

A NEW PARAMETER FOR THE DESCRIPTION OF SUBJECT ASSIGNMENT: THE TERM HIERARCHY

Carolina Rodríguez Juárez
Universidad de Las Palmas de Gran Canaria
crodriguez@dfm.ulpgc.es

The theoretical framework of Functional Grammar proposed by S. C. Dik highlights the relevance of implicational hierarchies for the different grammatical operations found in natural languages. In the case of the grammatical operation of Subject assignment, a group of priority hierarchies which predict the accessibility of term positions by virtue of their intrinsic properties have been claimed to directly impinge on the operation of Subject assignment. These typological hierarchies present term properties which are related to the term itself and to their referents. After balancing the postulates of Classical Dikkean Functional Grammar in general and the above-mentioned priority hierarchies in particular to a sampling of written material from the LOB Corpus, the main conclusion that emerges is that the internal structural complexity of the term must be recognised as a new relevant parameter for Subject assignment in English. Thus, I propose a new hierarchy, viz. the Term Hierarchy, which predicts the accessibility of term positions taking into account the internal structural complexity of the term in question.

Key words: accessibility, functional (discourse) grammar, implicational hierarchies, Subject assignment

1. Introduction

Within the model of Functional Grammar (henceforth FG) as outlined in Dik (1997a, b), the different grammatical operations which may be found in natural languages such as relativisation, Q-word questioning, anaphora, raising, and Subject/Object assignment are closely related to the concept of *accessibility*, which is crucial to understanding the possibilities terms have to take part in grammatical operations. In Classical Dikkean Functional Grammar, the notion of accessibility which Dik borrows from Keenan (1976, 1987) and Keenan and Comrie (1977) is defined 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. If it is inaccessible to O, there is apparently some accessibility constraint which prevents O from applying to T” (Dik 1997b: 356). The constraints which govern term accessibility have a hierarchical, functional, and intrinsic nature and 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. Thus, the properties restricting the accessibility of terms are

organised in a hierarchical fashion, with those features which are more accessible located in the first positions of such 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).¹

One of the original ambitions of Classical Dikkean Functional Grammar was to devise a grammar which apart from being pragmatically and typologically adequate was also psychologically adequate. However, it is only now that FG has developed into a new architecture known as Functional Discourse Grammar (FDG) that psycholinguistic findings and explanations have really been taken into account (Hengeveld 1997, 2004a/b; Hengeveld and Mackenzie [to appear]). In such a theoretical model, and with the main aim of achieving psychological adequacy, production is described as a top-down rather than bottom-up process in which the speaker’s intention is the starting point, leading to the proposal of a radical shift from sentence to discourse in the object of study. Another contribution of FDG is the postulation of a modular reorganisation of the different levels of representation which constitute the grammatical component (included in a wider theory of verbal communication, where a conceptual, a contextual, and output component are also envisaged) with the interpersonal, representational, and structural levels working simultaneously and individually organised into a hierarchical layering (Mackenzie and Gómez-González 2004). In this attempt to improve FG, 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).

The main concern of the current study is to study the syntactic function Subject with regard to a particular language, viz. English. 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 State of Affairs (henceforth SoA) (Dik 1997a: 251); rather, they are now regarded as grammatical notions which are the result of pragmatic and semantic choices at higher levels (Hengeveld 2004b: 373-74).

In order to provide a descriptive analysis of Subject in English, a written sampling gathered from the LOB corpus was analysed, which was made up of two groups of

¹ It must be noted that the Dikkean definition of accessibility differs from the mainstream sense of this term as conceived in psycholinguistics (Clark and Clark 1977: 474-76), where accessibility is ascribed to the production of utterances with special relevance for adjective ordering and phonological production. Thus, for instance, adjective accessibility determines that the most intrinsic and less subjective adjective will be placed closer to the noun; phonological accessibility predicts that the phonological shape of common words will be retrieved more readily than that of rare words.

constructions which differ in the constituent which carries Subject function.² On the one hand, an analysis was carried out of those instances in which the first argument (with any of the semantic functions belonging to it: Agent, Positioner, Force, Processed [Exp], Zero [Exp]) has been more accessible to Subject than any other term in the predication resulting in *active* constructions. The second group, by contrast, is constituted by structures in which Subject function has been assigned to a non-first argument (to second or third arguments) resulting in what has traditionally been described as *passive* constructions. The global sample includes 2,313 examples, of which 797 are examples of passive sentences and 1,516 are instances of active constructions.³

In a previous study (Rodríguez-Juárez 2003), the relevant Subject positions (as well as the Object positions of active sentences and the implicit or explicit *by*-phrases of passive constructions which could potentially be selected as Subject) were analysed with reference to different priority hierarchies which have been proposed as relevant to the grammatical operation of Subject assignment and which seem to directly influence the chance for a term to be assigned the function Subject (and also Object): the Definiteness, Person, Number, Animacy, Concreteness, and Entity Hierarchies (Dik

² The reason why the LOB corpus of British English was chosen for this study lies in the fact that its usefulness has been widely claimed (Kennedy 1991: 98; 1996: 218) and attested in a variety of research studies and grammars which have used it for the description and exemplification of various phenomena (Biber et al. 1998; Huddleston and Pullum 2002; Halliday and Matthiessen 2004). Moreover, I decided to initially 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.

³ In order to gather the relevant constructions needed for our analysis, different verbal *search-words* are used, which show a list of *concordances* including the relevant verb in its context and two or, if required, more lines before and after the search-word. These contexts include the Subject of active and passive constructions as well as the implicit or explicit *by*-phrase of passive verbal forms and the Object of active forms, which are the arguments which in the underlying representation of the clause could potentially be selected as Subject. The different types of search-words used are selected in terms of factors such as polarity, mood and finiteness, although the search is always conditioned by the various possibilities given by the retrieval tool (*Wordsmith Tools*) used to extract the relevant structures. In order to determine the sample size from the so-called population, a first stage was to carry out some disambiguation processes 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 (Blecu 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). These statistical methods were applied to each of the groups which constitute the population of passive and active constructions. Thus, the global number of passive and active sentences which according to the statistical methods should be analysed in order to work with a representative sample should be made up of 797 and 1,516 examples respectively, obtaining a global sample of 2,313 examples. For further information on the selection of this sample the reader is referred to Rodríguez-Juárez (2003: 189-97).

1997a: 279).⁴ A *hierarchy* is conceived in FG as a sequence “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.

Hierarchies, in addition, should be interpreted in two ways. On the one hand, they can be conceived as implicational universals which describe priorities that are typologically relevant and which typify the types of linguistic patterns which may be found across languages. Moreover, implicational hierarchies reflect aspects which differentiate natural languages 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 the constituent or constituents which can possibly be assigned Subject function in natural languages. The second aspect of 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: “(...) those items which do occur in a language (= the items preceding the cut-off point for that language in the hierarchy) will be used less and less frequently as we proceed through the hierarchy from left to right” (Dik 1997a: 33). It is within this second, more specific, descriptive interpretation that hierarchies are studied in the current investigation.

The results obtained from the descriptive analysis of the sampling of Subject positions show that, apart from the aforementioned priorities which influence the degree of accessibility for an argument to be assigned Subject function, a factor such as the internal complexity of the term might also condition the possibilities of Subject selection, with simple terms functioning as Subjects appearing more frequently than embedded complex constructions. Although the effect of the type of constructions on the organization of constituents within a clause has been noted and described by various scholars such as Keenan (1976) and Keenan and Comrie in the Accessibility

⁴ These hierarchies specify the following priorities: the Definiteness Hierarchy (definite > other specific > non-specific), the Person Hierarchy (first / second person > third person), the Number Hierarchy (singular number > plural number), the Animacy Hierarchy (human > other animate > inanimate force > other inanimate), the Concreteness Hierarchy (concrete entities > abstract entities), and the Entity Hierarchy (first-order entities > higher-order entities) (Dik 1997a: 279). All these hierarchies predict the term accessibility to Subject by virtue of their intrinsic properties, some of which are related to the referent of the term rather than to the term itself (Animacy, Entity, Concreteness), whereas others make reference to the grammatical operators of definiteness, number, and person (Definiteness, Person, and Number). Functional constraints, in addition, also condition the accessibility of terms to Subject and have been gathered in the Semantic Function Hierarchy (Arg-1 > Goal > Recipient > Beneficiary > Instrument > Location > Temporality) which predicts that those terms which carry any of the semantic functions belonging to the first argument will be the most accessible to Subject, followed in frequency and in level of difficulty by terms carrying the semantic functions Goal, Recipient, etc.

Hierarchy (1977: 66), determining universal constraints in relative clause formation, and Huddleston and Pullum with reference to information-packaging constructions in inversions, extrapositions, and existential, cleft, and passive clauses (2002: 1365-1447), here the formulation is attempted of a different priority hierarchy to be applied to the description of English with regard to the grammatical operation of Subject assignment, indicating the different degrees of accessibility of the constituents of a predication in terms of their internal structural constitution.

Thus, in the present paper a written corpus of English is analysed in terms of the accessibility of terms to Subject with a special focus on the degree of influence of the relevant priority hierarchies. The main conclusion ensuing from this analysis is that a further additional parameter is needed in the theory of perspective postulated by Dik as regards the structural complexity of terms (1997a: 254). Therefore, it is the main concern of this paper to provide an answer to the question of whether a new parameter in the study of Subject assignment in the English language is appropriate. In order to prove this hypothesis, in Section Two some preliminary assumptions that must be dealt with before attempting the proposal of a new hierarchy will be presented. Section Three is devoted to the presentation of the competing motivating factors which, supported by the analysis of relevant data, provided sufficient evidence for the formulation of the new hierarchy. The last section closes this paper with a proposal for a different type of intrinsic restriction which also conditions the accessibility of term positions to Subject and which predicts their degree of accessibility in terms of their internal constitution. These constraints are collected in the priority hierarchy which I have come to call the *Term Hierarchy*.

2. Terms: some preliminary assumptions

In order to attempt my own formulation of a further priority hierarchy as directly influencing the choice of Subject selection, it is advisable to first provide an account of the concept *term* as conceived within a functional approach. Terms are linguistic expressions whose main function is to refer to various types of entities which can be found in real or imaginary worlds. Within the theoretical framework of FG, there has been a reformulation of the ontological classification of entities as initially proposed by Lyons, with the addition of two types of entities to the first-, second-, and third-order entities (1977: 442-47). As a result, five types of entities are recognised in FG, each one carrying their own variable. First-order entities (x) have as referents individuals, things, and places which exist and can be conceived as being located in space. Second-order entities (e) are situated in space and time and designate SoAs describing Actions, Processes, States, and Positions, and “can be said to occur, begin, last and end; they can be perceived: watched, heard, felt, etc.; and they can be said to be sudden, gradual, violent, etc” (Dik 1997a: 292-93). The referents of third-order entities (X) are possible facts that can be situated neither in time nor space but that can be “believed, known, or thought; they can be reason for surprise or doubt; they can be mentioned, denied, and remembered; and they can be said to be true or false in relation to the occurrence of some SoA in some world” (Dik 1997a: 292-93). Fourth-order entities (E), on the other hand, are required by those predicates which designate speech acts such as giving advice

or permission, making an offer or a promise, ordering or warning, and asking a question or stating a fact. The fifth and last type of entities involved in the presentation of SoAs, zero-order entities (f), are associated with terms which refer to properties or relations which are attributable to a first-order entity. For an example of a term which refers to a zero-order entity, we have the following sentence taken from the sample:⁵

- (1) ... enthusiasms which sometimes brought him to the verge of absurdity, where he was saved by *his sharp wit*. (168/G16-107)

Apart from this classification of terms based on the type of entity they are used to refer to, a further classification depending on the internal constitution of such terms may be attempted, distinguishing between simple or primary terms and complex or secondary terms. Simple terms are simple nominal groups formed by nouns, personal pronouns, and demonstratives whose referents are prototypically first-order entities, but that may also be used to refer to higher-order entities. In the sample obtained from the corpus analysed here, different cases of examples could be found in which a simple term constituted by a third person singular pronoun was being used to pick out a second-, third-, and even fourth-order entity which had already been introduced in the previous discourse.

- (2) Senator Robertson's committee has to pass *Mr Weaver's nomination* before *it* (second-order entity) can be considered by the full Senate. (13/A01-105)

Examples of simple nominal groups headed by nouns such as *welcome/absolution*, *fact/thought* or *question/story* which refer to second-, third-, and fourth-order entities respectively were also registered:

- (3) General de Gaulle's official welcome last week to Britain's moves towards the six was taken as a friendly gesture in Whitehall. (8/A02-9)
 (4) Then a thought struck her. (1378/P20-108)
 (5) Against this industrial setting Mr Richardson has told Miss Delaney's story. (1061/C03-25)

As regards complex terms, there was evidence in the data of Siewierska's claim (1991: 33) that there seems to be a close connection between the internal complexity of terms and the type of entities they refer to, since all the terms analysed in the sample which refer to zero- or first-order entities are simple terms, whereas the types of entities complex terms generally refer to are second- (2.9%), third- (58.8%), and fourth-order

⁵ The references at the end of each example indicate (i) the number assigned to that example in the total corpus (168), (ii) the type of text from which the example was taken (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), and (iii) the number and line assigned to the text in the LOB corpus (in (1), 16 and 107, respectively).

(38.2%) entities. The internal constitution of complex terms is that of embedded constructions which, from a semantic point of view, act as restrictors (A: [Φ]), and which may be classified into embedded predications (e: [predication]), embedded propositions (X: [proposition]), and embedded clauses (E: [clause]) (Dik 1997b: 94). Furthermore, complex terms, as conceived by Dik, are required by different types of matrix predicates which impose semantic restrictions on the predicational, propositional, or clausal term which follows them. Thus, the types of matrix predicates which are relevant to Subject assignment in English are as follows. The group of predicates which require a complex predicational term may be subdivided on the one hand into directive (*order, ask*), practical manipulation (*force, cause*), and volitional (*want*) predicates, which from the semantic point of view require that the realisation of the SoA designated by the complex term be posterior (Post), and on the other hand, into predicates of direct perception (*see*), achievement (*manage, fail*), and phasal predicates (*begin, continue, stop*), whose semantic restrictions impose that the realisation of the SoA be simultaneous (Sim).⁶ As regards propositional terms, the types of matrix predicates which require them are those which express intellectual and emotional attitudes (*believe, presume; fear, hope*), manipulation (*convince, persuade, teach*), acquisition or loss of knowledge (*learn, know, forget*), and indirect mental perception (*see, hear*). Finally, the matrix predicates which require clausal terms are those designating speech acts such as *say* or *ask* (Dik 1997b: 96-113).

Complex terms may be further divided into finite embedded constructions and non-finite embedded constructions. The traditional correlate of finite embedded constructions is the subordinate nominal clause which requires the presence of an explicit subordinating marker (except in the case of *that*-clauses, where the conjunction *that* may not be specified in the sentence). Sentences six and seven below illustrate cases of Subject assignment to terms expressed by means of finite subordinate clauses, a *that*-clause and an interrogative clause introduced by a *wh*-word respectively:

- (6) That the Seljuks brought nothing but chaos and destruction to Asia Minor is not borne out by the facts. (430/G45-167).
- (7) What success Hahnemann had in Clarence's case is not known. (441/G06-103)

Any of the semantic and pragmatic functions which may be assigned to first-order simple terms may also be assigned to complex terms. As regards the type of syntactic function which may be assigned to an embedded construction, Dik manifests that "if the embedded construction occurs in second argument position, Subject assignment to it may or may not be possible" (1997b: 123). In fact, in those cases in which Subject has been assigned to an embedded term, the resulting construction is quite marked and fairly unnatural. As a result, at a higher level of the underlying clause structure, the

⁶ Dik presents a further classification for those predicates which require a predicational complement taking into account whether the SoA was finally accomplished or not (1997b: 114-15). Thus, he distinguishes between implicational predicatives such as *manage* (which imply the accomplishment or realisation of the SoA described in the complement), contra-implicative predicates such as *fail* (which imply that the SoA presented in the complex term was not finally executed) and non-implicative predicates such as *want* (which are neutral as to whether the SoA was finally realised or not).

reorganisation of the order of the constituents allows such a complex construction to be introduced by means of the *dummy it*, and be placed at the end of the embedded construction according to the LIPOC principle (*language-independent preferred order of constituents*) which among other things establishes that “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). This preference is attested in the corpus material as example eight illustrates, in which the *dummy it* occupies the Subject position and the embedded construction has been postponed and placed at the end of the main clause:⁷

- (8) *It is felt that the above correction is not entirely satisfactory* as it is based on fixed wing theory. (297/I73-133)

Non-finite embedded constructions are subdivided into infinitival constructions, which can function as first and second arguments as well as satellites; participle constructions which generally function as satellites; and nominalisations, which are constructions which present features which are typical of nouns and which can also have access to Subject and Object functions. The types of non-finite constructions which are relevant to the grammatical operation of Subject assignment are the first and last types. Among the group of infinitival constructions, Dik distinguishes between those in which all the argumental positions are explicitly represented by a term, which means that it is a *closed infinitival construction*, and those in which one of the argumental positions has been left unspecified, *open infinitival constructions* (1997b: 147-54). The Subject of closed infinitival constructions is introduced by means of the preposition *for* followed by an object form:

- (9) “Sir, it is all very well *for ‘Canuck’ to suggest* that there is no need for losing days at all for shops.” (2028/B24-150)

Non-finite clauses with Subject assignment introduced by the *dummy it* and as a result placed in final position in the sentence are normally linked to adjectival predicates rather than to verbal predicates (Dik 1997b: 149). In fact, in the present corpus no examples of closed infinitival construction were found in Subject position.

In open infinitival constructions, on the contrary, although the Subject is not explicitly encoded in the predication, it is semantically implicit, since it co-refers with the Subject of the matrix verb and can be treated as an example of zero anaphora (Dik 1997b: 148):

- (10) *Rehabilitation of refugees from East Pakistan* (e_i) still remains (Ae_i) to be accomplished. (169/G65-108)
 (11) More than once it had happened to *me* (x_i) that my reason for asking (Ax_i) to be excused attendance at St Bride’s on a given Saturday afternoon had been accepted as valid... (197/G14-138)

⁷ The LIPOC principle also predicts that no further material will occur after the subordinate clause, except if that material is as complex as or more complex than the material contained in the embedded clause (Dik 1997b: 127).

- (12) There are often few chairs on steamers which visit Adriatic islands, and *those few* (x_i) are shackled together, (Ax_i) to be queued for until a morose sailor consents to unlock them. (341/K22-12)

The covert Subject of open infinitival constructions does not always co-refer explicitly since on many occasions it has a generic value (*anybody*) which is represented by means of the operator (gx_i) indicating that “as a referent for this term, choose any entity which fulfils the selectional requirements imposed in this argument position” (Dik 1997b: 150). In the material analysed here, only one example of this type of generic Subject in an open infinitival construction has been attested:

- (13) (gx_i) *To be a little considerate about radios and gramophones and noise* generally is rated highly among good manners. (100/E26-115)

The other type of non-finite embedded construction relevant to the analysis of Subject assignment in English corresponds to nominalisations, which are characterised by presenting one or more properties regarded as typical of primary nominal terms. The first type of nominalisation proposed by Dik is that of *headed nominalisations*, whose nominal value comes from the noun around which the construction revolves. These nouns present the common feature of having a generic meaning such as *fact*, *belief*, *news*, *circumstance*, etc, and are normally followed by *that*-clauses. This last feature leads Dik to regard this type of nominal clauses as examples of finite subordinate clauses (1997b: 157), a suggestion which is followed here in the analysis of the Subject of a sentence such as the one exemplified in (14), which is regarded as a complex term realised by means of a finite subordinate clause:

- (14) ... and *the fact that Jones is British* will, he believes, reflect prestige upon Britain. (951/R01-11)

The second type of nominalisations which Dik calls *non-headed constructions* revolve around a nominalised verb and are realised by means of complex or secondary terms which refer to second-, third-, or fourth-order entities such as *John’s denial of the charges* and *the presidential elections* (1997b: 159). These terms have encoded the semantic and morphosyntactic features of first-order expressions as the Principle of Formal Adjustment predicts: “Derived, secondary constructions of type X are under pressure to adjust their formal expression to the prototypical expression model of non-derived, primary constructions of type X” (Dik 1997b: 158). Nominalised verbal predicates are initially derived from predicate formation rules, although in subsequent stages they may become lexicalised and, as a consequence, appear listed in the lexicon: “this is only one instance of a quite general process of loss of productivity and subsequent lexicalization” (Dik 1997b: 168).

Although fundamentally using Dik’s division between simple and complex terms as a basis when formulating this new hierarchy, the classification proposed by Martín-Arista was also adopted, which distinguishes two types of nominalisations depending on whether the nominalised verbal predicate has been encoded as a noun (lexical nominalisation), or as a non-finite form in infinitive or –ing (syntactic nominalisation) (1999: 184). As a result, and bearing in mind all these considerations, I analyse all the

examples of expressions which contain simple nouns, pronouns, and lexical nominalisations whose structure is that of a nominal group as simple or primary terms. Secondary or complex terms include the examples of syntactic nominalisations which have the constitution of finite and non-finite embedded constructions.

3. Formulation of the Term Hierarchy: competing motivations, evidence and facts

As Kuno has pointed out (1976: 438), syntactic manifestations should be accounted for not exclusively from a syntactic standpoint, but from a semantic perspective which might explain the behaviour of syntactic constructions. Thus, in this section, the results obtained from the descriptive analysis of the data are shown in the light of the types of terms which have been selected as Subject. These percentages provide solid evidence which will be translated into the shape of the Term Hierarchy. Apart from the purely syntactic analysis of Subject positions, this article will deal with the way in which the other intrinsic properties relevant in the operation of Subject assignment in English behave when the prediction claimed by the Term Hierarchy has been violated. Finally, other types of competing motivations which might also influence the higher degree of accessibility to Subject of simple terms over complex ones are considered, namely the organisation of information structure and the ease of language processing.

Table 1 shows the results obtained from the descriptive analysis of the types of term which have been selected as Subject in the whole sample of active and passive constructions and reveals that the more frequently attested type of Subjects has the internal structure of simple terms (98.4%), whereas only 1.6% of the Subjects analysed present the internal structure of a complex term. Thus, at this stage the Term Hierarchy that I intend to substantiate could be provisionally formulated as follows:

simple terms > complex terms or primary terms > secondary terms.

| Subject | Total corpus | |
|---------------|--------------|-------|
| | No. | % |
| Simple terms | 2277 | 98.4% |
| Complex terms | 36 | 1.6% |

Table 1. Types of Subjects

A more detailed analysis of complex terms furnishes interesting evidence as regards the type of complex term which could be assigned Subject function (Table 2). Thus, in the case of Subject assignment to a non-first argument (passive constructions) only two instances of Subject assignment to a finite subordinate clause in an initial position are registered (examples (6) and (7) above) as opposed to 31 cases with extraposed Subject introduced by anticipatory *it* (see example 8). In addition, the realisation of non-finite subordinate Subjects is even less frequent, as can be shown by the fact that only two examples are found, one of which presents the Subject term in initial position (example 13 above). Furthermore, in most of the cases in which such terms are assigned Subject

function, the resulting passive construction is rather marked as a result of the reorganisation of the constituent order of the elements partaking in the underlying representation of the predication described. Likewise, the study of the active constructions shows that only one example has undergone Subject assignment to a finite clause (example 14 above), supporting the proposed hypothesis that complex terms are less accessible to Subject than simple terms and that, when this is the case, such constructions are held to be rather marked. Finally, and as far as complex terms are concerned, the figures obtained in the analysis of the data enable us to conclude that finite embedded constructions have access to Subject (1.5%) more frequently than non-finite embedded ones (0.1%).

| Embedded clauses | Passives | | Actives | | Total corpus | |
|------------------|----------|------|---------|------|--------------|------|
| | Nº | % | Nº | % | Nº | % |
| Finite | 33 | 1.4% | 1 | 0.1% | 34 | 1.5% |
| Non-finite | 2 | 0.1% | 0 | 0.0% | 2 | 0.1% |

Table 2. Subject accessibility of embedded constructions

The most common linguistic pattern which has been attested in the total corpus of active and passive constructions, as far as the analysis of the Term Hierarchy is concerned, is associated with cases in which the two terms which could potentially be assigned Subject function (Subject and Object of active constructions and Subject and *by-phrase* of passive constructions) have the internal constitution of simple terms (ST) (see Table 3); this pattern (ST > ST) has been registered in 82.2% of the cases. In those instances in which both a simple term and a complex term (finite clause (FC) or non-finite clause (NFCL)) compete with the possibility of being assigned Subject function, the simple term is assigned Subject function in 16.2% of the examples, and it is only in a very low percentage (1.6%) that a complex term (particularly a finite embedded construction (1.5%)) is more accessible than a simple term (for concrete examples see sentences number 6, 7, 8 and 13).

| Linguistic Patterns: Subj > Obj / Subj > by-phrase | No. | Global % |
|---|------|----------|
| ST > ST | 1902 | 82.2% |
| ST > FCL | 230 | 9.9% |
| ST > NFCL | 145 | 6.3% |
| FCL > ST | 34 | 1.5% |
| NFCL > ST | 2 | 0.1% |
| Total corpus | 2313 | 100% |

Table 3. Term hierarchy: linguistic patterns (global corpus)

In the light of the preceding results, the Term Hierarchy could be reformulated as follows: simple terms (nominal groups) > complex terms (finite clauses) > complex terms (non-finite clauses). This scale is intended here as a tool to study the frequency distributions of the different types of terms which according to the internal constitution of the argument might be selected as Subject in English. Tables 4 and 5 break down the percentages of passive and active constructions separately:

| Terms: Subj > by-phrase | No. | % |
|-----------------------------------|------------|----------|
| Irrelevant cases: ⁸ | | |
| ST > ST | 762 | 95.6% |
| Unfulfilled: | | |
| FCL > ST | 33 | 4.1% |
| NFCL > ST | 2 | 0.3% |

Table 4. Validity of the Term Hierarchy in passive constructions

| Terms: Subj > Obj | No. | % |
|-----------------------------|------------|----------|
| Irrelevant cases: | | |
| ST > ST | 1140 | 75.2% |
| Fulfilled: | | |
| ST > NFCL | 145 | 9.6% |
| ST > FCL | 230 | 15.2% |
| Unfulfilled: | | |
| CLF > ST | 1 | 0.1% |

Table 5. Validity of the Term Hierarchy in active constructions

In those instances in which the Term Hierarchy is unfulfilled because a complex term has been selected as Subject instead of a simple term, the priorities established by the Animacy, Entity, and Concreteness Hierarchies are also highly violated due to the fact that the complex terms which are assigned Subject function are mainly abstract, inanimate, third-, and fourth-order entities, in contrast with the other term which could have also been assigned Subject, viz. a concrete first-order entity which in most of

⁸ All those examples in which the Subject and the Object of active sentences and the Subject and *by-phrase* of passive sentences present the same feature (i.e. the two terms are either simple terms or complex terms, or in the case of the other priority hierarchies, the two terms are definite, or third-persons, or abstract) have been described as irrelevant for the present study since it is not possible to see which feature has won out the other. However, for the calculation of the global conclusions taking into account the whole population, the irrelevant examples are counted as instances in which the hierarchy is not violated, and as a result, are included in the group of examples in which the prediction established by the hierarchy is fulfilled (Table 7).

the cases has a human referent (Table 6).⁹ Those instances in which a human term has not been assigned Subject function appear to be conditioned by the priority presented in the Definiteness Hierarchy which predicts the higher degree of accessibility of a definite term (example 15) or a specific indefinite one (the embedded construction in example 16) over a specific indefinite and a non-specific indefinite term which from the semantic point of view is not specially significant (by anyone/by people).

- (15) Mr Hugh Gaitskell (specific, definite) argued that no final decision should be taken until a conference of commonwealth Premiers had been held (specific, indefinite). (2077/A06-182)
- (16) It may well be argued [by someone / people (non-specific, indefinite)] at this point that the above type of reasoning is all very well for simple two-element circuits (specific indefinite). (276/J69-113)

Most of the examples related to the study of the Person Hierarchy are irrelevant in the present discussion, since both terms competing for Subject position are third persons. The only three cases in which the hierarchy has been violated, because a third person has had access to Subject instead of a singular or plural first person, belong to learned and scientific texts (represented by J in the LOB corpus) which are normally characterised by an impersonal style. It is obvious that the differences in style and register, and especially the differences between written and spoken English, reflect differences in the type of lexical associations and grammatical constructions used. In fact, this paper is aimed to serve as a stepping stone for further projects in which spoken language could be analysed in terms of the Term Hierarchy (as well as the other priority hierarchies).¹⁰ As for the restrictions on lexical associations, it seems appropriate to highlight that the type of constructions in which the Term Hierarchy has not been

⁹ In Table 6, “d” stands for definite; “i” for indefinite, “-s” for non-specific, “t” for third person, “f” for first person, “in” for inanimate entity, “h” for human, “sg” for singular, “pl” for plural, and ordinals “o, 1st, 2nd, 3rd, 4th” for zero-, first-, second-, third-, and fourth-order entities. The percentages of fulfilment and violation have been calculated taking into account only the relevant examples in which the analysed arguments present different properties.

¹⁰ Biber et al. (1998: 73-76; 101-05) present a corpus-based study in which they analyse the distribution and function of *that*-clauses and *to*-clauses in two registers, viz. conversation and academic prose, and conclude that in general terms *that*-clauses are more common in conversational English but relatively rare in academic prose, whereas *to*-clauses show the same distribution in the two registers but are regarded as less common than *that*-clauses in conversation. These distributional differences seem to be constrained in part by differing lexical associations, which are more restricted in the case of *that*-clauses than in the case of *to*-clauses, since the latter may be complements not only to verbal predicates but also to a large number of adjectives. In fact, the grammatical associations of these constructions show that in extraposed constructions *to*-clauses are more common in academic writing than in conversation (and more frequent than extraposed *that*-clauses). This is due mostly to the fact they have a very strong association with adjectival predicates which allow the static packaging of information which is normally preferred in academic prose. As discussed in section 2 in this paper, in the current research, however, only verbal predicates are taken into account because these allow alternative choices of Subject assignment, whereas *to*-clauses as complements of adjectives do not.

accomplished feature matrix predicates which require both propositional terms (example 17) and clausal terms (example 18):¹¹

- (17) It was learned that Lord Home will afterwards go to stay privately with Queen Elizabeth, the Queen Mother at Birkhall. (27/A29-126)
- (18) Finally, it may be mentioned that as the Portuguese pushed their exploratory voyages down the west coast of Africa, they added the acquisition of negro slaves ... (284/J58-124)

| Hierarchies | Properties | No. | % | Fulfilled | Unfulfilled | Irrelevant |
|------------------------|------------|-----|-------|---------------|--------------|------------|
| Definiteness | d>i | 1 | 2.9% | 31 (88.6%) | 4 (11.4%) | 1 |
| | i>d | 4 | 11.4% | | | |
| | i>i | 1 | ----- | | | |
| | i>i(-s) | 30 | 85.7% | | | |
| Person | t>f | 3 | 100% | 0 (0.0%) | 3 (100%) | 33 |
| | t>t | 33 | ----- | | | |
| Animacy | in>h | 34 | 100% | 0 (0.0%) | 34 (100%) | 2 |
| | in>in | 2 | ----- | | | |
| Number | sg>pl | 34 | 100% | 34 (100%) | 0 (0.0%) | 2 |
| | sg>sg | 2 | ----- | | | |
| Entity- Abstraction | 2nd>1st | 1 | 2.9% | 0 (0.0%) | 34 (100%) | 2 |
| | 3rd>1st | 20 | 58.8% | | | |
| | 4th>1st | 13 | 38.2% | | | |
| | 3rd>3rd | 1 | ----- | | | |
| | 3rd>0 | 1 | ----- | | | |

Table 6. Behaviour of the priority hierarchies when the Term Hierarchy has been unfulfilled

So far, we have seen that the higher or lower degree of accessibility of terms to Subject is conditioned by purely syntactic constraints as has been presented in the Term Hierarchy, but this restriction on the complexity of terms is also motivated by factors of a different nature which condition the order in which constituents are presented. One

¹¹ Biber et al. (1998: 103) account for the differing registering patterns found between *that*-clauses and *to*-clauses in terms of the kind of verbal predicates complemented by these complex clauses. Thus, the higher level of occurrence of *that*-clauses in conversation is directly associated to their common co-occurrence with three matrix predicates: *think*, *say* and *know*: “These [predicates] are used to report two of the most important activities and states of humans: what they think/know, and what they said!”, hence they are not used very often in academic writing. Examples of *that*-clauses in academic prose report research findings by using verbal predicates such as *show*, *indicate*, *suggest*, etc.

of these competing motivations is associated with the ease of language processing. Thus, complex terms are considered to be more difficult to produce and process (partly due to their 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 and LaPolla 1997: 201). Moreover, the fact that non-finite clauses as Subject are much less frequent and therefore more marked constructions than finite clauses (thus occupying the right extreme of the hierarchy) may be explained in relation to the difficulty in processing these constructions in comparison with finite clauses. These do not show such a high degree of syntactico-semantic compression, in the sense that these are typically introduced by subordinating markers and have explicit Subjects; by contrast, non-finite clauses do not exhibit their Subject in most of the cases and are not introduced by means of subordinating conjunctions. Finally, language acquisition processes also seem to reflect the relationship between linguistic complexity and language production in the sense that it has been shown that children take longer to learn complex language devices (Clark and Clark 1977: 338).

On the other hand, the selection of a particular type of term as Subject may also be motivated by the way the information is organized within the overall discourse. Recent works on how information structure conditions syntactic structure have been written among others by Halliday (1967, 1985), Kuno (1972a, 1972b), Prince (1981), Davison (1984), Lambrecht (1994), Van Valin and LaPolla (1997), and Halliday and Matthiessen (2004). Thus, for instance, Davison claims that 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 (1984: 806-09). 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.¹² This tendency justifies the higher degree of occurrence of extraposed embedded clauses as opposed to the basic ones, a phenomenon which is further motivated by the fact that it is easier to process the subordinate clause in those instances with extraposition than when the complex clause is in the initial Subject position (Huddleston and Pullum 2002: 1382; 1403-05). These underlying motivations which explain the higher degree of accessibility to Subject of simple terms over complex terms are more in the line of the postulates of FDG, since it combines language and cognition/processing and resorts to discourse related phenomena as competing factors determining Subject selection.

4. Conclusion

It is an uncontroversial fact, at least within Classical Dikkean Functional Grammar, that the accessibility of a term to Subject function is conditioned and restricted by

¹² In Classical Dikkean terminology this tendency is referred to as the LIPOC principle (see section two in this paper).

hierarchical, functional, and intrinsic constraints. These restrictions are presented in the form of implicational priority hierarchies which are associated with semantic and intrinsic properties attributed to such terms and which were briefly presented in the introduction to this paper. In this study, one of the main concerns is to prove the hypothesis which suggests that a new parameter in the study of Subject assignment on the English language is possible, and indeed necessary, as regards the internal constitution of terms. Moreover, the higher or lower degree of accessibility of terms to Subject function in terms of their internal constitution is also supported by other competing motivations such as ease of language processing and information structure. Thus, taking all these underlying factors into account, a new priority scale, the Term Hierarchy, is presented, which predicts term accessibility to Subject function in terms of its structural internal complexity. The formulation of this new priority hierarchy also observes the distinction between simple and complex terms, with simple terms being more accessible to Subject position than complex ones (i.e. finite and non-finite embedded constructions). Schematically, this hierarchy could be represented like this: simple terms > finite complex terms > non-finite complex terms.

The applicability and validity of this hierarchy is tested in the analysis of written data both in those constructions in which a first argument has been assigned Subject function (active sentences) and in those in which a non-first argument has been more accessible to Subject (passive sentences). The analysis entertained in this paper can be taken to provide incontrovertible evidence that the accessibility of a complex term to Subject is less frequent and more marked (1.6%) than Subject assignment to a simple term (98.4%) both in active and in passive constructions (Table 7).

| Terms | Passives | | Actives | | Total corpus | |
|-------------|----------|----------|---------|----------|--------------|----------|
| | No. | % global | No. | % global | No. | % global |
| Fulfilled | 762 | 32.9% | 1515 | 65.5% | 2277 | 98.4% |
| Unfulfilled | 35 | 1.5% | 1 | 0.0% | 36 | 1.6% |

Table 7. Validity of the Term Hierarchy: global conclusions

In this paper I hope to have made some contribution towards a better understanding of the theory of perspective in general and Subject assignment in particular through the postulation of a new parameter within those intrinsic constraints on the accessibility of terms to Subject function assignment in English.

Works Cited

- Biber, Douglas, Susan Conrad, and Randi Reppen 1998: *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge UP.
- Blecua, José Manuel, Gloria Clavería, Carlos Sánchez, and Joan Torruella, eds. 1999: *Filología e Informática: Nuevas Tecnologías en los Estudios Filológicos*. Barcelona: Milenio y Universidad Autónoma de Barcelona.

- Clark, Herbert H. and Eve V. Clark 1977: *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Davison, Alice 1984: "Syntactic markedness and the definition of sentence topic." *Language* 60.4: 797-846.
- Dik, Simon C. 1997a: *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Ed. Kees Hengeveld. Berlin and New York: Mouton de Gruyter.
- 1997b: *The Theory of Functional Grammar. Part 2: Complex and Derived Constructions*. Ed. Kees Hengeveld. Berlin and New York: Mouton de Gruyter.
- Halliday, Michael A. K. 1967: "Notes on transitivity and theme in English." *Journal of Linguistics* 3: 37-81 and 199-244.
- 1985: *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, Michael A. K. and Christian Matthiessen 2004: *An Introduction to Functional Grammar*. 3rd ed. London: Edward Arnold.
- Hengeveld, Kees 1997: "Cohesion in Functional Grammar." *Discourse and Pragmatics in Functional Grammar*. Ed. John H. Connolly et al. Berlin and New York: Mouton de Gruyter. 1-16.
- 2004a: "The architecture of a Functional Discourse Grammar." *A New Architecture for Functional Grammar*. Eds. J. Lachlan Mackenzie and M^a de los Ángeles Gómez-González. Berlin and New York: Mouton de Gruyter. 1-21.
- 2004b: "Epilogue." *A New Architecture for Functional Grammar*. Eds. J. Lachlan Mackenzie and M^a de los Ángeles Gómez-González. Berlin and New York: Mouton de Gruyter. 365-78.
- Hengeveld, Kees and J. Lachlan Mackenzie (to appear): "Functional Discourse Grammar." *Encyclopaedia of Language and Linguistics*. 2nd ed. Ed. Keith Brown. Oxford: Elsevier.
- Huddleston, Rodney and Geoffrey K. Pullum 2002: *The Cambridge Grammar of the English Language*. Cambridge: Cambridge UP.
- Keenan, Edward L. 1976: "Towards a universal definition of 'subject of'." *Subject and Topic*. Ed. Charles N. Li. New York: Academic. 303-33.
- 1987: *Universal Grammar: 15 Essays*. London: Croom Helm.
- Keenan, Edward L. and Bernard Comrie 1977: "Noun phrase accessibility and universal grammar." *Linguistic Inquiry* 8: 63-99.
- Kennedy, Graeme 1991: "Between and trough: The company they keep and the functions they serve." *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. Eds. Karin Aijmer and Bengt Altenberg. London and New York: Longman. 95-110.
- 1996: "The corpus as a research domain." *Comparing English Worldwide: The International Corpus of English*. Ed. Sidney Greenbaum. Oxford: Clarendon. 217-26.
- Kuno, Susumu 1972a: "Functional Sentence Perspective: a case study form Japanese and English." *Linguistic Inquiry* 3: 269-320.
- 1972b: "Pronominalization, reflexivization, and direct discourse." *Linguistic Inquiry* 3: 161-96.
- 1976: "Subject, theme, and the speaker's empathy – a reexamination of relativization phenomenon." *Subject and Topic*. Ed. Charles N. Li. New York: Academic. 417-44.
- Lambrecht, Knud 1994: *Information Structure and Sentence Form*. Cambridge: Cambridge UP.
- Lyons, John 1977: *Semantics*. 2 vols. Cambridge: Cambridge UP.
- 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." *A New Architecture for Functional Grammar*. Eds. J. Lachlan Mackenzie and M^a de los Ángeles Gómez-González. Berlin and New York: Mouton de Gruyter. 179-95.
- Mackenzie, J. Lachlan and M^a 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." *Nuevas perspectivas en gramática funcional: Los estándares de adecuación de la teoría*. Ed. Christopher Butler et al. Barcelona: Ariel Lingüística. 173-217.
- Neuber, Keith A. 1980: *Needs Assessment: A Model for Community Planning*. Newbury Park: SAGE.
- Prince, Ellen F. 1981: "Toward a taxonomy of given-new information." *Radical Pragmatics*. Ed. Peter Cole. New York and London: Academic. 223-55.
- Rodríguez-Juárez, Carolina 2003: "La asignación de la función Sujeto en inglés: jerarquías y prioridades." (unpublished PhD dissertation). La Laguna: Department of English and German Philology, University of La Laguna.
- Siewierska, Anna 1991: *Functional Grammar*. London: Routledge.
- Van Valin, Robert D., Jr. and Randy J. LaPolla 1997: *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge UP.