MARKEDNESS AND THE HIERARCHY
OF SUBJECT PROTOTYPICALITY*

Francisco Javier Martín Arista
Beatriz Martínez Fernández
Laura Caballero González
Elisa González Torres
Universidad de La Rioja

ABSTRACT: *This journal article engages in the realization of the English subject from a quantitative perspective and takes steps towards enlarging the framework of nominal prototypicality set up by Langacker (1991: 142). The hierarchy of subject prototypicality that follows from a quantitative analysis for independent clauses is Nominal > Pronominal >Complex > Dummy > Clausal > Adjectival > Adverbial > Prepositional; for dependent clauses this hierarchy is re-interpreted as follows: relative pronoun > other pronouns > noun > complex > dummy > clausal. At the theoretical level, a quantitative approach to subject prototypicality contributes to the study of frequency markedness and prototypicality judgements on the one hand and substantive markedness on the other.*

KEYWORDS: *functionalism, cognitive linguistics, markedness, subject*

RESUMEN: *Este artículo se ocupa de la realización del sujeto en inglés desde una perspectiva cuantitativa y constituye un paso adelante en el estudio de la prototipicidad nominal tal y como ésta se entiende en Langacker (1991: 142). La jerarquía que se sigue del análisis cuantitativo para las oraciones independientes es* Nominal > Pronominal >Complex > Dummy > Clausal > Adjectival > Adverbial > Prepositional*; para las oraciones dependientes la jeraraquía que se propone es* relative pronoun > other pronouns > noun > complex > dummy > clausal. *En el plano teórico, un análisis cuantitativo de la prototipicidad del sujeto contribuye al estudio de la marca textual y los juicios de prototipicidad de una parte y al estudio de la marca sustantiva de otra.*

PALABRAS CLAVE: *funcionalismo, lingüística cognitiva, marca, sujeto*

1. INTRODUCTION

Martín Arista (1998a), offers a theory-independent approach to the English subject by putting forward that leakage, blocking properties and realization must be taken into account -along with Keenan's (1976) coding and control properties- in the definition of this grammatical function. In this journal article we examine the issue of the realization of the English subject from a quantitative perspective and take steps towards enlarging the framework of nominal prototypicality set up by Langacker (1991: 142). It is not our purpose, however, to offer an exhaustive study of nominal realization, not even of the realizations of the subject function. On the contrary, we report a piece of programmatic research that constitutes the beginning of a research project that will involve the following stages:

(i) Discussion of the qualitative aspects of nominal realization, definition of variables for a quantitative study of nominal realization, and analysis of the variables on a preliminary corpus. The results of the first stage are rendered in this paper.
(ii) Study of a larger corpus of subject realizations that includes, at least, the general, technical, scientific and literary registers of the English language;
(iii) Study of a corpus of object realizations that has the same characteristics as the one described in (ii);
(iv) Comparison of subject and object realizations.

Throughout these stages, our aim is to draw conclusions on the following respects:

(i) The relationship between structural, frequency and cognitive markedness;
(ii) The hierarchy of nominal prototypicality;
(iii) The hierarchy of subject and object topicality;

Methodologically, we derive our inspiration from Van Valin and LaPolla (1997), Dik (1986), Taylor (1989), Croft (1990, 1991), Langacker (1987), Nuyts (1992) and Givón (1995). More specifically, we adopt a cognitive-functional position that is characterized by three underpinnings:

(i) Stating that knowledge of language is knowledge is tantamount to saying that knowledge is knowledge of language.
(ii) Given (i), a functional grammar will improve its level of psychological adequacy by incorporating insights and generalizations from the cognitive field.
(iii) In spite of (i) and (ii), a functional grammar will proceed, both for methodological and theoretical reasons, from pragmatics and semantics, through syntax, towards the incorporation of pre-linguistic notions into the descriptive apparatus of the theory.

A remark in point is that, even though all the linguists quoted above share the view that the structure of language is externally controlled (Kuno 1980), they have different priorities when seeking linguistic explanation. Our position in this respect is that, however central context -in line with Givón (1989)- or meaning -in tune with Langacker (1987)- are, grammar certainly exists; and it is probably more elaborate than a symbolic interface between cognition and phonology (Langacker 1987); and more complex than a set of discourse principles (Givón 1989). As regards the case under scrutiny, namely the realization of the English subject, we will build the case that the structure of lexical and syntactic categories has explanatory power for a functional grammar of English[1].

2. THE APPEAL OF THE QUALITATIVE APPROACH

Very briefly, Langacker's (1991: 145) contention is that the prototypical realization of any nominal construction is the nominal phrase. Example (1) illustrates this proposal:

(1)
a.   This building
b.   No questions
c.   The central station
d.   A green armchair
e.   Those two £5 notes
f.   The three German students in this class

In these noun phrases, a syntactic head specifies a semantic type, while a syntactic sequence of a compulsory determiner plus an optional modifier grounds an instance of that type. Langacker justifies his choice of prototypicality on the basis of meaning-form iconicity: grounding and type are iconically reflected in the syntactic constituency represented by the labeled bracketing representations that follow:

(2)
a.   [[This]Det [building]Head]NP
b.   [[No]Det [questions]Head]NP
c.   [[The]Det [central]Mod [station]Head]NP
d.   [[A]Det [green]Mod [armchair]Head]NP
e.   [[Those]Det [two]Det [[£]Head [5]Det]Mod [notes]Head]NP
f.   [[The]Det [three]Det [German]Mod [students]Head [[in] [[this]Det [class]Head]NP]PP]NP

In the same sphere, Langacker (1991: 148) finds the following nominals to constitute deviations from the prototype:

(3)
a. Money, two pictures, rejected options
b. John Smith, I, she
c. Your turning up surprised me
d. The fact that the Prime Minister has not resigned after the bribe scandal will not help the party in the next election
e. All the students, both lawyers

Indeed, some nominals, like those in (3.a), lack grounding; other nominals, such as *John Smith* in (3.b), are construed as designating a single instance of a type; the case is the same with the personal pronouns *I* and *she* in (3.b): they profile a single instance of a type; complex nominals resulting from nominalization, like (3.c), or apposition, as in (3.d) also diverge from the nominal prototype. Finally, nominals containing *both* or *all*, like those in (3.e), share with complex nominals originating in apposition the characteristic of the double grounding.

For the purposes of our research, Langacker's proposal has three drawbacks: firstly, no further distinction is drawn among nominals realized by a single element, which, in our opinion, may not show the same degree of prototypicality; secondly, no attention is paid to some nominal realizations, such as dummies; and, thirdly, syntactic complexity plays a minor role since no difference is made, for instance, between relatively simple instances like (2.a) and much more complex ones, such as (2.f).

3. SEMANTICS AND SYNTAX IN THE DEFINITION OF THE NOMINAL PROTOTYPE

At this point of the discussion, a word of caution is necessary: Langacker (1991) speaks of *nominal constructions*, rather than *subjects*; this is a significant point of divergence with respect to Langacker's proposal that must be borne in mind when reading the remainder of this paper: we take the subject to be the prototypical syntactic function of the noun phrase. In our opinion, this remark can be gathered from intransitive constructions, in which the only noun phrase that is present bears the grammatical function subject. The semantic and syntactic considerations that follow also point in this direction.

Dik (1989: 101) and Downing and Locke (1992: 153) share a very restricted view of the notion of agentivity, which is prototypically associated with the subject. For these authors, who, according to Siewierska (1991: 91) come very close to Van Valin and LaPolla (1997), the agent is one of the functions that can be borne by the only argument of one-place verbal predicates like *die,* and by the most central argument of more-than-one-place verbal predicates like *buy*. The following examples illustrate Downing and Locke's inventory of the semantic participants that can be coded by the subject in English[2]:

(4)
a.	Agent:	Jack slammed the door
b.	Force:	The flood blocked the tunnel
c.	Causative Agent:	The workers narrowed the lane
d.	Affected:	The building collapsed
e.	Experiencer:	Tim watched the kidnappers
f.	Phenomenon:	The news caused much concern
g.	Relator:	The plan was a success
h.	Possessed:	That camera is mine
i.	Sayer:	Who said so?
j.	Existent:	There is a book on the shelf
k:	Recipient:	The public were sent the tickets

For the sake of the argument, Downing and Locke's semantic function inventory might be simplified along the following lines: their Agent, Causative Agent and Sayer share the features [+human] and [+control] and could therefore be unified under the Agent function; Downing and Locke's Force and Phenomenon could also be subsumed under the former function, since both are [-human], and [+intent] but [-control]; Affected and Experiencer resist any kind of simplification since the Experiencer, which is [+human] can be either volitional (*Tim watched the kidnappers*) or non-volitional (*Tim saw the kidnappers*), whereas the Affected is [-human] and [-control]; finally, the participants Existent and Recipient might be dispensed with because they seem to be directly motivated by -and exclusively associated with- syntactic phenomena like existential and copulative constructions. Summarizing, a simplified framework for the first argument could rely on the process participants Agent, Force, Affected and Experiencer. Such framework would not be very far from Givón's (1984: 87), who speaks of three semantic roles as subject candidates: the Agent, the Dative and the Patient. Agents and datives occupy the top position of the noun hierarchy Entity>Temporal>Concrete> Animate>Human (Givón 1984: 56): the agent is prototypically a human participant who deliberately initiates the action, whereas the dative, being a human participant as well, does not hold responsible for the action. Patients occupy the bottom position of the noun hierarchy: they are typically inanimate and thus non-human and undergo the process initiated by the Agent. Let us set an example of each:

(5)
Agent:	The boss has dismissed me
Dative:	I don't know the answer
Patient of state:	The bike is in the garden
Patient of change:	The sofa has faded

These examples give us a clue for interpreting the noun hierarchy in another way: agents are morphologically more marked than datives and datives, in turn, more marked

than patients. From the syntactic perspective, however, the assignment of the subject to the agent is less marked than the assignment of this grammatical function to the patient.

As we have remarked above, Dik's view of agentivity is, unlike Givón's, very restricted. For Dik (1989: 101; 1997: 12), the Agent is, exclusively, the entity that controls an Action. Other semantic functions typically associated with the First Argument are: Positioner (the entity that controls a Position), Force (the entity that originates a Process), Processed (the entity that undergoes a Process), and Zero (the entity associated with a State). The following instances illustrate these semantic functions:

(6)
a. **Agent**:   Charlie was washing up
b. **Positioner**:   John was lying on the bed
c. **Force**:   The wind opened the gate
d. **Processed**:   The ship sank
e. **Zero**:   The car is in the garage

The parameter [± control] draws a distinction between the Agent, which is [+control], and the Forcer, which qualifies as [-control]. The parameter [± intent] distinguishes the function Force, which gives rise to an instigated process, from the function Processed, where no instigation applies.

These three proposals bear on the idea that certain participants are more directly associated with the grammatical function subject. In other words, the functionalist field has favoured a semantic approach to the definition of subject prototypicality and has stated, for instance, that the Agent participant makes a more prototypical subject than the Zero participant. However, as the methodological underpinnings we have given in the introduction to this paper reflect, we opt for an approach that associates meaning and structure and, consequently, does not exclude the latter. As is explained more fully in Martín Arista (1998a), the categories that realize the syntactic function subject do contribute to the definition of this linguistic phenomenon. The English subject can be realized by various phrases and clauses (Downing and Locke 1992: 34, among others): nominal phrases, adjectival phrases, adverbial phrases and prepositional phrases and nominal clauses. This is illustrated by (7):

(7)
a.   The picture was sold yesterday
b.   The very wealthy should pay more taxes
c.   Upstairs is the ideal place
d.   At twelve may turn out the only option
e.   That you showed up made Jim very angry

Even simple realizations like the ones in (7) allow us to speak of a lexical and syntactic category hierarchy. Such a hierarchy would be motivated by the following claims:

(i) It seems beyond a doubt that noun phrases are more prototypical subjects than noun clauses: not in vain is reference the cross-linguistically attested function of the noun; and predication the prototypical pragmatic function of the clause (Croft 1991: 99). In other words, prototypical predications are about nouns, not about other predications. This is confirmed by the fact that noun clauses cannot occur as subjects in interrogatives, as the following example shows (Huddleston 1989: 64):

(8)
a. That John was late annoyed them
b. *Did that John was late annoy them?

(ii) The nominal realization is more prototypical than adjectival, adverbial or prepositional realizations. As regards the adjectival realization, we take this one to involve a metaphorical extension of the characteristics coded by the adjective onto the being prototypically referred to by the noun:

(9)
a. The unemployed should not be given the cold shoulder
b. The elderly play a central role in national politics

(iii) When one has to decide on prepositional phrases, one may proceed in a syntactically oriented way: noun phrases are more prototypical realizations of the English subject than prepositional phrases, because noun phrases are structurally simple, whereas prepositional phrases are structurally complex (in fact, they consist of a noun phrase plus a prepositional element governing the presence of the noun phrase):

(10)
a. [The house]NP
b. [[In] [the house]NP]PP

(iv) Some remarks are in order as regards adverbial phrases. We may reach a conclusion on the degree of prototypicality of different adverbial classes on the grounds of the structure of the clause: the lower the level of the hierarchical structure of the clause (Dik 1989; Hengeveld 1990) to which an adverbial class is usually attached, the more prototypical the realization of the subject as a given adverbial class is. For example, adverbials of Manner, like the one in bolds in (11.a), or Direction, attached to the level of the predicate, give rise to more prototypical subjects than adverbials of Place or Time, like the one in

(11.b), which are associated with the layer of the predication; adverbs that realize predication adverbials, in turn, constitute more prototypical subjects than proposition adverbs, such as that in (11.c):

(11)
a. In this case quietly implies some sort of abduction
b. Yesterday was not the right moment for resigning
c. Certainly causes some difficulties in this discussion

Notice that no difference in terms of clause layer can be drawn between (11.c) and (11.a): the attachment to the proposition or the illocution level can hardly be recovered from expression like these ones. Put another way, illocution adverbs cannot show up in subject position, which reinforces our view of the prototypicality of adverbial subjects in terms of their belonging in a certain clause layer. Moreover, as example (12) makes it clear, the reference potential of these subjects is severely constrained, which is iconically reflected by the fact that they can hardly collocate with a verb different from the copula[3]. Again, this fact supports our approach to this question on the basis of clause layers: the higher in the clause layers we move, the more difficult it is for an adverb to collocate as the subject with a non-copular verb, as in (12):

(12)
a. Here honestly is the same as candidly
b. ?[The word] honestly raises another discussion
c. *Honestly came as a surprise

Indeed, the higher up the layers of the clause we go, adverbial subjects become more utterance independent, and thus denotative if we follow Lyons' (1995: 79) terminology; whereas prototypical subjects are utterance dependent, that is, referential[4].

(v) As Martín Arista (1998b) remarks, there are formal and functional arguments for grouping the extraposition of the nominal segment of dummy *it* and dummy *there* constructions on the one hand and the extraposition of the postmodifier on the other. Both categories represent a non-prototypical realization of the subject, being syntactically discontinuous and, in the case of dummies, double-headed. Three cases in point are presented below:

(13)
a. There is somebody sleeping in my car
b. It is obvious that he will stay
c. A much more exciting night started than I had expected

(vi) Finally, nothing has been said so far concerning complex subjects. Our intuition is that complex nominals give rise to less prototypical subjects than simple nominals. Suffice it so give these examples of a copulative subject, an adversative subject, a distributive subject and a distributive negative subject, respectively:

(14)
a. **Jimbo and Sean** won the prize
b. **Not Jimbo but Sean** won the prize
c. **Either Jimbo or Sean** won the prize
d. **Neither Jimbo nor Sean** won the prize

We close this section by summarizing these judgements on subject prototypicality in the following hierarchy, which will be tested against quantitative evidence in the following section:

(15) **The hierarchy of subject prototypicality**:
Nominal>Pronominal>Complex>Dummy>Clausal>Adjectival> Adverbial>Prepositional

This hierarchy is interpreted in the following terms: a subject realized by a nominal noun phrase constitutes a more prototypical realization of the syntactic function subject than a subject realized by a noun phrase whose head is a pronoun; nominal and pronominal subjects are more prototypical than complex, dummy and clausal subjects, which, in turn, represent more prototypical realizations of this grammatical function than adjectival, adverbial and prepositional subjects. In sum, nominal subjects are the most prototypical and prepositional subjects give rise to the least prototypical realization of this syntactic function in English.

4. INTRODUCING QUANTITATIVE DATA

The quantitative evidence we provide in this study of the realization of the English subject has been drawn from *The Guardian* (Friday November 21, 1997, pages 1-3); Peter Newmark's *A Textbook of Translation* (New York: Prentice Hall, 1988, pages 3-12); Tom Sharpe's *Grandchester Grind* (London: Pan Books, 1995, pages 1-9); and the technical texts in Louis Trimble's *English for Science and Technology* (Cambridge: Cambridge University Press, 1989, pages 16-56). We have taken these sources to represent, respectively, the general, academic, literary and technical registers of English. We have examined 250 examples of each register, which amounts to 1,000. The examples have been analyzed in the light of the variables contained in the hierarchy of subject prototypicality given in (15).

A glimpse at these examples yields the data that are summarized in Table 1: simple subjects constitute 92% of the total. Within simple subjects, the number of instances of adjectival, adverbial, prepositional and non-finite clause subjects is irrelevant: these four categories together represent 0.3% of the total amount. Within other subject realizations, the number of occurrences of dummy subjects and coordinated subjects is similar: 3.5% versus 3.7% respectively. Finally, clausal subjects amount to 0.8%.

REALIZATION	EXAMPLES	PERCENTAGE
PRONOMINAL	43O	43%
NOMINAL	487	48.7%
ADJECTIVAL	10	0.1%
ADVERBIAL	10	0.1%
PREPOSITIONAL	0	0%
NON-FINITE CLAUSE SUBJECT	10	0.1%
SIMPLE SUBJECT TOTAL	**920**	**92%**
COMPLEX SUBJECT	37	3.7
DUMMY SUBJECT	35	3.5
CLAUSAL SUBJECT	8	0.8
OTHER REALIZATIONS TOTAL	**80**	**8%**
TOTAL	**1000**	**100%**

Table 1: *Data summary*

Although the variable independent/dependent clause was not initially included within the hierarchy given in (15), some data provided by the analysis of this parameter are worth considering.

REALIZATION	EXAMPLES		PERCENTAGE	
	INDEP	DEP	INDEP	DEP
PRONOMINAL	257	173	25.7%	17.3%
NOMINAL	429	58	42.9%	5.8%
ADJECTIVAL	1	0	0.1%	0%
ADVERBIAL	1	0	0.1%	0%

PREPOSITIONAL0	0	0%		0%
NON-FINITE CLAUSE SUBJECT	-	1	-	0.1%
SIMPLE SUBJECT TOTAL	**688**	**232**	**68.8%**	**23.2%**
COMPLEX SUBJECT	29	8	2.9%	0.8%
DUMMY SUBJECT	32	3	3.2%	0.3%
CLAUSAL SUBJECT	7	1	0.7%	0.1%
OTHER REALIZATIONS TOTAL	**68**	**12**	**6.8%**	**1.2%**
TOTAL	**756**	**244**	**75.6%**	**24.4%**

Table 2: *Independent vs. dependent clauses (absolute)*

Even though an absolute analysis like the one offered in Table 2 shows that the number of different realizations of the subject does not vary substantially when one moves from independent to dependent clauses, a relative analysis, such as the one given in Table 3, is far more revealing. To start with, the relative analysis makes it clear that the pronominal realization of the subject in dependent clauses (70.9%) doubles the percentage of pronominal subjects in independent clauses (33.9%). Conversely, nominal subjects are far more frequent in independent clauses (56.7%) than in dependent clauses (23.7%). Simple subjects, being the general rule, are even more abundant in dependent clauses (95%, versus 90.8% in independent clauses). Clausal subjects are equally scarce in independent and in dependent clauses. Finally, whereas the number of coordinated subjects is similar in both categories of clauses, dummy subjects seem to occur more often in independent clauses (4.2%) than in dependent ones (1.2%).

REALIZATION	PERCENTAGE	
	INDEPENDENT	**DEPENDENT**
PRONOMINAL	33.9%	70.9%
NOMINAL	56.7%	23.7%
ADJECTIVAL	0.1%	0%
ADVERBIAL	0.1%	0%
PREPOSITIONAL	0%	0%
NON-FINITE CLAUSE SUBJECT	-	0.4%
SIMPLE SUBJECT TOTAL	**90.8**	**95**

COMPLEX SUBJECT	3.8%	3.2%
DUMMY SUBJECT	4.2%	1.2%
CLAUSAL SUBJECT	0.9	0.4%
OTHER REALIZATIONS TOTAL	**8.9**	**4.8**
TOTAL	**99.7%**	**99.8%**

Table 3: *Independent vs. dependent clauses (relative)*

Among pronominal subjects, third person personal pronouns (including the singular and plural number) constitute a vast majority (38.3%). The first person personal pronoun follows in the ranking (22%). In dependent clauses, the realization of the subject by means of a relative pronoun stands out clearly (17.4%). In relative terms, the relative pronoun realization of the subject in dependent clauses totals 43.2% of pronominal subjects, a significant percentage. Other categories are marginal, with, perhaps, the exception of second person pronouns (9.7%). These data are displayed in Table 4:

REALIZATION	PERCENTAGE
FIRST PERSON	22%
SECOND PERSON	9.7%
THIRD PERSON	38.3%
DEMONSTRATIVE	4.1%
POSSESSIVE	0%
INDEFINITE	4.8%
RELATIVE (CLAUSAL)	1.3%
RELATIVE (DEPENDENT)	17.4%
INTERROGATIVE	1.8%
TOTAL	**99.4%**

Table 4: *Pronominal subjects (relative)*

Most nominal subjects in the corpus qualify as morphologically unmarked nouns (both in the singular and the plural). Nevertheless, the percentage of noun realizations (26.6%) does not stick out clearly with respect to the percentage of determiner-head realizations of the subject (22.3%). What is beyond a doubt is that these two realization together amount to practically one half of the total (48.9%). The determiner-head and the

determiner-modifier-head realizations together total 25%. The other fourth is broken down into several categories of postmodification: discontinuous postmodifier (0.4%), modifier-head-postmodifier (1%), determiner-modifier-head-postmodifier (3.4%), head-postmodifier (5.3%) and determiner-head-postmodifier (14.9%). It is probably worth considering that premodification does not seem to be favored by these data with respect to postmodification. These data are gathered in Table 5:

REALIZATION	PERCENTAGE
NOUN (UNMARKED)	26.6%
NOUN (MARKED)	0.2%
DETERMINER-HEAD	22.3%
MODIFIER-HEAD	12.5%
DETERMINER-MODIFIER-HEAD	12.5%
HEAD-POSTMODIFIER	5.3%
DETERMINER-HEAD-POSTMODIFIER	14.9%
MODIFIER-HEAD-POSTMODIFIER	1%
DETERMINER-MODIFIER-HEAD-POSTMODIFIER	3.4%
DISCONTINUOUS POSTMODIFIER	0.4%
TOTAL	**99.1%**

Table 5: *Nominal subjects (relative)*

No distinction has been drawn in this section as regards register. As we have remarked above, the percentages rendered in these tables represent the average. For the percentages in terms of the four registers considered we refer the reader to the appendix.

5. DISCUSSION

It is a well-known fact that it was Trubetzkoy in his article "Die phonologischen Systeme" (1931) who coined the term *Merkmalhaltigkeit*, the property of having a certain mark. In the 1940's and 1950's, Jakobson expanded the concept of markedness not only to accommodate morphological and semantic features but also to deal with non-binary oppositions[5]. Current work in markedness theory revolves around two key ideas:

(i) Markedness delimits some categories that are more accessible than others and thus encodes a preference for language acquisition. This is the line that the Chomskian paradigm has taken. Indeed, as Battistella (1996: 121) has pointed out, the Principles and Parameters theory has identified, more or less

consistently, the unmarked options with core grammar and the marked options with the periphery.

(ii) The functional field has opted for relating prototypes and hierarchies to markedness evidence. In this line, Croft (1990: 124; 1991: 56) has stated that the unmarked form represents a cluster of grammatical values on a number of parameters, given that grammatical categories seldom occur in isolation. This cluster represents the prototype category. A hierarchy of prototypicality links prototypical categories to less-prototypical categories and non-prototypical categories. This is the kind of approach that we adopt in this paper.

With respect to the criteria for determining markedness, Croft (1990, 1991) distinguishes the frequency criterion, the structural criterion, the behavioral criterion and the distributional criterion. For the purposes of this discussion, we will consider that occurrence (the frequency criterion) and minimal marking (the structural criterion) run parallel. This parallelism constitutes, in fact, one of the main sources of external motivation for a functional grammar: iconicity. We will take iconicity to provide non-quantitative evidence in markedness judgements[6]. We will add Givón's (1995: 28) substantive markedness to the aforementioned criteria.

Substantive markedness, which may be re-labeled *cognitive markedness*, accounts for Langacker's (1991: 318) prototypicality choices. Langacker regards the grammatical accessibility of subjects as a symptom of their conceptual prominence. In other words, this grammatical function is realized by more categories than, say, the object, which is less conceptually-prominent. As Croft (1990: 71) puts it, syncretization (morphological and categorial syncretization, in this case) coincides with unmarkedness. In the same vein, Langacker (1991: 143) puts forward that the prototypical nominal incorporates a head and a determiner since the semantic relationship between these two components of the nominal is intrinsic while the semantic relationship between the head and the modifier is extrinsic, and largely dependent on the kind of mental contact that the speech act participants have established between these subconstituents. While the realization determiner-head is unmarked with respect to determiner-modifier-head on the grounds of cognitive and structural criteria, the quantitative analysis shows that the realization determiner-modifier-head (or determiner-head-postmodifier) is not textually marked (the frequency criterion) with respect to the sequence determiner-head. Moreover, the plain noun occurs with a similar frequence.

The issue of pronominal subjects hangs in the balance: on the one hand, the frequency criterion reveals that third person pronominal subjects are unmarked with respect to first person pronominal subjects; on the other hand, the substantive criterion seems to indicate that both first and third person are non-prototypical because they profile a single instance of a type (Langacker 1991: 143). All in all, Langacker remarks that the prototypical expression makes reference to a third person participant within a given setting.

Even though Langacker does not consider nominal prototypicality from the angle of clause dependency, the quantitative analysis of this variable has proved very telling: the pronominal realization of the subject of dependent clauses is unmarked with respect to the nominal realization. In independent clauses, on the contrary, the nominal realization seems to be, from a quantitative perspective, the unmarked option (this is especially the case with relative pronouns). These statements come in the wake of Givón's (1993 vol. II: 334) subject continuity, which is defined as *reduced subject marking*.

Simple subjects represent the unmarked option both in dependent and independent clauses (but far more clearly in the former). Our quantitative approach corroborates Langacker's cognitive judgement that complex subjects are not prototypical. In this respect, it is worth noticing that the three criteria of markedness point in the same direction. Our quantitative analysis has also revealed that dummy and clausal subjects are textually marked with respect to complex subjects (which was also stated in qualitative terms in Martín Arista (1998b)). This fact can be put down to the existence of syntactic discontinuity in dummies, unlike complex subjects, whose subconstituents occur in an adjacent fashion. Adjectival, adverbial and prepositional subjects, which involve the same kind of metaphorical extension as clausal subjects and thus imply complex cognitive processes, are textually and cognitively marked, which has been demonstrated by our quantitative analysis.

6. Summary and conclusions

The discussion above stresses the necessity of enlarging the corpus of analysis so that the subject realizations that show the lowest frequency of occurrence can be adequately dealt with. As regards the variables used in this analysis, the ones we have used have sufficed, but it would not be out of place to consider the variable Saxon/Norman genitive in the discussion of modification in future research.

With reference the hierarchy of subject prototypicality advanced in (15), our quantitative analysis has proved it valid, although, at the same time, it has shown that the variable independent/dependent is of great relevance. For this reason, the hierarchy of subject prototypicality can be re-interpreted for dependent clauses as follows: relative pronoun>other pronouns>noun>complex>dummy>clausal. In the case of independent clauses, the hierarchy given in (15) may be re-written in the following terms: simple>complex>dummy>clausal. Other realizations do not seem to be frequent enough to be worth considering (at least until the study of a larger corpus throws light on this respect).

At the theoretical level, a quantitative approach to subject prototypicality focuses on the subtleties of markedness and, more specifically, on the intricacies of frequency markedness and prototypicality judgements on the one hand and substantive markedness on the other. We have relied on a kind of analysis that links syntactic structure and categorization to prototypicality theory via markedness considerations. Our point is that

the methodology used in this paper contributes to the integration of language structure and language meaning, including the aspects normally included under the heading *cognitive*.

NOTES

1. See Kalisz and Kubinski (1997) for a more detailed comparison between the explanatory aims of Cognitive Grammar and Functional Grammar.
2. See Halliday (1985: 131) for a slightly different inventory. Foley (1993) and Van Valin and Wilkins (1996) are also worth seeing with reference to the Actor, the Undergoer and the Effector.
3. But see Goossens (1992) on predicationality and the copula.
4. See Dik (1997: 103) on propositional and predicational terms.
5. For more detailed information on the history of markedness theory, see Andrews (1990).
6. See Haiman (1985: 73), Östman (1989: 158) and Croft (1990: 165) on this concept.

REFERENCES

Andrews, E. 1990. *Markedness Theory*. Durham: Duke University Press.

Battistella, E. 1996. *The Logic of Markedness*. Oxford: Oxford University Press.

Croft, W. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

Croft, W. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: Chicago University Press.

Dik, S. 1986. "On the notion 'functional explanation'". *Belgian Journal of Linguistics* 1: 11-52.

Dik, S. 1989. *The Theory of Functional Grammar I: The Structure of the Clause*. Dordrecht: Foris.

Dik, S. 1997. *The Theory of Functional Grammar II: Complex and Derived Constructions*. Berlin: Mouton de Gruyter.

Downing, A. and P. Locke. 1992. *A University Course in English Grammar*. Hemel Hempstead: Prentice Hall International.

Foley, W. 1993. "The conceptual basis of grammatical relations". *The Role of Theory in Language Description*. Ed W. Foley. Berlin: Mouton de Gruyter.131-174.

Foley, W. ed. 1993. *The Role of Theory in Language Description*. Berlin: Mouton de Gruyter.

Givón, T. 1984/1990. *Syntax: A Functional-Typological Introduction* (2 vols.). Amsterdam: John Benjamins.

Givón, T. 1989. *Mind, Code and Context*. New York: Erlbaum.

Givón, T. 1993. *English Grammar: A Function-Based Introduction* (2 vols.). Amsterdam/Philadelphia: John Benjamins.

Givón, T. 1995. *Functionalism and Grammar*. Amsterdam/Philadelphia: John Benjamins.

Goossens, L. 1992. "Graded predicationality, semanticization and *be* in a functional grammar of English". *Belgian Journal of Linguistics* 7. 57-73.

Haiman, J. 1985. "Symmetry". *Iconicity in Syntax*. Ed J. Haiman. Amsterdam: John Benjamins. 73-95.

Haiman, J. ed. 1985. *Iconicity in Syntax*. Amsterdam: John Benjamins.

Halliday, M. 1985. *An Introduction to Functional Grammar*. London: Arnold.

Hengeveld, K. 1990. "The hierarchical structure of utterances". *Layers and Levels of Representation in Language Theory*. Eds J. Nuyts and A.M. Bolkestein. Amsterdam: John Benjamins. 1-24.

Huddleston, R. 1989. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.

Kalisz, R. and W. Kubinski. 1997. "Functional Grammar vs. Cognitive Linguistics: an attempt at discovering similarities and differences". *Working Papers in Functional Grammar* 64. Amsterdam: University of Amsterdam.

Keenan, E.L. 1976. "Towards a Universal Definition of Subject". *Subject and Topic*. Ed Ch. Li. New York: Academic Press. 303-333.

Kuno, S. 1980. "Functional Syntax". *Syntax and Semantics 13: Current Approaches to Syntax*. Eds E. Moravcsik and J. Wirth. New York: Academic Press.117-135.

Langacker, R. 1987. *Foundations of Cognitive Grammar I: Theoretical prerequisites*. Stanford, Ca.: Stanford University Press.

Langacker, R. 1991. *Foundations of Cognitive Grammar II: Descriptive Application*. Stanford, Ca.: Stanford University Press.

Li, Ch. ed. 1976. *Subject and Topic*. New York: Academic Press.

Lyons, J. 1995. *Linguistic Semantics*. Cambridge: Cambridge University Press.

Martín Arista, J. 1998a. "Beyond coding and control: leakage, categorization and blocking in the definition of the English subject". *Miscelánea* 19. 111-138.

Martín Arista, J. 1998b. "English dummy constructions revisited: a functional approach". *Estudios Ingleses de la Universidad Complutense* 6. 65-80.

Moravcsik, E. and Wirth, J. eds. 1980. *Syntax and Semantics 13: Current Approaches to Syntax*. New York: Academic Press.

Nuyts, J. 1992. *Aspects of a Cognitive-Pragmatic Theory of Language*. Amsterdam: John Benjamins.

Nuyts, J. and A.M. Bolkestein eds. 1990. *Layers and Levels of Representation in Language Theory*. Amsterdam: John Benjamins.

Östman, J. 1989. "Testing iconicity". *Belgian Journal of Linguistics* 4. 145-163.

Shibatani, M. and S. Thompson eds. 1996. *Grammatical Constructions. Their Form and Meaning*. Oxford: Oxford University Press.

Siewierska, A. 1991. *Functional Grammar*. London: Routledge.

Taylor, J. 1989. *Linguistic categorization*. Oxford: Clarendon Press.

Van Valin, R. and D. Wilkins. 1996. "The Case for 'Effector': Case Roles, Agents and Agency Revisited". *Grammatical Constructions. Their Form and Meaning*. Eds M. Shibatani and S. Thompson. Oxford: Oxford University Press. 289-322.

Van Valin, R. and R. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge University Press.

APPENDIX: REGISTER PERCENTAGE (GENERAL, ACADEMIC, TECHNICAL AND LITERARY)					
	GENERAL	ACADEMIC	TECHNICAL	LITERARY	AVERAGE
1. MAIN CLAUSES	%	%	%	%	%
1.1. SIMPLE SUBJECT					
1.1.1. PRONOMINAL SUBJECT					
1.1.1.1. FIRST PERSON	4	5.2	5.6	16	7.7
1.1.1.2. SECOND PERSON	2.4	4.4	0	4.8	2.9
1.1.1.3. THIRD PERSON	12	9.2	5.2	18.8	11.3
1.1.1.4. DEMONSTRATIVE	2.4	0.8	2	0	1.3
1.1.1.5. POSSESSIVE	0	0	0	0	0
1.1.1.6. INDEFINITE	1.2	0.4	1.6	1.6	1.2
1.1.1.7. RELATIVE (CLAUSAL)	0	1.2	0	1.2	0.6
1.1.1.8. INTERROGATIVE	0	1.2	0	1.6	0.7
1.1.2. NOMINAL SUBJECT					
1.2.1.1. NOUN					
1.2.1.1.1. UNMARKED	14.4	4.8	8.4	16	10.9
1.2.1.1.2. MARKED (GEN.)	0	0	0	0	0
1.2.1.2. DET-HEAD	12	13.6	6.8	2.8	9.4
1.2.1.3. MOD-HEAD	5.2	7.2	9.2	0.8	5.6
1.2.1.4. DET-MOD-HEAD	6	3.2	12	0.4	5.4
1.2.1.5. HEAD-PPOSTMOD	3.6	2.8	2	0.8	2.3
1.2.1.6. DET-HEAD-POSTMOD	2.8	6.4	14.8	3.6	6.9
1.2.1.7. MOD-HEAD-POSTMOD	1.2	0	0.4	0	0.4
1.2.1.8. DET-MOD-HEAD-POSTMOD	0.8	1.2	4.4	0	1.6
1.2.1.9. DISCONTINUOUS POSTMOD	0.4	0.4	0	0	0.2
1.1.3. ADJECTIVAL SUBJECT	0.4	0	0	0	0.1
1.1.4. ADVERBIAL SUBJECT	0.4	0	0	0	0.1
1.1.5. PREPOSITIONAL SUBJECT	0	0	0	0	0
1.2. COMPLEX SUBJECT					
1.2.1. COPULATIVE: A AND B	1.6	3.6	4	0.8	2.5
1.2.2. ADVERSATIVE: NOT A BUT B	0	0	0.4	0	0.1
1.2.3. DISTRIBUTIVE: EITHER A OR B	0	0.4	0.8	0	0.3
1.2.4. DISTR. NEG.: NEITHER A NOR B	0	0	0	0	0

	GENERAL	ACADEMIC	TECHNICAL	LITERARY	AVERAGE
1.3. DUMMY SUBJECT					
1.3.1. DUMMY THERE	4	2.4	0.8	0.4	1.9
1.3.2. DUMMY IT	0.8	0.4	0	4	1.3
1.4. CLAUSAL SUBJECT					
1.4.1. FINITE CLAUSE	0.4	0.4	0	0	0.2
1.4.2. NON-FINITE CLAUSE	0	0.8	1.2	0	0.5
2. DEPENDET CLAUSES					
2.1. FINITE CLAUSES					
2.1.1. SIMPLE SUBJECT					
2.1.1.1. PRONOMINAL SUBJECT					
2.1.1.1.1. FIRST PERSON	2.4	1.2	0	3.6	1.8
2.1.1.1.2. SECOND PERSON	0.8	1.6	0	2.8	1.3
2.1.1.1.3. THIRD PERSON	3.6	5.2	4.8	7.2	5.2
2.1.1.1.4. DEMONSTRATIVE	0.8	1.2	0	0	0.5
2.1.1.1.5. POSSESSIVE	0	0	0	0	0
2.1.1.1.6. RELATIVE	6.8	9.6	11.2	2.4	7.5
2.1.1.1.7. INDEFINITE	0	3.2	0	0.4	0.9
2.1.1.1.8. INTERROGATIVE	0	0.4	0	0	0.1
2.1.1.2. NOMINAL SUBJECT					
2.1.1.2.1. NOUN					
2.1.1.2.1.1. UNMARKED	2.8	0.8	0.8	4	2.1
2.1.1.2.1.2. MARKED (GEN.)	0.4	0	0	0	0.1
2.1.1.2.2. DET-HEAD	0.8	0.8	0.8	3.6	1.5
2.1.1.2.3. MOD-HEAD	0.4	0.4	1.2	0	0.5
2.1.1.2.4. DET-MOD-HEAD	0.8	0.8	0.4	0.8	0.7
2.1.1.2.5. HEAD-POSTMOD	0.8	0.4	0	0	0.3
2.1.1.2.6. DET-HEAD-POSTMOD	0.8	0	0.4	0.4	0.4
2.1.1.2.7. MOD-HEAD-POSTMOD	0.4	0	0	0	0.1
2.1.1.2.8. DET-MOD-HEAD-POSTMOD	0.4	0	0	0	0.1
2.1.1.2.9. DISCONTINUOUS POSTMOD	0	0	0	0	0
2.1.1.3. ADJECTIVAL SUBJECT	0	0	0	0	0
2.1.1.4. ADVERBIAL SUBJECT	0	0	0	0	0

	GENERAL	ACADEMIC	TECHNICAL	LITERARY	AVERAGE
2.1.1.5. PREPOSITIONAL SUBJECT	0	0	0	0	0
2.1.2. COMPLEX SUBJECT					
2.1.2.1. COPULATIVE: A AND B	1.6	0.4	0.4	0.8	0.8
2.1.2.2. ADVERSATIVE: NOT A BUT B	0	0	0	0	0
2.1.2.3. DISTRIBUTIVE EITHER A OR B	0	0	0	0	0
2.1.2.4. DISTR. NEG.: NEITHER A NOR B	0	0	0	0	0
2.1.3. DUMMY SUBJECT					
2.1.3.1. DUMMY THERE	0	0.4	0	0.4	0.2
2.1.3.2. DUMMY IT	0	0	0.4	0	0.1
2.1.4. CLAUSAL SUBJECT					
2.1.4.1. FINITE CLAUSE	0	0	0	0	0
2.1.4.2. NON-FINITE CLAUSE	0	0.4	0	0	0.1
2.2. NON-FINITE CLAUSES					
2.2.1. UNMARKED (ACCUSATIVE)	0.4	0	0	0	0.1
2.2.2. UNMARKED(GENITIVE)	0	0	0	0	0
2.2.3. MARKED (INTRODUCED BY FOR)	0	0	0	0	0
TOTAL	100	100	100	100	100