, Processed and Zero."

Table 3. Rate of appearance of the semantic functions of the A1

| A1 → | Agent | | Positioner | | Force | Processed |
|---|---|---|---|---|---|---|
| | Nº | % | Nº | % | Nº % | Nº % |
| Passives | 628 | 78.8% | 81 | 10.1% | 50 6.3% | 38 4.8% |
| Actives | 980 | 64.6% | 285 | 18.8% | 109 7.2% | 142 9.4% |
| SoA→ | Actions 69.5% (1608) | | Positions 15.8% (366) | | Processes 14.7% (339) | |

Finally, it should be highlighted that no examples of Subject assignment to Beneficiary were found in the data, which in a sense reflects the degree of linguistic insecurity as to the acceptability of such Beneficiary Subjects in English (Dik 1997a: 268).

In conclusion, the analysis of the data shows that as far as Subject assignment is concerned the prediction claimed by the SFH is correct for the English language and this fact can be formulised as SubjAssig $(T)_{A1> Goal> Rec > (Ben)}$. However, in the special case of trivalent predicates the prediction is not respected because Recipient Subjects outrank Goal Subjects.

### 4.3. Intrinsic constraints in Subject assignment

A term's intrinsic properties may constrain its accessibility to grammatical operations and determine the degree of priority of some terms over others in those operations. This type of restriction is represented in the formula O (T: < P >): "the operation (O) may only be applied to terms (T) with the intrinsic properties P" (Dik 1997b: 359). The relevant intrinsic properties are represented in scales which establish priorities and which have been given the name of priority hierarchies. These hierarchies have been proposed as typologically relevant and present a diversity of properties which highlight the pluridimensional approach adopted by the functional theoretical model in the study of Subject assignment.

According to Dik, the following hierarchies present relevant intrinsic factors which restrict the operation of Subject assignment (1997a: 279): Definiteness, Number, Animacy, Concreteness, Person, and Entity. As I pointed out in section 2, I have proposed a further hierarchy to the study of Subject assignment, the Term Hierarchy, which predicts the accessibility of term positions taking into account the internal structural complexity of the term. In the following sections I present the results of the analysis of the relevant data for each of the priority hierarchies separately.

4.3.1. The Definiteness Hierarchy

The properties presented in the Definiteness Hierarchy "definite > other specific > non-specific" (Dik 1997a: 37) are not intrinsic to the referent of the term, but to the term itself: "definiteness … [is] only contingent to nouns, dependent on the properties of the terms in which they are used" (Siewierska 1991: 187) or as Dik himself manifests: "Definiteness is … a property intrinsic to the term as such" (1997b: 359).

The Subject terms which were studied both in the passive and active constructions were analysed in terms of the notions of definiteness (d) and specificity (s) which are closely related to the Definiteness Hierarchy. In this way, the means of presentation of the hierarchy can be reformulated, establishing that a definite term (whether specific or not) is more accessible to Subject assignment than an indefinite term which is specific, and this, in turn, is more accessible than an indefinite non-specific term: d (±s) > i (+s) > i (-s).

The results of the analysis of the entire sample indicate that 86.3% of the examples have a definite Subject (example (5)) as opposed to 13.7% of indefinite Subjects, of which 7.0% have specific reference and 6.7 % non-specific reference. Table 4 breaks down the results obtained for passive and active constructions and shows that the perspective adopted by the speaker when presenting an SoA is highly constrained by this hierarchy since the prediction established by the Definiteness Hierarchy is highly respected in both types of constructions: 95.6% in passive sentences and 93.5% in active sentences. Examples (5) to (8) illustrate cases in which the hierarchy has been violated.

5) When one considers that the permanent committee and all the subcommittees (d) are similarly helped by advisers (i(-s)), one gets some idea of the formidable body of international yachtsmen who descend upon London for this occasion (558/E18-84).
6) No one (i(-s) can blame Harold McMillan (d) for trying to reach the elusive goal (1123/B20-193).

7) It rustled faintly as it moved, and then a sharper breath of wind (i (+s)) caught it (d) and hurried it away (1316/P15-148).
8) Someone (i(-s)) ... had sent her (d) this damnable note... (1393/P12-172).

Table 4. Validity of the Definiteness Hierarchy

| Definiteness | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 519 | 95.6% | 677 | 93.5% |
| Violated | 24 | 4.4% | 47 | 6.5% |
| Total | 543 | | 724 | |

In conclusion, the following formula summarises the accessibility of terms to Subject as regards the SFH and the Definiteness Hierarchy in the English language:

$$\text{SubjAssig } (_{\text{e/structural}}(\text{T})_{\text{A1 > Goal > R > (Ben)}} < [\text{d } (\pm \text{ s}) > \text{i } (+ \text{ s}) > \text{i } (- \text{ s})] >)$$

4.3.2. The Person Hierarchy

The intrinsic properties encoded in the Person Hierarchy are closely related to the concepts of empathy and iconicity, since, as Kuno (1976: 433) suggests, it will be easier for the speaker to show empathy for himself/herself than for the hearer, and in turn more difficult to show empathy for a third person than for himself/herself or for the hearer. In the same line, Siewierska claims that the Person Hierarchy is a manifestation of iconicity which is closely related to the natural attention flow, which "refers to the actual development of events in the real world, the basis for the perception of naturalness being temporal order" (1991: 105-106). The order of preference formulated in the Person Hierarchy is iconic in the sense that it reflects the interest of human beings in themselves, their interlocutors and other humans respectively and in the fact that the different types of SoAs are usually presented from the point of view of the human participants involved in such events and situations rather than from the point of view of other non-human or inanimate entities also involved in them.

Table 5 shows different results for active and passive constructions as far as the Person Hierarchy is concerned: {1, 2} > 3 or Speech Act Participant > Non-Participant (Dik 1997a: 36).

Table 5. Validity of the Person Hierarchy

| Person | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 37 | 34.9% | 366 | 83.0% |
| Violated | 69 | 65.1% | 75 | 17.0% |
| Total | 106 | | 441 | |

Thus, in the case of passive examples, there is a higher rate of cases in which the preferred sequential order has been violated (65.1%), always together with the violation of some other hierarchies (Definiteness, Animacy, Concreteness, Entity and Term Hierarchies) (example (9)), as opposed to active constructions in which the hierarchy is fulfilled in 83.0% of the examples (example (10)):

9) Children of age (10 and 11 years old) (3) were chosen [by us (1)] on the assumption that parental interest and curiosity would be at their height… (268/J29-07).
10) I (1) should be setting a better example (2) (1550/A25-87).

The reason why there is a high number of passive examples in which the Person Hierarchy is not respected may be due to the fact that passive constructions allow for the presentation of the SoA from a point of view distant from the speaker especially in academic and scientific texts where there is a general tendency to avoid the presentation of events from the perspective of the first person.

4.3.3. The Animacy Hierarchy

The Animacy Hierarchy presents intrinsic constraints perceived through the referents of the terms which are associated with the following properties: human (h) > other animate (an) > inanimate force (f) > inanimate (in) (Dik 1997a: 37).

The analysis of the data has revealed that the fulfilment of the Animacy Hierarchy in the operation of Subject assignment is highly dependent on the type of argument which has had access to Subject. If the Subject function has been assigned to a first argument, the hierarchy is fulfilled in 96.5% of the cases, whereas in those examples in which a non-first argument is in Subject position the percentage is reduced to 5.6% (Table 6).

Table 6. Validity of the Animacy Hierarchy

| Animacy | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 31 | 5.6% | 1065 | 96.5% |
| Violated | 522 | 94.4.% | 39 | 3.5% |
| Total | 553 | | 1104 | |

The discrepancy between active and passive constructions as regards the Animacy Hierarchy is intrinsically related to the various semantic roles and to the type of predication involved. Thus, it can generally be claimed that the semantic function Agent, which is always associated to a human entity, is very often mapped onto the first argument (A1) both in active and passive constructions, as was shown in table 3, followed in a lower percentage by the semantic functions of Positioner, Force and Processed. The semantic function Goal, which is generally mapped onto the A2, is associated to an inanimate entity in 92.8% of the examples in which the A2 has been selected as Subject (passive constructions), as opposed to only 5.0% of human entities, 2.0% of non-human animate entities and 0.2% of inanimate forces. When the A2 has been selected as Object (active constructions), this is linked to an inanimate entity (as opposed to a human entity as Subject) in 74.0% of the examples, to a human entity in 24.9% of the examples (of which 22.4% of the sentences also have human Subjects) and to animate and inanimate forces in very low percentages (1.0% and 0.1% respectively). This kind of relationship between semantic functions and argument types and their association to the various animacy features explains why in passive constructions the Animacy Hierarchy is not respected in 94.4% of the examples, and, in contrast, active constructions comply with the hierarchy in 96.5% of the cases.

Another reason which could justify the discrepancy between active and passive constructions as regards the Animacy Hierarchy has to do with the lexico-semantic properties of the predicates involved in the predication. Thus, the low percentage of fulfilment in passive constructions may be associated with the type of SoAs which most often allows Subject assignment to a non-first argument: Events and especially Actions (see Table 3). The predicates which describe Actions require the presence of an explicit or implicit human Agent (the internal-A1 in passive constructions and the Subject in active sentences) which projects the result of the verbal action into another constituent (A2: Goal) which can also be human but that in most of the cases is an inanimate entity. This sequential ordering, however, is not followed in passive constructions since an A2

(especially an inanimate Goal) has been selected as Subject. Such instances of violation of the Animacy Hierarchy are associated to instructions-giving predicates (example (11)) which explain how to carry out actions or processes (probably in order to avoid imperative active constructions) and to scientific and academic contexts in which the use of the first person is avoided because of formality reasons and conventions (example (12)).

11) The amount of the backward tilt and also the height is left to personal choice and the width is best taken [by you (h)] directly from the work to ensure a good fit (566/E04-31).
12) This regression analysis was also carried out [by us (h)] for each school separately (661/J38-104).

Moreover, in most of the passive examples (77.4%), the human Agent (the internal-A1) is indefinite as opposed to a definite Subject, which shows that the Definiteness Hierarchy is more determinant than the Animacy Hierarchy in passive constructions. In active constructions, however, the Animacy Hierarchy is complied in most of the cases because the presentation of the SoA reflects the linear projection of a human definite A1 towards and A2 (and even an A3) which very often presents definite or indefinite inanimate features.

Nevertheless, it should be observed that in *get*-passives the Animacy Hierarchy is obeyed in 93.1% of the examples. The high rate of cases in which the examples of *get*-passives comply with the prediction established by the Animacy Hierarchy is due to the fact that this type of passive with *get* is generally restricted to human Subjects whose referents are somehow negatively affected by the action described in the verb (example (13)).

13) Len and Busk got away. They're making for the north. If they (h) get picked up, they won't grass (774/N18-78).

All in all, it can be concluded that the levels of fulfilment of the Animacy Hierarchy vary as to the type of constituent which has been selected as Subject. Thus, if an A1 is selected as Subject, high levels of fulfilment are expected: SubjAssig ($_{e/structural}$ (T=A1)$_{SFH}$ < [h > an > f > in]>). On the contrary, if a non-first argument is selected as Subject, the level of fulfilment of the Animacy Hierarchy decreases importantly: SubjAssig ($_{e/structural}$ (T=A2/3)$_{SFH}$ < [in > h > an > f]>).

4.3.4. The Number Hierarchy

Table 7 shows the specific results of the analysis of the relevant data with regard to the validity of the Number Hierarchy in Subject assignment which predicts that singular terms are more accessible to Subject than plural:

Table 7. Validity of the Number Hierarchy

| Number | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 452 | 95.0% | 172 | 44.2% |
| Violated | 24 | 5.0% | 217 | 55.8% |
| Total | 476 | | 389 | |

Although Dik includes this hierarchy among those which are relevant in the process of Subject assignment (Dik 1997a: 279), he admits having some doubts about its validity in comparison with the other priority hierarchies due to the fact that there are not enough relevant data which may confirm or deny this reality: "I have not included the category of Number in the hierarchies given in (31)-(39) [Person, Animacy, Gender, Definiteness, Semantic Functions, Syntactic Functions, Pragmatic Functions], because there is a lack of data on the role of Number in the relevant priorities" (Dik 1997a: 38, n. 7).

The Number Hierarchy also shows different results for active vs. passive constructions, being almost always fulfilled in the case of Subject assignment to a non-first argument as opposed to only 55.8% of examples of active sentences in which the hierarchy is violated. It is interesting to highlight that in the cases of violation of the Number Hierarchy in active constructions, the other relevant priority hierarchies show high levels of fulfilment and in 89.4% of the examples in which plural terms have been assigned Subject over singular terms the Number Hierarchy is the only one which has been unfulfilled, indicating that the property presented by this hierarchy is not as determinant as the features included in the rest of the priority hierarchies.

4.3.5. The Entity Hierarchy and the Concreteness Hierarchy

The Entity Hierarchy is connected with the types of entities that terms can refer to and may be represented as "first-order entities > higher-order entities" (Dik 1997a: 279). First-order entities ($x_i$) refer to physical objects, individuals and places located in space, and are more accessible than second-order entities ($e_i$),

which make reference to SoAs, and these in turn are more accessible than those which describe a possible fact ($X_i$: third-order entities), which will at the same time be more accessible than fourth-order entities ($E_i$), which are associated with speech acts. On the other hand, the Concreteness Hierarchy predicts that terms which refer to concrete entities are more accessible to Subject than those that make reference to abstract entities: concrete entities > abstract entities (Dik 1997a: 279). The reason why both these hierarchies are dealt with in the same section responds to the fact that it could be observed that their intrinsic properties are closely connected, and as a result it may be claimed that they could both be united under just one hierarchy: the Entity-Concreteness Hierarchy.

The analysis of the sample as far as the Entity Hierarchy is concerned has revealed that this hierarchy should be reformulated so that it can include zero-order entities, which make reference to properties or relations typically associated with first-order entities, and which, although in a low percentage, can also have access to Subject as in example (14) or may be explicitly encoded as the internal-A1 of a passive construction (example (15)):

14) … and its bow front gave it [a chest of drawers] an elegance which (zero) pleased them both (1st) (1504/P16-115).
15) ... enthusiasms which sometimes brought him to the verge of absurdity, where he (1st) was saved by his sharp wit (zero) (168/G16-107).

Thus, it was felt that these types of entities should be grouped with the higher-order entities since they are in fact less accessible to Subject than first-order entities, and be placed to the right of the hierarchy under a different name: first-order entities > other-order entities (or non-first-order entities). No specification among the non-first order entities was necessary since in fact 98.8% of the examples include a first-order entity in the description of the predication. However, it was possible to predict different degrees of accessibility among the non-first-order entities when these compete with a first-order entity. In this way, I can conclude that second-order entities are the ones which most frequently have access to Subject, followed by third-order entities and fourth-order entities, and all these are in turn more accessible than zero-order entities.

For the analysis of the sample in terms of the Concreteness Hierarchy, I have followed Quirk *et al*., who claim that concrete entities are those which are "accessible to the senses, observable, measurable" (1985: 5.3) whereas abstract nouns are those which refer to "unitary phenomena (such as events) on the one hand, or to states, qualities, activities, etc on the other" (1985: 5.58). The examples of embedded noun clauses (both finite and non-finite) have been given the property abstract: "semantically these clauses are normally abstract; i.e. they refer to such abstractions as events, facts, dates and ideas rather than to percep-

tible objects" (Quirk *et al*. 1985: 15.2).

These two types of intrinsic constraints overlap in the sense that first-order entities must necessarily be concrete whereas non-first-order entities must necessarily be abstract, which means that these two hierarchies could conflate in the following way: first-order concrete entities > other-order abstract entities (first-order concrete entities > or non-first-order abstract entities).

The results obtained from the study of the relevant data as far as the Entity-Concreteness Hierarchy is concerned show that the hierarchy is obeyed in different degrees depending on the type of argument that is in Subject position. In the cases in which a first argument has had access to Subject (active constructions), the hierarchy is obeyed in 97.1% of the examples (example 16). However, if a non-first argument has been assigned Subject function (passive constructions, example (17)) this percentage is reduced to 7.9% (Table 8).

16) It was suspected [by us (1st-concr.)] that the lack of freedom in the drag hinges was the possible cause (3r-abstr.) (332/J73-15).
17) You (1st-concr.) must try what people do in hot countries (2nd-abstr.) (2135/F31-170).

Table 8. Validity of the Entity-Concreteness Hierarchy

| Entity-Concreteness | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 23 | 7.9% | 679 | 97.1% |
| Violated | 269 | 92.1% | 20 | 2.9% |
| Total | 292 | | 699 | |

The behaviour of the Entity-Concreteness Hierarchy in passive constructions may be explained in relation to the type of predicates involved. Relational and psychological predicates, for instance, normally require abstract (non-first order) entities, whereas action predicates tend to occur with concrete entities. Thus, in passive constructions with relational and psychological predicates which are common in scientific and academic discourse, it will be more likely that the Subject will be an abstract non-first order entity, which justifies the high number of examples of violation of the hierarchy (97.1%)

4.3.6. The Term Hierarchy

As I briefly mentioned before, I have suggested a new parameter in the study of term accessibility to Subject assignment in virtue of their internal complexity, the Term hierarchy, which claims that simple terms are more accessible to Subject than finite complex terms and these in turn more than non-finite complex terms (see Rodríguez-Juárez 2003: 393-409; 2006).

Terms are linguistic expressions whose function is to refer to different types of entities in the real and imaginary worlds and which can be divided into two types depending on their internal constitution: simple (or primary) terms and complex (or secondary) terms. Following Dik (1997a) and Martín-Arista (1999: 184), simple terms are those whose structure is that of a simple nominal group whose referent is prototypically a first-order entity but which could also refer to any of the higher-order entities, as in example (18), in which the Subject refers to a second-order entity. Complex terms, which refer to second, third and fourth-order entities, have the internal structure of embedded finite (example (19)) and non-finite (example (20)) constructions:

18) Senator Robertson's committee has to pass Mr Weaver's nomination before it can be considered by the full Senate (13/A01-105).
19) What success Hahnemann had in Clarence's case is not known (441/G06-103).
20) There are often few chairs on steamers which visit Adriatic islands, and those few are shackled together, to be queued for until a morose sailor consents to unlock them (341/K22-12).

Apart from the purely syntactic constraints which may condition the higher or lower degree of accessibility of terms to Subject, other competing motivations such as ease of language processing may condition the order in which constituents are presented. Thus, complex terms are considered to be more difficult to produce and process (partly due to its abstract nature), and, consequently, in information planning, speakers will tend to choose a structure which the hearer will be able to interpret with the least amount of processing effort (Van Valin – Lapolla 1997: 201). Moreover, the fact that non-finite clauses as Subject show a higher degree of syntactico-semantic compression and are as a result more difficult to process than finite clauses, which are typically introduced by subordinating markers and have explicit Subjects, explains why non-finite terms as Subject are much less frequent and as a result more marked constructions, thus occupying the right extreme of the hierarchy. Another competing motivation underlying the preference of simple terms as Subjects is associated with the way in which the information is organized within the overall dis-

course. Thus, for instance, noun phrases (i.e. simple terms) are the constituents that are typically used as topics in Subject position, although in lower percentages other constituents such as prepositional phrases, adverbs and even whole clauses (complex terms) might be marked as topics (Davison 1984: 806-809). This fact might also explain why simple terms are less marked as Subjects and, as a result, are more frequent than embedded constructions which are more marked and consequently less usual and more complex.

The lower degree of occurrence of complex terms in Subject position is also due to the general tendency in information packaging constructions to place heavy constituents at or towards the end of the clause as established by the LIPOC principle (language-independent preferred order of constituents): "other things being equal, constituents prefer to be placed in order of increasing complexity … Clitic < Pronoun < Noun Phrase < Adpositional Phrase < Subordinate clause" (Dik 1997b: 127).

Thus, the analysis of the data as far as the complexity of terms is concerned has shown that the accessibility of terms to Subject function in terms of their structural complexity may be presented in the Term Hierarchy which predicts that simple terms are more accessible to Subjects than finite complex terms and these in turn are more accessible than non-finite complex terms: simple terms (nominal groups) > complex terms (finite clauses) > complex terms (non-finite clauses). In fact, the validity of the Term Hierarchy is verified in 95.6% of the passive examples and in 99.9% of the active constructions analysed in the data (Table 9).

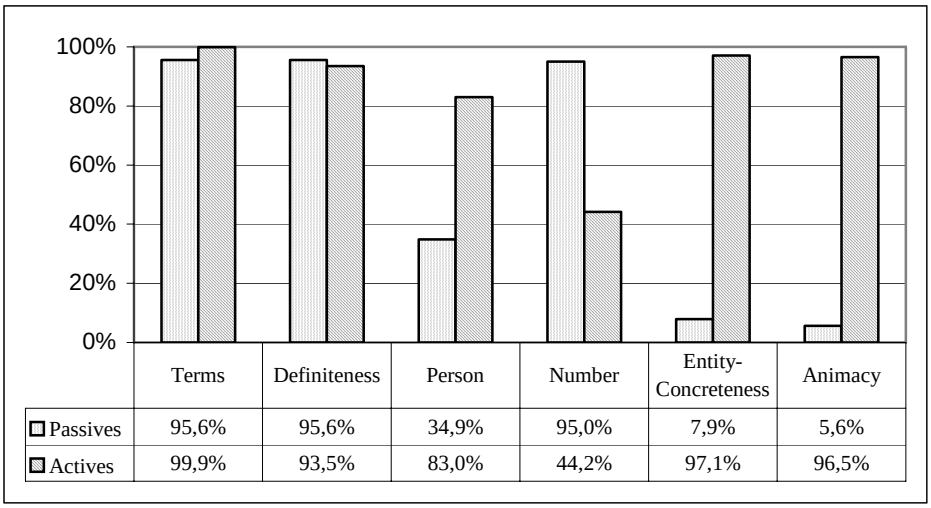
Table 9. Validity of the Term Hierarchy

| Terms | Passives | | Actives | |
|---|---|---|---|---|
| | No. | Global % | No. | Global % |
| Fulfilled | 762 | 95.6% | 1515 | 99.9% |
| Violated | 35 | 4.4% | 1 | 0.1% |
| Total | 797 | | 1516 | |

## 5. Conclusion: The Prioritising Hierarchy

It seems evident from the results obtained from the analysis of the sample of active and passive constructions that the different priority hierarchies presented as relevant in the grammatical operation of Subject assignment in the English language operate at the same time, establishing different degrees of satisfaction among themselves and showing that some of the properties gathered in the hierar-

chies take priority over others in the accessibility of a particular term to Subject. On the other hand, this study has also demonstrated that the level of fulfilment of the individual hierarchies varies on many occasions in virtue of the type of argument which has had access to Subject. Figure 1 sketches the degree of satisfaction of the different hierarchies in both active and passive constructions separately.




| | Terms | Definiteness | Person | Number | Entity-Concreteness | Animacy |
|---|---|---|---|---|---|---|
| Passives | 95,6% | 95,6% | 34,9% | 95,0% | 7,9% | 5,6% |
| Actives | 99,9% | 93,5% | 83,0% | 44,2% | 97,1% | 96,5% |

Figure 1. Subject assignment: validity of the priority hierarchies

The following examples illustrate the levels of fulfilment of these hierarchies in those predications in which an A1 has had access to Subject (example (21)) and also in the case that a non-first argument (passive) has been more accessible to Subject (example (22)).

21) It was the week before Jane was due to go on holiday that they saw the chest of drawers. (1749/P16-113)
22) The radioactivity measurements were made with a Mullard MX 123GM tube… (308/J12/139)

Taking these percentages of fulfilment as a basis, I have tried to formulate a new hierarchy of hierarchies which may predict the existing preferences among the different hierarchies both for active and passive constructions and which I have called the Prioritising Hierarchy. In the case of passive constructions, this new hierarchy may be initially presented as Terms > Definiteness > Number > Person > Entity-Concreteness > Animacy. For active constructions, the Prioritising Hierarchy presents the following order: Terms > Entity-Concreteness >

Animacy> Definiteness > Person > Number.

Apart from the purely numerical findings which motivate such hierarchical ordering, other coincidences among the different priority hierarchies are observed, which permit their organisation into three differentiated groups according to the type of priority presented. The first group which could be organised in the Prioritising Hierarchy is the Term Hierarchy, which presents intrinsic restrictions related to syntactic factors associated with the internal structural complexity of the potential Subject term itself. The second group shows constraints which are expressed grammatically and are associated with term operators indicating distinctions in the semantic domain with regards to the notions of definiteness, person and number. Finally, the Entity-Concreteness Hierarchy and the Animacy Hierarchy are grouped together since they both present intrinsic restrictions which are attributed to the referents of terms, rather than to terms themselves.

Thus, I can now attempt the formulation of the Prioritising Hierarchy which reflects and predicts the dominance of some priority hierarchies over others in Subject function assignment in English, taking into account that the higher or lower influence of these three groups as regards Subject selection differs depending on the kind of term which has been assigned Subject function. Thus, in the case of active constructions, the Prioritising Hierarchy adopts the following sequential ordering in which syntactic and grammatical aspects are more dominant than referential properties:

Term > {Entity-Concreteness / Animacy}> {Definition / Person / Number}

In passive constructions, on the contrary, the group of hierarchies which includes referential properties is more dominant than the group associated with term operators, and thus, the linear ordering of the Prioritising Hierarchy for those constructions in which a non-first argument has been selected as Subject is as follows:

Term > {Definition / Number / Person} > {Entity-Concreteness / Animacy}

Both in the case of active and passive constructions, within the group of hierarchies whose properties belong to the grammatical domain, the Definiteness Hierarchy is the one which principally conditions the point of view adopted in the presentation of the SoA over Person and Number.

To conclude, the analysis of the sample in terms of the hierarchical, functional and intrinsic constraints which are relevant in Subject assignment provides relevant data so as to be able to predict and express these restrictions on

the accessibility of terms to Subject assignment in English by means of a formula which summarises all these requirements:

$$\text{SubjAssig } [_{\text{e/structural}} (\text{T})_{\text{SFH: A1>Goal>Rec>(Ben)}} < \text{Prioritising Hierarchy} >]$$

This formula should be read in the following way: the grammatical operation of Subject assignment in English is restricted to terms which form part of the predication (e) and which carry one of the semantic functions presented in the SFH; likewise, the intrinsic properties attached to such terms compete with each other and manifest a scale of dominance which has been framed in the Prioritising Hierarchy. This new hierarchy predicts that the properties related to the internal and structural complexity of the term (Term Hierarchy) are more determinant than the properties associated with grammatical operators such as definiteness, number and person (Definiteness > Person > Number) in the case of passive constructions, and these in turn take priority over those properties related to the referent of the terms (Entity-Concreteness / Animacy). The description of Subject assignment in active constructions, on the contrary, differs in the sense that referential properties are more determinant than properties associated with grammatical operators.

## REFERENCES

Blecua, José. M – Gloria Clavería – Carlos Sánchez – Joan Torruela (eds.)
1999 *Filología e informática. Nuevos tecnologías en los estudios filológicos*. Barcelona: Milenio and Universidad Autónoma de Barcelona.

Brown, K. (ed.)
forthcom. *Encyclopedia of language and linguistics.* (2nd edition.) Oxford: Elsevier

Butler, Christopher – Ricardo Mairal – Javier Martín-Arista – Francisco J. Ruiz-de-Mendoza
1999 *Nuevas perspectivas en Gramática Funcional*. Barcelona: Ariel Lingüística.

Davison, Alice
1984 "Syntactic markedness and the definition of sentence topic", *Language* 60/4: 797-846.

Dik, Simon C.
1989 *The Theory of Functional Grammar*: *Part 1*: *The structure of the clause*. Dordrecht: Foris.
1997a *The Theory of Functional Grammar*: *Part 1*: *The structure of the clause*. Berlin: Mouton de Gruyter.
1997b *The Theory of Functional Grammar*. *Part 2*: *Complex and derived constructions*. Berlin: Mouton de Gruyter.

García-Ferrando, Manuel
1985 *Introducción a la estadística en sociología*. Madrid: Alianza Universidad.

Givón, Talmy
1984 *Syntax*: *A functional-typological introduction*. (2 vols.). Amsterdam – Philadelphia: John Benjamins.
Greenberg, Joseph H.
1966 *Language universals*: *With special reference to feature hierarchies*. The Hague: Mouton.
Groot, Casper de – Kees Hengeveld
2005 *Morphosyntatic expression in Functional Grammar*. Berlin – New York: Mouton de Gruyter.
Hengeveld, Kees
2004a "The architecture of a Functional Discourse Grammar", in: J. Lachlan Mackenzie – Mª de los Ángeles Gómez-González (eds.), 1-21.
2004b "Epilogue", in: Lachlan Mackenzie – Mª de los Ángeles Gómez-González (eds.), 365-378.
2005 "Dynamic expression in Functional Discourse Grammar", in: Casper de Groot – Kees Hengeveld (eds.), 53-86.
Hengeveld, Kees – J. Lachlan Mackenzie
forthc.a "Functional Discourse Grammar", in: K. Brown (ed.).
forthc.b *Functional Discourse Grammar*. Oxford: Oxford University Press.
Johansson, Stig – Eric Atwell – Roger Garside – Geoffrey Leech
1986 *The tagged LOB corpus*: *Users' manual*. Bergen: Norwegian Computing Centre for the Humanities.
Keenan, Edward L.
1976 "Towards a universal definition of 'subject of'", in: Charles N. Li (ed.), 303-333.
1987 *Universal Grammar*: *15 essays*. London: Croom Helm.
Keenan, Edward L. – Bernard Comrie
1977 "Noun phrase accessibility and universal grammar", *Linguistic Inquiry* 8: 63-99.
Kuno, Susumu
1976 "Subject, theme, and the speaker's empathy – a reexamination of relativization phenomena", in: Charles N. Li (ed.), 417-444.
Li, Charles N. (ed.)
1976 *Subject and topic*. New York: Academic Press.
Mackenzie, J. Lachlan
2000 "First things first. Towards an Incremental Functional Grammar", *Acta Linguistica Hafniensia* 32: 23-44.
2004 "Functional Discourse Grammar and language production", in: J. Lachlan Mackenzie – Mª de los Ángeles Gómez-González (eds.), 179-195.
Mackenzie, J. Lachlan – Mª de los Ángeles Gómez-González (eds.)
2004 *A new architecture for Functional Grammar*. Berlin and New York: Mouton de Gruyter.
Martín Arista, Javier
1999 "Expresión y motivación en gramática funcional", in: Christopher Butler *et al*., 173-217.
Martínez-Vázquez, Montserrat (ed.)
1996 *Gramática contrastiva Inglés-Español*. Huelva: Publicaciones de la Universidad.

Neuber, Keith A.
1980 *Needs assessment. A model for community planning*. Newbury Park: SAGE Publications.
Quirk, Randoph – Sidney Greenbaum – Geoffrey Leech – Jan Svartvik
1985 *A comprehensive grammar of the English language*. London: Longman.
Rodríguez-Juárez, Carolina
2003 La Asignación de la Función Sujeto en Inglés: Jerarquías y prioridades. [Unpublished Ph.D. dissertation, University of La Laguna.]
2006 "A new parameter for the description of Subject Assignment: The Term Hierarchy", *ATLANTIS* 28/1: 71-88.
Ruiz-Yamuza, Emilia
1996 "Voz, función sintáctica y función pragmática en la predicación", in: Montserrat Martínez-Vázquez (ed.), 199-228.
1999 "Sintaxis, semántica y pragmática de la asignación de sujeto en pasiva en griego clásico". *Revista Española de Lingüística* 28/2: 333-358.
Siewierska, Anna
1991 *Functional Grammar*. London: Routledge.
Van Valin, Robert D. – Randy J. Lapolla
1997 *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
Vossen, Piek (ed.)
2001 *EuroWordNet top-ontology*, available at http://www.illc.uva.nl/EuroWordNet/corebcs/ewnTopOntology.html#_Toc419884306 (date of access 17 June 2005).
Wolvengrey, Arok
2005 "Inversion and the absence of grammatical relations in Plains Cree", in: Casper de Groot – Kees Hengeveld (eds.), 419-445.

INTERNET SOURCES

Artstein, Ron
1998 "Hierarchies", available at http://privatewww.essex.ac.uk/~artstein/
Hengeveld, Kees – Tanneke Heesakkers
2004 "Valency in Functional Discourse Grammar", available at http://home.hum.uva.nl/oz/hengeveldp/ (date of access: 3 February 2005).

Appendix 1

Table 1. Passive with *be*: Declaratives (positive / modals)

| GENRES | EXAMPLES (total sample) | | SAMPLE BY GENRE |
|---|---|---|---|
| A | 830 | (9 .6%) | 37 |
| B | 468 | (5.4%) | 21 |
| C | 230 | (2.6%) | 10 |
| D | 386 | (4.4%) | 17 |
| E | 933 | (10.7%) | 41 |
| F | 746 | (8.6%) | 33 |
| G | 1,197 | (13.8%) | 53 |
| H | 762 | (8.8%) | 34 |
| J | 2,113 | (24.4%) | 93 |
| K | 227 | (2.6%) | 10 |
| L | 243 | (2.8%) | 11 |
| M | 53 | (0.6%) | 2 |
| N | 184 | (2.1%) | 8 |
| P | 184 | (2.1%) | 8 |
| R | 103 | (1.2%) | 5 |
| TOTAL | 8,659 | (100%) | 383 |

Table 2. Passive with *be*: Declaratives (negative)

| GENRES | EXAMPLES (total sample) | | SAMPLE BY GENRE |
|---|---|---|---|
| A | 10 | (7.1%) | 7 |
| B | 7 | (5.0%) | 5 |
| C | 4 | (2.8%) | 3 |
| D | 9 | (6.4%) | 7 |
| E | 19 | (13.7%) | 14 |
| F | 14 | (10.0%) | 10 |
| G | 20 | (14.3%) | 15 |
| H | 12 | (8.5%) | 9 |
| J | 24 | (17.1%) | 18 |
| K | 5 | (3.5%) | 4 |
| L | 6 | (4.3%) | 4 |
| M | 1 | (0.7%) | 1 |
| N | 7 | (5.0%) | 5 |
| P | 1 | (0.7%) | 1 |
| R | 1 | (0.7%) | 1 |
| TOTAL | 140 | (100%) | 104 |

Table 3. Passive with *be* + adverb + participle

| GENRES | EXAMPLES (total sample) | | SAMPLE BY GENRE |
|---|---|---|---|
| A | 64 | (8.5%) | 22 |
| B | 44 | (5.8%) | 15 |
| C | 26 | (3.4%) | 9 |
| D | 38 | (5.0%) | 13 |
| E | 73 | (9.7%) | 25 |
| F | 73 | (9.7%) | 25 |
| G | 126 | (16.7%) | 44 |
| H | 55 | (7.3%) | 19 |
| J | 176 | (23.3%) | 61 |
| K | 19 | (2.5%) | 7 |
| L | 17 | (2.2%) | 6 |
| M | 2 | (0.3%) | 1 |
| N | 14 | (1.8%) | 5 |
| P | 17 | (2.2%) | 6 |
| R | 11 | 11(1.4%) | 4 |
| TOTAL | 755 (100%) | 755 (100%) | 262 |

Table 4. Passive with *be* + pronoun + participle

| GENRES | EXAMPLES (total sample) | SAMPLE BY GENRE |
|---|---|---|
| A | 1 (9.1%) | 1 |
| B | 1 (9.1%) | 1 |
| C | 0 (0.0%) | 0 |
| D | 0 (0.0%) | 0 |
| E | 0 (0.0%) | 0 |
| F | 1 (9.1%) | 1 |
| G | 1 (9.1%) | 1 |
| H | 0 (0.0%) | 0 |
| J | 0 (0.0%) | 0 |
| K | 4 (36.3%) | 4 |
| L | 1 (9.1%) | 1 |
| M | 0 (0.0%) | 0 |
| N | 1 (9.1%) | 1 |
| P | 1 (9.1%) | 1 |
| R | 0 (0.0%) | 0 |
| TOTAL | 11 (100%) | 11 |

Table 5. Passive with *be* + singular and plural common noun + participle

| GENRES | EXAMPLES (total sample) | SAMPLE BY GENRE |
|---|---|---|
| B | 1 (33.3%) | 1 |
| D | 2 (66.7%) | 2 |
| TOTAL | 3 (100%) | 3 |

Table 6. Passive with *be* + singular proper noun + participle

| GENRES | EXAMPLES (total sample) | SAMPLE BY GENRE |
|---|---|---|
| B | 1 (100%) | 1 |
| TOTAL | 1 (100%) | 1 |

Table 7. Passive with *be* + article (the, no, a, an, every) + singular and plural common noun + participle

| GENRES | EXAMPLES (total sample) | SAMPLE BY GENRE |
|---|---|---|
| D | 1 (25.0%) | 1 |
| F | 1 (25.0%) | 1 |
| J | 1 (25.0%) | 1 |
| L | 1 (25.0%) | 1 |
| TOTAL | 4 (100%) | 4 |

Appendix 2

Table 1. Active transitive: verb + noun

| TYPES OF NOUNS + EXAMPLES IN THE GLOBAL CORPUS (4,927) | SAMPLE (370) |
|---|---|
| singular common noun (2,210) | 165 |
| plural common noun (1,189) | 89 |
| abbreviated unit of measurement unmarked for number (63) | 5 |
| singular proper noun (1,453) | 109 |
| singular locative noun with word initial capital (5) | 1 |
| plural proper noun (7) | 1 |

Table 2. Active transitive: verb + pronoun

| TYPES OF PRONOUNS + EXAMPLES IN THE GLOBAL CORPUS (7,010) | SAMPLE (379) |
|---|---|
| possessive pronoun (14) | 1 |
| personal pronoun. 1st pers. sing. acc. (*me*) (805) | 43 |
| personal pronoun. 1st pers. plural acc. (*us*) (348) | 19 |
| personal pronoun. 2nd pers. (*you*. *thou*. *thee*. *ye*) (772) | 42 |
| personal pronoun. 3rd pers. sing. acc. (*it*) (1,878) | 101 |
| personal pronoun. 3rd pers. sing. acc. (*him*. *her*) (1,992) | 107 |
| personal pronoun. 3rd pers. plural acc. (*them*. ‘*em*) (752) | 41 |
| nominal pronoun (*anybody*. *anyone*. *anything*; *everybody*. *everyone*. *everything*; *nobody*. *none*. *nothing*; *somebody*. *someone*. *something*) (448) | 24 |
| nominal pronoun + genitive (1) | 1 |

Table 3. Active transitive: verb + determinant + noun

| TYPES OF DETERMINANTS + NOUN + EXAMPLES IN THE GLOBAL CORPUS (19,280) | SAMPLE (392) |
|---|---|
| singular article (*a*. *an*. *every*) (4,412) | 89 |
| singular or plural article (7,855) | 159 |
| singular determiner (*another*. *each*. *that*. *this*) (942) | 18 |
| plural determiner (*these*. *those*) (179) | 4 |
| singular or plural determiner (*any*. *enough*. *some*) (530) | 11 |
| determiner/double conjunction (*either*. *neither*) (39) | 1 |
| post-determiner (*few*. *fewer*. *former*. *last*. *latter*. *least*. *less*. *little*. *many*. *more most*. *much*. *next*. *only*. *other*. *own*. *same*. *several*. *very*) (588) | 12 |
| others (9) | 1 |
| pre-qualifier (quite. rather. such) (139) | 3 |
| pre-quantifier (all. half) (277) | 6 |
| pre-quantifier/double conjunction (both) (62) | 1 |
| cardinal (781) | 18 |
| possessive determiner (3,467) | 69 |