University of Venice

WORKING PAPERS IN LINGUISTICS

Vol. 5, n. 1

1995

edited by Laura Brugè

Centro Linguistico Interfacoltà
Università Ca' Foscari di Venezia
S.Croce 2161 - 30135 Venezia
Tel. 39-41-5241642, Fax 39-41-718259

Table of Contents:

Switching from Narrative to Legal Genre

Rodolfo Delmonte and Denise Dibattista
University of Venice

The paper presents a linguistically based approach to text understanding related to texts belonging to legal genre and highlighting some of their properties in the area of semantic and inferential processes. In particular, referential properties of general nouns are discussed and shown at work in an excerpt taken from a EC directive. Proposals to deal with the same problem by Strzalkowski, Cercone, Woods, Brachman are presented, discussed and compared. Our approach is embodied in a system called GETA_RUN presented at various international conferences and freely available from our ftp site.

1. *Introduction*

This paper deals with the analysis of written texts in legal genre with the aim to highlight linguistic and cognitive differences in its surface and deep form from narrative genre.

Generally speaking, narrative texts deal with entities actually existing and events which actually took place in the real world - disregarding for the moment fairy tales and science fiction stories. In the analysis of narrative texts, the backbone is represented by a temporal sequence or timeline to which the various events making

University of Venice
Working Papers in Linguistics
vol. 5, n.1; 1995

up the story are related, and a limited number of topics introduced and then referred to during the text which attract the focus of attention of reader or listener.

In legal texts, there is neither an internal plot, nor a main protagonist to be highlighted, nor a temporal progression. Events are all related to a hypothetical instantiation of some (pre)condition which is the obligational body of the legal argumentation, usually also addressed by the title or the heading of the bill or directive, and there are definitions which express it in more detail.

As to entities, legal texts are characterized by the implicit presence of a universal quantifier which determines the interpretation of referential expression in terms of classes of individuals rather than sets or single individuals. For instance, the council directive we have used to elaborate the computational model concerns the liability of producers for defective products. However, as we shall see in more detail below, our ontology is made up of three types of entities: generic entities, classes of individuals, individuals and sets.

We shall not be dealing with the specificity of the legal lexicon, which however requires a specialized terminology or in which the common vocabulary is used with a different meanng from narrative texts.

Switches are placed in the system at all level of computation and we shall discuss them one by one in relation to each level of representation.

2. *Syntactic switches*

A switch at syntactic level will typically regard the peculiarities of legal sentential structures, which might affect the way our parser takes decisions, for instance, when in presence of conjunctions or punctuation marks; in presence of constituents which might be regarded both as complements or as adjuncts, like PPs, but when appearing in a given position in legal texts, they can only be interpreted in a given way; pecualirities of postmodification, when for instance an implicit participial adjunct is coordinated with a PP, and so on.

Generally speaking, legal texts are more complex from a structural point of view than narrative texts. Sentences tend to be very long - in particular definitional ones -

and very heavy stylistically. This is usually due to the need to express complex concepts which are better expressed when connection are kept within the same utterance than when they are expanded at discourse level, intersententially.

Here below is a list of typical problems at a structural level which require a specialized set of rules, activated by a switch:

a. Enumerations or NP lists;
b. PP adjuncts in between auxiliary/modal and main verb;
c. Double subordination;
d. Shall as modality operator;
e. Lack of 'that' complementizer deletion that might induce garden paths with NP structure;
f. High number of passive constructions with a systematic omission of the agent that contrary to what happens in narrative texts need not be introduced semantically; also there is a heavy use of passive constructions coupled with modal operators;
g. Premodifiers with restrictive import, rather than as free adjuncts;
h. Ellipsis and use of other implicit subordinate structures like absolute constructions or generic participials;
i. Postmodification with restrictive relative clauses and participials;
j. Parataxis.

At syntactic level, clauses and sentences are connected with a preferred tendency towards parataxis. However, also hypotactic relations are commonly used whenever condensation and precision are the main purposes.
Therefore, in a legislative text, the scarce occurrence of connective elements usually coexist with multiple subordinate clauses or sentences introduced in a position unusual in ordinary language.

The use of paratactic structures is functional to legislative language, which tends to explicitness and exhaustiveness whenever possible.
Nevertheless, the repetition or juxtaposition of elements may lead to long, sometimes redundant, sentences. Actually, it has been observed that not always the desirability of simplified legislative language is obtained through the use of short sentences.

The problems of semantic connections between sentences, and the necessity of syntactic condensation of legislative texts, are solved, to a great extent, by means of highly nominal constructions, reduced clauses, nominalizations, passives with deleted agents, and embedded sentences.

2.1. *Reduced clauses*

An important syntactic feature is the prevalence of complex nominal phrases with a high occurrence of postmodification.

Restrictive relative clause is the more frequent case of postnominal modification. This is justified by its function of defining or reducing the meaning of the antecedent noun phrase, a further step towards precision and clarity required by legal genre.
For such reasons, restrictive relative clauses are recurrent in legal definitions whenever the legislator intends to limit the class of individuals to whom a rule applies. Furthermore, in legislative texts, these kind of clauses are generally correspondent to conditional sentences:

(1) Any person who imports a product for sale shall be responsible as a producer.

(2) If a person imports a product for sale he shall be responsible as a producer.

It is evident that example (1) is equivalent to (2). The latter is a defining conditional, it is not used to make hypothesis or predictions, but provides the precise meaning of a term which is general and not specific. On the particular form and entailments of conditional clauses we will dwell upon below. Non-restrictive relative clauses are rare in legislative texts and used only when there is no possibility of ambiguity:

(3) The Joint Committee shall be composed on principle of twelve members, five representing the Committee of Ministers and seven representing the Consultative Assembly, the latter number to include the President of the Consultative Assembly, who shall be a member ex-officio.

Example (3) is the only instance of non-restrictive relative clause found in our texts and it is of a simple, not ambiguous interpretation, since the only antecedent of the relative pronoun may be "the President of the Consultative Assembly". As a result, these clauses are largely substituted with coordinate sentences and the repetition of the subject expression is preferred to the relative pronoun.

Postmodification is also realized through non-finite clauses, particularly present and past participles, which emphasize once more, the legislator's attention to condensation. All the instances present the antecedent head identical to the deleted subject of the non-finite verb clause: "the Treaty establishing [...]", "a proposal amending [...]", "damage resulting from [...]"; "regulations issued", "convention ratified", "the rights conferred", etc.

In the examples the *-ing construction* represents the active voice, whereas the *-ed construction* is linked with the passive voice. Passives with deleted agents actually are a common means of legal language to give thematic prominence to the element in subject position and therefore achieve depersonalization and abstractness:

(4) Before deciding upon a request mentioned under paragraph 1 above the requested body may, in view of general interests involved, hear the competent public authorities.

The use of premodification with participle is noteworthy: there is hardly any instance of premodification by present participle ("distinguishing feature" is one of the few instances present), while the examples of premodification by past participle ("a finished product", "the injured person", "the proposed measure", etc.) are more common. Since present and past participles maintain the same reference to active and passive voice as in postmodification, this prenominal forms reflect, in their syntactic use, the legislative preference for an impersonal, detached tone.

2.2. *Nominalization*

A particular case of postmodification is given by prepositional phrases that are used to qualify the head noun. Even though in other genres (narrative, instructional texts, etc.) a noun that specify a preceding phrase is commonly placed in adjective position, legal language opts for postmodifying of-phrases whenever possible:

(5) The declaration will become effective on the first day of the month following the expiration of a period of three months after the date of its reception by the Secretary General.

A remarkable occurrence of such linguistic behaviour is given by the presence of nominalization, in our case reception in example (5).

Because of its frequency, the tendency towards nominalization of legal language is considered the more relevant lexical and syntactic characteristic. It consists in conveying the functions of the verb to the noun and it is realized in lexicon through deverbal and verbal nouns.

Deverbal nouns are forms obtained by adding to the verbs particular suffixes such as: *-ation* ("presentation", "circulation", "implementation", etc.); *-ment* ("arrangement", "apportionment", etc.); *-al* ("proposal", "refusal", "dismissal", etc.); *-ing* ("stock-farming", "processing", "leasing", etc.). In legal genre deverbal nouns have only an abstract meaning, while in other genres (narrative, instructional texts, dialogues) they can be used also for concrete nouns.

It must be observed that this lexical process has clear syntactic properties: the prepositional phrase headed by "of" may correspond functionally either to the subject (e.g. "the functioning of the common market"), or to the object of the transformed verb ("protection of the consumer").

Nominalization has essentially two functions: in the first place it is, together with participial phrases, the principal means of controlling the length of the sentence and obtain condensation. It must be observed that it does not always achieve its purpose since sentences using the nominal style are frequently wordier than the corresponding verbal style.

In the second place, nominalization has a relevant consequence: a noun-like element in a sentence may be moved more freely than a verb and can be placed in subject position. With this lexical device the writer objectifies the content and makes it possible to treat it as something abstracting from specific personal references. This circumstance is in line with the tendency of legislative discourse to do away with the temporal perspective and to achieve generality and abstractness.

2.3. *Passive Voice*

Passives constructions, may be regarded as a useful means to obtain condensation and depersonalization. However the high frequency of passive occurrence (see ex. 6)) and the complex use of multiple passives, require a further justification.

(6) The provisions of Articles 1,2,3 and 4 of this Protocol shall be regarded as additional articles [...]. This Protocol shall be ratified at the same time [...]. The instruments of ratification shall be deposited [...]. Done at Paris.

(7) A declaration made in accordance with this article shall be deemed to have been made in accordance with paragraph 1 of Article 63 of the Convention.

It cannot be denied that passives with deleted agent give to legal genre its typical communicative mode, namely the impersonal and detached tone, but they also contribute to its performative function .

Austin (1962), in his established work on performatives states that a verb in the passive voice is an indubitable performative, which is usually found in formal and legal occasions. Furthermore, the "hereby" criterion, used to detect performativity, is applicable to passives. As Austin suggests, when "hereby" can be inserted, it indicates that the written utterance is the instrument affecting the act of authorising, forbidding, etc. and is therefore performative.

Although passives with deleted agent are found also in conditional clauses, here they cannot have a performative function.

(8) Any member of the Council of Europe may withdraw by formally notifying the Secretary General of its intention. Such withdrawal shall take effect at the end of the financial year in which it is notified. [...] If the notification is given in the last three months of the financial year, it shall take effect at the end of the next financial year.

In accordance with the explanation proposed by Vander Linden (1994) for instructional texts, the use of agentless passives seems to depend on whether the action has been mentioned before or not. If an action or an event has already been introduced in the previous text, it takes the present tense, agentless passive form, as in (8); otherwise a simple present, active form is used.

2.4. *Embeddings*

Embedding is a device frequently employed in legislative texts to provide condensation and precision. It consists in the insertion of a modifying phrase or clause, often introduced by a subordinating conjunction, within the clause structure. The adverbial function of embeddings is realized by finite-verb clauses, as in example (9); non-finite clauses, as in example (10); and prepositional phrase, as in example (11):

(9) The following may bring a case before the Court, provided that the High Contracting Party concerned, if there is one, or the High Contracting Parties concerned, if there is more than one, are subject to the compulsory jurisdiction of the Court.

(10) The liability of the producer may be reduced or disallowed when, having regard to all the circumstances, the damage is caused both by a defect in the product and by the fault of the injured person.

(11) The same shall apply, in case of an imported product, if this product does not indicate the identity of the importer referred to in paragraph 2, even if the name of the producer is indicated.

Embedded clauses with finite verbs, like in (9), are usually of conditional type and are often placed next to the elements they qualify, thus contributing to the precision of the text.

Nonetheless, other recurrent embeddings are verbless clauses like "if necessary", "if possible", etc., where the assumed subject is an impersonal "it" referring to the main clause as a whole. In this case the text appears heavier and a potential source of ambiguity. In turn, (10) is an embedding with a non-finite clause, which may give rise to problems concerning the identification of its subject, since it does not occur in the main clause.

Lexical-Functional Grammar, the theory that we adopt as our framework, treats this kinds of clause as Closed Adjuncts, which may modify the event described by the main predicate. They are subject to anaphoric or arbitrary control. It means that PRO (called "big pro") the morphologically unexpressed subject of non-finite clauses, has an antecedent in the main clause or, in case no controller is available, it is given arbitrary interpretation (see: Bresnan (1982); Delmonte (1988; 1992)).

In example (10) the antecedent of the PRO is not recoverable from the text and, moreover, the temporal reference is non-specific, we are therefore in presence of the typical conditions that give rise to arbitrary interpretation. In legislative texts this use of non-finite embedded clauses is common, even though it is sometimes considered redundant and confusing for a textual comprehension.

Finally, adverbial modifiers may be used as prepositional phrases as in (11). They have the typical adverbial mobility and so may be inserted quite freely. However, they are usually close to the elements they qualify, thus acting as substitutes for longer clausal structures.

The present analysis of embedded structures suggests a feature of legislative language already mentioned: a constant tension between condensation and precision, which may become a potential source of ambiguity.

Differently from narrative texts, the occurrence of anaphoric links between sentences is scarce: pronouns are found only when no ambiguity may arise and, in

most cases, when their antecedents are in the same sentence; finally, substitution, ellipsis and other reference devices are not common.

Coherence is usually given by the repetition of lexical items, which has the function of making explicit the meaning of a possible reference, therefore avoiding the risk of ambiguity that the use of pronominal anaphors might cause.

3. *PP attachment and Parsing*

Legal texts are very rich in adjuncts at all levels of syntactic structure, so it is necessary to give a detailed analysis of the mechanisms underlying the treatment of complements and adjuncts from a parsing point of view. Consider now the wellknown problem of PP attachment or Syntactic Closure (see Delmonte, 1984), which concerns the way in which Prepositional Phrases modifiers or arguments of a given lexical head should be dealt with in a parsing scheme. In the literature we have two different proposals: one grammatical called "garden-path" (see Frazer) and the other purely semantic called "incremental-interactive" (see Steedman & Altmann); we are only interested in the grammatically based one, which we illustrate briefly here below:

A. The Garden Path proposal has the following main features (see Kennedy et al.(1989)):

- choose the first available analysis, or words in a sentence are incorporated into complete syntactic structures at the earliest possible opportunity (i.e. the structural description is developed "on-line", word-by-word);
- do not postulate any unnecessary node - attachment of words within a structure is invariably achieved in a way which minimises the number of nodes demanded (the principle of Minimal Attachment - MA);
- if consistent with the rules of the grammar, attach each incoming word into the phrase currently being analyzed - or new words are

incorporated, wherever this is possible grammatically, into the current clause or phrase being processed (the principle of Late Closure - LC).

In a typical minimal pair example such as the following,

(12) a. She positioned the dress on the rack
b. She wanted the dress on the rack

the two principles mentioned above, MA and LC, make conflicting predictions: in (12a) MA predicts that the PP "on the rack" shall not be assigned as adjunct of the head "dress" but as locative oblique argument of the main verb "position", thus complying with the general criteria of economicity and psychological efficiency based on grammatical issues; however the principle LC would make just the opposite prediction, since there is no grammatically motivated criteria to prevent the PP to be computed locally as a semantically compatible adjunct of the head "dress". In version (12b) the two principles will make just the opposite predictions, so that MA would predict wrongly that the PP attaches to the main verb as a locative adjunct or oblique argument, whereas only LC would apply correctly.

It is a fact that examples such as these are genuinely ambiguous cases and there is no way to state general principle which could apply to both equally well, producing the best efficient result. Even if we try to maximixe on the fact that in a. we already know what the Verb Guidance (see Mitchell (1989)) is, i.e. we expect an oblique PP to be present somewhere in the following structure, we still need to take a stance as to whether we rely on a

- selecting procedure, or structure assembly process;
- monitor procedure, or structure checking process;

where the second one may involve rejecting the initial structure, backtracking and reassembling a new structure (see Mitchell, 126).

3.1. *The Parser*

The parser we use is a rule-based parser, however, being a context-sensitive grammar it incorporates naturally linguistic restrictions which make it particularly

attractive. It is not a principle-based parser in the sense of Chomsky (1986), in that there are no external principles which apply to syntactic representations in order to check their grammaticality. We could dub our parser as a global approach to linguistic parsing, in that linguistic restrictions are integrated in the parsing schema. In this sense, all selectional semantic filtering operations are carried out as soon as possible. This is allowed especially by the overall theoretical framework offered by LFG, which however we enrich, as far as syntactic c-structure and lexical representation are concerned. In particular, LFG already provides a very restricted way to deal with grammatical phenomena because of its modular structure: lexical phenomena are dealt with at a l-structure, where lexical variables are bound; syntactic phenomena are dealt with at c-structure, where syntactic variables are bound; finally, anaphoric phenomena are treated at f-structure level, where pronominal binding takes place. At the same level, also semantic compatibility between adjuncts is computed. We shall discuss the theoretical background in the following Chapter II.

A strict comparison with principle-based parsing as presented in Berwick et al.(1991) is, in our opinion, out of place for two reasons: most of the parsers presented are toy-parsers in the sense that they are not intended for parsing real extended texts; they simply try to show the feasibility of the principle and parameters approach. None of the parser actually implement the theory in a step by step fashion, due to the fact that in order to mimic Move-a in a real parsing schema D-structure should be produced first and then traces should appear where NP movement applies. However, the choice of all the parser builders reviewed in the book is to work directly on some version of S-structure. A secondary but nonetheless important feature of principle-based parsers is that, with the notable exception of Dorr's parser, all other parsers work on a two-stage mode, i.e. the structure building process is kept separate and precedes the interaction with universal principles. Dorr's parser is a modular system which tries to make use of semantic restrictions, such as thematic roles constraints on arguments of a given predicate. However, as the author herself admits, the system is still too much "syntactically driven", thus failing to capture a number of structural ambiguities of semantic nature (see Dorr, 179). X-bar theory in itself and the configurational approach to sentence and utterance analysis is deficient in its ability to solve structural ambiguities due to attachment of adjuncts. All major projections, NP, PP, AP, QP, CP, can either constitute an (optional) argument or an adjunct: in turn, adjuncts, being syntactically independent and unrestricted by either

subcategorization or argument structure, require the presence of a semantically consistent network of compatibility tests to be applied before attachment can be decided. As a matter of fact, any grammatical theory in itself, be it LFG or GB/Barriers or GPSG, lacks adequate principles for interpreting the role played by adjuncts in the structure building process, simply because adjuncts are only semantically and not syntactically restricted. However, it is clear to us, that the solution to the problem of adjuncts within systems like LFG or GPSG is much more perspicuous, simply because they are feature-based systems. Some of the solutions adopted in our system are discussed in this chapter.

Here we would like to comment on the most attractive computational features of our parser. Being a DCG, the parser is strictly a top-down, depth-first, one-stage parser with backtracking: differently from most principle-based parsers presented in Berwick (1991), which are two-stage parsers, our parser computes its representations in one pass. This makes it psychologically more realistic. The final output of the parsing process is f-structure which serves as input to the binding module and logical form: in other words, it constitutes the input to the semantic component to compute logical relations. In turn the binding module may add information as to pronominal elements present in the structure by assigning a controller/binder in case it is available. As to the most important features of DCGs, we shall simply quote from Pereira and Warren (1980) conclusions, in a comparison with ATNs:

"Considered as practical tools for implementing language analysers, DCGs are in a real sense more powerful than ATNs, since, in a DCG, the structure returned from the analysis of a phrase may depend on items which have not yet been encountered in the course of parsing a sentence. ... Also on the practical side, the greater clarity and modularity of DCGs is a vital aid in the actual development of systems of the size and complexity necessary for real natural language analysis. Because the DCG consists of small independent rules with a declarative reading, it is much easier to extend the system with new linguistic constructions, or to modify the kind of structures which are built. ... Finally, on the philosophical side, DCGs are significant because they potentially provide a common formalism for theoretical work and for writing efficient natural language systems." (ibid, 278).

Grammatical principles are located within the parser at various levels, in force of the principle that as soon a principle becomes applicable it is instantiated. LFG makes

use of X-bar notation but not in a principled way: for instance, there is no Extended Projection Principle which requires the presence of a Subject NP in the IP level representation. In languages like Italian, IPs may be rewritten simply by VP in case the SUBJect NP is computed in VP internal position, as happens with unaccusatives. Beside, for a maximal projection to be instantiated at c-structure level, the lexical head must be present. Spec-Head agreement is performed very smoothly by means of Prolog internal mechanism of variable instantiation.

Both categorial and subcategorization information is used to guide the parsing process deterministically: in particular, subcategorization both for verbs and adjectives is accessed from the lexical form of the predicate as soon as it is available. When the verb is analysed, syntactic categories like transitive, unaccusative, psychic etc. are used to guide the construction of the VP. As for adjectives, their Pred is inspected in order to decide whether to look for a given complement, which can be a constituent like a PP or a VP in the input string. Subsequently, semantic features may be used to further select these constituents as being appropriate complements to the head.

Selectional restrictions in the form of inherent semantic features associated to any lexical category are used throughout the parsing process. In the building of an NP, a PP is accepted as possible argument or adjunct in case semantic compatibility tests are passed. These tests are performed before the complete structure of the NP is built, and other constituents may be present in the final structure. When building a relative adjunct, semantic features of the head noun are percolated into the embedded open proposition in order to speed up the search for the controlled element, or landing site. This is paramount in languages like Italian, where SUBJect NPs may be freely inverted in postverbal position.

Generally speaking, the same semantic testing procedures are used for adjuncts: they are tested locally for compatibility. This is due to the fact that adjuncts in Italian may be freely interposed between any major constituent and the verb, as well as between the verb and the OBJect NP, differently from what happens in English where strict adjacency is respected (see Delmonte (1987)). We shall concentrate only on the PP attachment problem and basically on the way in which we cope with the need to pass information which has been acquired in a previous step of computation and use it to guide the parsing process.

3.2. *Two mechanisms at work*

Suppose we have to use the information that "position" is a verb which requires an oblique PP be present lexically in the structure, as results from a check in its lexical form. We take the verb in I position and then open the VP complement structure, which at first builds a NP in coincidence with "the dress". However, while still in the NP structure rules, after the head has been taken, a PP is an option freely available as adjunct.

In our parser we always give priority to what is locally possible and thus we favour a "selecting or structure assembly procedure", rather than the second procedure which would allow for backtracking, seen that this is a very risky and time-consuming option. What kind of mechanism do we incorporate in order to make the right decision also in the second example, where the main verb "want" does not guide any argument requirement in the following structure building process.

We have implemented two look-ahead mechanism which are used in the PP building rule and are always triggered, be it from a position where we have a noun as head and we already built part of the corresponding constituent structure; be it from a position where we have a verb as head and we want to decide whether our PP will be adequate as argument rather than as adjunct - in the latter case it will become part of the Adjunct Set.

Mechanism 1

- *Cross Compatibility Check*

This mechanism requires the head semantic features or inherent features to be checked against the preposition, which in turn activates a number of possible semantic roles for which it constitutes an adequate semantic marker. For instance, the preposition "on" is an adequate semantic marker for "locative" semantic role, this will cause the compatibility check to require the presence in the governing heading of inherent or semantic features that allow for location. A predicate like "dress" is computed as an object which can be assigned a spatial location, on the contrary a predicate like "want" is computed as a subjective intensional predicate which does not require a spatial location. However, in order to take the right decision, the CCC must be equipped with the second mechanism we implemented;

Mechanism 2

- *Argument Precedence*

The second mechanism we implemented is a global or metavariable which allows the parser to satisfy the subcategorization requirements in any constituent it finds itself at a given moment if the parsing process - relatively only to PPs though. Suppose that after taking "positioned" as the main verb, this mechanism is activated, by simply copying the requirements on PP oblique locative present in the lexical form associated with the predicate "position" in the lexicon, in the AP metavariable. As soon as the NP "the dress" is opened, after taking "dress" as N at the head position, the parser will meet the word "on", which allows for a PP adjunct. While in the P head position, the parser will fire the CCC mechanism first to see whether the preposition is semantically compatible, and in case it is, the second AP mechanism will be fired. This will cause the system to do the following steps:

i. check whether the metavariable is empty or not;
ii. and in case it is instantiated, to control the semantic role associated with it;
iii. to verify whether the P head is a possible semantic marker for that semantic role: in our case, "on" is a possible semantic marker for "locative" semantic role;
iv. finally to cause the parser to fail on P as head of a PP adjunct of the head noun;
v. produce a closure of NP which obeys Minimal Attachment principle.

3.3. *Some examples*

In our text there is a great number of examples which can be used as empirical evidence for the need to use lexical information in order to reduce parsing loads resulting from structure monitoring, or rather backtracking procedures. Our examples are taken from the text included as an Appendix at the end of the paper: we mark decision points with a bar:

a. Council directive | **of** july 1985 | **on** the approximation | **of** the laws, | regulations and | administrative provisions | **of** the Member States | concerning liability | **for** defective products.

At the first boundary we have "of" which is non semantically marked and no prediction is available, so that the default decision is to apply Late Closure, which turns out to be the correct one. When the second preposition is found we are in the NP of the PP headed by "of", and we have taken the date "1985": this will cause the CCC to prevent the acceptance of the preposition "on" as a semantically compatible marker thus preventing the construction of the NP headed by "approximation".
Notice, that in case that would be allowed, the NP would encompass all the following PPs thus building a very heavy NP: "the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products". In case the parser had a structure monitoring strategy all this work would have to be undone and backtracking would have to be performed. Remember that the system does not possibly know where and how to end backtracking unless by trying all possible available combination along the path. In our case, the presence of a coordinate structure would render the overall process of structure recoverability absolutely untenable.

Another important decision has be taken at the boundary constituted by the participial head "concerning": in this case the CCC will take the inherent features of the head "States" and check them with the selection restrictions associated in the lexical form for the verb "concern". Failure in this match will cause the NP "the Member States" to be closed and will allow the adjunct to be attached higher up with the coordinated head "laws, regulations and administrative provisions". In this case, all the inherent features are collected in a set that subsumes them all and can be used to fire CCC.
Notice that the preposition "for" is lexically restricted in our representation for the noun "liability", and the corresponding PP that "for" heads interpreted as a complement rather than as an adjunct. We include here below the relevant portion of each utterance in which the two mechanisms we proposed can be usefully seen at work. We marked with a slash the place in the input text in which, usually when the current constituent is a NP a decision must be taken as to whether causing the parser

to close (MA) or to accept more text (LC) is actually dependent upon the presence of some local trigger. This trigger is mostly a preposition; however, there are cases in which, see e., f., h., i., the trigger is a conjunction or a participle introducing a reduced relative clause. Coordinate NPs are a big source of indecision and are very hard to be detected if based solely on syntactic, lexical and semantic information. For instance, e. can be thus disambiguated, but h. requires a matching of prepositions; In the case represented by i. we put a boundary just before a comma: in case the following NP "the Member State" is computed as a coordination - which is both semantically, syntactically and lexically possible, the following sentence will be deprived of its lexical SUBJect NP - in this case, the grammar activates a monitoring procedure independently so that backtracking will ensue, the coordinate NP destroyed and the comma computed as part of the embedded parenthetical (which is in turn an hypothetical within a subordinate clause!!). Notice also that a decision must be taken in relation to the absolutives headed by apast participle which can be intended as an active or a passive past participle: in the second case the head noun would have to be computed as an OBJect and not as a SUBJect

b. a differing degree of protection of the consumer | **against** damage caused by a defective product | **to** his health or property

c. in all member states | **by** adequate special rules, it has been possible to exclude damage of this type | **from** the scope of this directive

d. to claim full compensation for the damage | **from** any one of them

e. the manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part | **and** any person

f. The liability of the producer | **arising** from this directive

g. any person who imports into the community a product | **for** sale, hire or any form of distribution | **in** the course of his business

h. both by a defect in the product | **and** by the fault of the injured person

i. However, if... the commission does not advise the Member State | **concerned** that it intends submitting such a proposal | **to** the council |, the Member State

4. *Semantics and Discourse Model*

The main tenet of the theory supporting the construction of our system is that it is possible to reduce access to domain world knowledge by means of contextual reasoning, i.e. reasoning triggered independently by contextual or linguistic features of the text. It is sensible to assume that when understanding a text a human reader or listener does make use of his encyclopaedia parsimoniously. Contextual reasoning is the only way in which a system for Natural Language Understanding should tap external knowledge of the domain. In other words, a system should be allowed to perform an inference on the basis of domain world knowledge when needed and only then. In this way, the system could simulate the actual human behaviour in that the access to extralinguistic knowledge is triggered by contextual factors independently present in the text and detected by the system itself.

It is also our view that humans understand texts only whenever all the relevant information is supplied and available. Descriptive and narrative texts are usually self-explanatory - not so, literary texts - in order to allow even naive readers to grasp their meaning. Note that we are not here dealing with spoken dialogues, where a lot of what is meant can be left unsaid or must be implicitly understood.

In the best current systems for natural language understanding, linguistic components are kept separate from knowledge representation, and work which could otherwise be done directly by linguistic analysis is duplicated by the inferential mechanism. Linguistic representation is usually mapped onto a logical representation which is in turn fed onto the knowledge representation of the domain in order to understand and validate a given utterance or query.

Thus the domain world model must be priorly built, usually in view of a given task the system is set out to perform. However, it is clear that this modelling is

domain and task limited and no generality whatsoever is achieved from it. In some of these systems, the main issue is how to make the two realms interact as soon as possible in order to take advantage of the inferential mechanism to reduce ambiguities present in the text or to allow for reasoning on linguistic data, which otherwise couldn't be understandable.

We assume that an integration between linguistic information and knowledge of the world can be carried out at all levels of linguistic description and that contextual reasoning can be thus performed on the fly rather than sequentially. This does not imply that external knowledge of the world is useless and should not be provided at all: it simply means that access to this knowledge must be filtered out by the analysis of the linguistic content of surface linguistic forms and the abstract representations of the utterances making up the text. One of the major problems in every attempt at real text understanding has been that of controlling the inferences associated with the interpretation process. In our opinion as well as in Partee's and Dahlgren's opinion, separate semantic levels - lexical level, grammar and parser, pronominal and discourse level, temporal level - conspire to produce the global interpretation. Indeed, discourse analysis requires a highly flexible structure where each individual source of semantic knowledge contributes its own local inference which is then combined at a higher level. In this perspective, semantic interpretation is carried out incrementally and the whole process requires cooperation among separate semantic modules in order to produce a coherent inferential process. Such an approach makes the reasoning process both tractable and computationally efficient.

In particular, as in Dahlgren's approach, we assume that an autonomous syntactic component draws upon lexical semantic information for tackling problems such as prepositional phrases attachment, constituency interpretation and its coupling to argument functions or adjunct functions, and word sense disambiguation. The remaining components, however, in our system do not work in parallel, but are fed by the output of the parser, onto which pronominal binding and quantifier scoping is carried out, also using semantic interpretation. The remaining upper modules are essentially semantic interpretation modules including temporal reasoning where spatiotemporal inferential processes are fired; discourse topic hierarchy construction to enable reference resolution at discourse level (see Delmonte & Bianchi, (1994)), and finally the building of a discourse model of facts.

Semantic peculiarities concern genre specialization that affect the way in which the interpretation process is tied up with linguistic forms. It is a fact that linguistic structure and form is the first means by which meaning is conveyed to language users, and it will be shown that it is often contingent upon genre.

Generally speaking, building a Discourse Model is a precondition for any reader or generic addressee of the contents of a legal text to enable reference to entities and events contained in the text. A DM is clearly only a part of the overall process of understanding which makes heavy use of background mutual knowledge on the side of the adressee in order to carry out the complex inferences required by this genre.

In line with current assumptions about the ontological status of entities and events referred to in any legal texts, we also assume that they are spatiotemporally disanchored and they have no counterpart in the real present world: in other word, their referential status is abstracted from spatiotemporal restrictions. Nonetheless, the DM will be represented as a list of facts and sits very much in the same way in which it would happen in narrative texts. The reason for this is very simple, the DM constitutes the semantic informational representation of the linguistic structure of any text or discourse. It aims at simulating the processes underlying anaphora and reference resolution within the text, thus registering and storing information in a given formal format which is the actual interpretation related to entities and events contained in the text.

For instance, a norm, directive or other legal text might contain Obligations which are expressed by a specialized use of the modal "shall" and as is the case with Permissions which are highlighted by the presence of the modal "may", are treated as facts in the DM; in turn, Hypotheses do not count as extensional objects and are carefully set apart from Conditions which assume a similar syntactic pattern but have a different semantic marker.

If we take reference to classes as the norm in legal language we still must allow for individuals - the current directive is one such case - or sets, the commission or the European community, or still the Member states.

5. *Reference and General Nouns*

As for human entities, we want to differentiate generic from defining descriptions or properties; in turn these descriptions, which might be exemplified by "consumer" or "producer" in our case, are roles attached to default properties or general nouns like "person" or "man" which are computed as generic entities. A side effect of the special use of referential expressions in legal texts is their lack of referential content: in other words, legal texts do not refer extensionally to entities in the world but deal with generic entities or with classes of individuals; proper nouns may only refer to institutions and as such they can be treated as collective entities or sets; common nouns may deal with abstract concepts belonging to the vocabulary of law or to concrete individuals or objects which could be found in the world but needn't do so. Usually, when common nouns are introduced in a directive or bill they are individuated by a specific definition: this is particularly so, when those nouns are the addressee or the main topic to which the norm applies.

The general noun "person" should be referentially empty in legal contexts: consider its lack of gender specification and its inherent generic nature if compared to common nouns related ro social roles. It is heavily used to refer either to actual or to potential addressees: in particular, we might be talking about "producers" and "consumer", and then go on talking about "the injured person" or "two or more persons" which can cause damage. In fact, "person"can be used both with modifiers and without them, as a definite NP or as an indefinite one. When it is used as a definite NP, it requires some modifier to be properly coreferred; no bare NPs can be created in association with such a generic head.

While common nouns may refer either to individuals or to classes, singular definite NPs with a generic head cannot be computed as individuals but only as classes and this must be inferred from the modifier or adjunct that specifies it adequately.

Plural definite NPs with generic heads are computed as classes, and can be used to corefer to other singular generic nouns or to specific nouns which they share default properties with.

Differently from what happens in narrative texts, legal texts introduce indefinite NPs not to assert the existence of some individual or set but to refer to classes or to generic entities.

Strzalkowski and Cercone (1989) in their work introduce the problem of reference from a semantic computationally feasible point of view. This is how they define their proposal:

> "... we introduce a layered model of reality (the universe) as perceived by a discourse participant, and define relative singularity of objects in this universe as an abstraction class of the layer-membership relation. Subsequently, linguistic descriptions and names are classified as **singular**, **measurably singular**, or **non-singular** depending upon what they are assumed to denote in the universe. The relationship between objects referred to in discourse and classified into different layers (levels) of the universe model has a particular significance for resolution of certain types of cohesive links in text. We call these links remote reference because they cross level boundaries." (p.172)

The tripartipe classification of entities in the model reflects our proposal where we have generic entities, classes and individuals. As in S & C they are partially ordered by the relation "more informative than": intuitively, they introduce a relation of relative singularity among objects that allows us to break down the universe of objects into classes or levels, where a lower level L0 consist of manifestations, extensions, instances of objects at level L1. In fact S & C always speak of the existence of three main levels, one for each type of denotation: if L0 contains objects with a measurably singular interpretation, L-1 may contain the same object with a specific or singular interpretation - if possible at all, and L+1 will contain referring expressions - nouns, definite descriptions, pronouns but also other nominals denoting objects which are non-singular. If level L+1 contains generalizations of objects from L0, then level L-1 will contain their specializations of extensions (ibid., 177). Any non singular object can potentially be decomposed into instances in many ways, depending upon the relation that bind the two objects, or coordinate - part_of, instance_of, genus, specimen etc.; also, and more generally, whenever a higher-level

object is decomposed with two different coordinates, the resulting sets of instances need not belong to the same level (ibid., 178).

However, we will assume that natural language users, when introducing referring expressions in a text or discourse will abide the wellknown maxim that requires them not to be ambiguous. In particular, in our model the spatiotemporal location indeces assigned according to semantic and linguistic criteria are the only possible coordinates but cannot be used to partition our discourse model into layers or levels, with respect to some current level, corresponding to the level of reference at a present point of discourse. We do not find any convenience in introducing such layers into the overall computation of entities and their reference. In turn S & C do not describe in detail how one could compute remote reference in discourse, apart from establishing an obvious definition that relates two objects by means of remote reference in case they have different kinds of singularity. Naming an object previously existing in the discourse model may end up simply with a coreferring operation, or with the need to create a supercontext in case it is less informative than the existing entity, or a subcontext in case it is an instance of the previously existing entity (ibid., 182).

Now consider the problem of general nouns: person, man, individual etc. They constitute the most common and most frequent noun in our texts and this is due to their referential properties: they can pick the maximum set of human entities available when used with no attribute or modifier; they can also be used to corefer to single classes of individuals by adequate modification. In other words, the concept "person" will be present at different levels of generality in the DM and it will be instantiated with a different semantic identifier according to its "dattrs" (see Woods and Brachman). Reference to one or the other of the instantiations of "person" will depend upon the way in which the discourse entity is described, in other terms on its "dattrs", i.e. attributes, properties, parts, constituents, features, relations and so on. Notice that in our DM every entity has a description in terms of situation semantics, with a polarity and a spatiotemporal anchorage in terms of indices for main spatial and main temporal location.

6. *Inferential processes*

In our system, the problem of assigning properties to entities in narrative genre is carried out in two main phases: first only referring expressions (ref_exs) asserted as best candidates for topichood and ranked as Main, Secondary, Expected or Potential Topics by the Topic Hierarchy Mechanism are considered. This might or might not apply to legal genre, on the basis of the assumption that there is no plot nor main character to focus on. However, legal texts contain pronominals which refer in the previous discourse segment and cannot be done away with. In case no such coreferent is found, the remaining DM is searched for. Generally speaking, there are four possible ways in which the problem might be solved:

CASE A
there is already at least one such entity in the Discourse Model (DM) with the same class restriction which could be used to corefer or cospecify the current ref_ex disregarding its possible contextually determined properties and considering only the class or name that has been used as first or Initial Description (ID);

CASE B
the entity to be picked up as possible coreferent or cospecifier must be semantically equal: i.e. we consider number of the current ref_ex and if plural, check whether the entity has the same cardinality or is asserted as a class; and if singular look only for individuals;

CASE C
the entity and the ref_ex must have the same attributes, roles or other properties assigned to it in the DM and present in the current ref_ex;

CASE D
works as case C and in addition it looks only for properties associated to entities asserted as Topics of discourse in the previous text; in case no such property is present, a new entity is entered in the DM.

It is easy to notice that, case D is the most restrictive but also the most genre bound procedure: in fact, it is only useful in case the text is a narrative one.

The first question is discovering whether the current ref_ex is already present in the discourse model. If the problem at hand were that of matching the current linguistic form of the predicate or concept with those present in the discourse model (DM) this would be easily solved. However this procedure is clearly insufficient and leads to mistaken matches in case cardinality, definiteness, or simply the modifiers of the current linguistic use of a certain entity present in the DM does not coincide semantically with it.

The overall framework is further complicated by the presence of generic entities and by the fact that what we are dealing with are usually entities which might be easily subsumed by them or constitute a specific subclass. A "consumer" or a "producer" is clearly such a subclass in case the general noun "person" has already been introduced in the DM. However, problems will arise when the contrary applies, i.e. whenever a general noun is introduced with specific properties added by adjunction or modification: in this case it would both represent a subclass of a generic entity and a coreferent of a some class of individuals or role.

The general problem to be solved might be coped with by the following procedure:

1. find all semantic identifiers associated in the DM with the predicate of the current referring expression starting from the most recent ones in the stack of ids' associated with some previous topic;

1.i search all ids' starting from the most recent one and make a list;

1.ii then for each id, look whether the current predicate head is associated to some property in the DM;

a.1 no identifier associated with the head, goto 2.

a.2 there is at least one identifier associated with the head;

b. verify whether the set of adjuncts, modifiers and other property functional assigners, MODS associated to the current head is empty or not

b.1 if it is empty, goto f.;

b.2 if it is nonempty, find all predicate heads associated to MODS;

c. search the DM for properties associated to the predicate head other than the head itself and the default properties;
c.1 if no property is found in the DM goto f.;
c.2 if the list of properties in the DM is nonempty:
d. make the intersection between the two sets of properties: set1 from MODS, set2 from DM;
d.1 if the intersection is empty goto e.;
d.2 if the intersection is nonempty and the property/ies found is/are equal to the one/s contained in set1 then fail;
d.3 else continue.
e. check in the external knowledge of the world whether the two sets contain properties which are synonymous or which are inferentially derivable from other properties;
e.1 if there are some such properties, remove them from the list;
e.2 if not continue;
f. check preliminarly whether the two sets of properties are nonempty; then, check number, and if plural check cardinality and then
f.1 if it is the same assert the set of properties in MODS;
i. a singular is compatible with inds;
ii. a plural is compatible with sets and classes
•in legal genre, a singular is compatible both with singular and with classes!!
f.2 if it is different fail;
g. search for other ids associated to the current head;
g.1 if no other id is present goto end;
g.2 if some other id is present goto b.

2. find all semantic identifiers associated in the discourse model (DM) with the predicate of the current referring expression which are not included in the list of the topic identifiers;

2.i if it is empty, goto 3.

2.ii if it is nonempty goto b.

3. find all semantic identifies associated in the External Knowledge Base with the predicate of the current referring expression;

3.i if it is empty, goto end.

3.ii if it is nonempty goto b.

end.

Let's now go back to general nouns: in any text, "person" should be introduced effectively only whenever a number of possible specific entities which might be subsumed by "person" already exist in the DM. In the texts we analyzed, this is what happens: in particular, the Council Directive we used and implemented which is discussed in mode detail at the end of this paper, has the following textual structure:

1. introduce main topics and addressee of its contents, which are "producers" and "consumers" living in the "Member States"; the directive concerns the protection of consumes and the producers' liability in relation to injury or damage which might result from defective products. Importers of products who present themselves as producers are also regarded as such.
2. subsequently, both main topics are coreferred by means of the general noun "person".

On first appearance of common nouns which might be subsumed by "person", the system checks whether there is any such entity in the DM: in the affirmative case it simply inherits its identifier and in the negative case, a generic entity is asserted in the DM. This entity is not an individual nor a set but has the following properties:
- it may subsume other generic entities of the same kind;
- it may be used to infer the nature of references to individuals by means of "person" and some specific attribute or property.

In particular, in the following text we find "several persons liable" and "injured person". The first referring expression is computed as "producer" and the second one as "consumer". In order to get this interpretation, which is the one intended in the Directive, an inferential process must be carried out on the basis of external knowledge of the world. However, the trigger to start this process is constituted by the attribute "injured" which is computed in the property checker algorithm as a

possible new property to be associated to some entity already existing in the DM. The search starts from a referring expression that has some attribute or modifier which is a property not yet asserted for the corresponding entity. Consider the case of "injured person", the algorithm by looking into the DM recovers the identifier associated to "consumer" on the basis of the inference that he is a "person" which is expressed in the following list of facts:

```
class(_, Id)
class(_, Ind)
fact(_, isa, [_:Ind, class:person], 1, _, _)
fact(_, ist_of, [_:Ind, class:man], 1, _, _)
fact(_, role, [consumer, Id, Ind], 1, _, _)
fact(_, isa, [_:Id, class:consumer], 1, _, _)
in(_, Id, Ind)
```

When the Id of "consumer" is recovered from the DM by the property checker that looks for correspondences between literal predicates and their relations in the DM, the "injured" attribute and the Id is passed on to the following inference engine that collects knowledge of the world associated to the trigger and checks to see whether it applies to the current identifier and its properties in the DM.

```
infer_process(Trigger, Id):-
  infer_trig(Trigger, Props),
  infer_rels(Id, Props).
```

This is done by means of an inferential process that takes as input external knowledge of the world, where we deposit information related to implicit knowledge, mutual knowledge and specialized information which could be part of T-box component in a knowledge base understanding system. In those inheritance networks (see Woods (1978)), each concept that the system understands is represented in a network of concepts, which have to cope with the problem of internal recursion - i.e. of a concept defined in terms of another concept. Our taxonomy is only a small list of facts and the portion that interests us now is represented as follows:

```
infer_trig(injured, [cause(damage, 0), has(protection, 1), liable(Id, 0)]).
infer_trig(liable, [cause(damage, 1), has(liability, 1)]).
```

Each term is made up of a trigger, "injured", "liable" which are the properties by means of which the general noun "person" is modified; in turn each trigger is associated with a set of relations and properties which have as argument either a class predicate and a polarity, or a variable and a polarity.

In the text, we learn that "injured person" is only used to corefer to "consumer"; we also know that in order to be interpreted as "injured" a person has not to have caused the damage, nor to be "liable" for it, and finally be the one that is given "protection". By definition, then, a person is liable in case "he caused damage" and in case " he has liability".

```
infer_rels(Id, [Prop(Id, 1)|Props]):-
  fact(_,Prop, [_:Id], 1, _, _),
  infer_rels(Id, Props)
  ;
infer_rels(Id, [Prop(Id, 0)|Props]):-
  (not fact(_,liable, [_:Id], 1, _, _))
  ;
  fact(_,liable, [_:Id], 0, _, _))
  )
  infer_rels(Id, Props)
  ;
infer_rels(Id, [Rel(Prop, 1)|Props]):-
  fact(_, Rel, [_:Idx, _:Id], 1, _, _),
  fact(_, isa, [_:Id, class:Prop], 1, _, _)
  infer_rels(Id, Props)
  ;

infer_rels(Id, [Rel(Prop, 0)|Props]):-
  fact(_, Rel, [_:Idx, _:Id], 0, _, _),
  fact(_, isa, [_:Id, class:Prop], 1, _, _)
  ).
```

7. *DM and inferences: an example*

We will show here below the DM relatively only to three entities: consumer, producer and their general noun "person" subsuming both. In utterance 4 with first appearance of "consumer and producer" the system generates a generic entity "person" which subsumes both: the inclusion relation is registered by the fact in (_, A, B) where A is the semantic identifier or initial description of the superset or superobject, and B is the subsumed entity:

[Whereas approximation of the laws of the 'Member' 'States' concerning the liability of the producer for damage caused by the defectiveness of his products is necessary, because the existing divergences may entail a differing degree of protection of the consumer, against damage caused by a defective product to his health or property.]
ent(infon102, id22)
fact(infon103, isa, [arg:id22, class:person], 1, univ, univ)
class(infon104, id23)
class(infon105, id24)
fact(infon106, cause, [causer:id5, arg:id24], 1, univ, univ)
fact(infon107, inst_of, [ind:id24, class:abstract_state], 1, univ, univ)
fact(infon108, isa, [ind:id24, class:damage], 1, univ, univ)
fact(infon109, against, [arg:id23, malefactive:id24], 1, univ, univ)
fact(infon110, inst_of, [ind:id23, class:man], 1, univ, univ)
fact(infon111, isa, [ind:id23, class:consumer], 1, univ, univ)
in(infon112, id23, id22)
fact(infon113, role, [consumer, id23, id22], 1, univ, univ)
fact(infon114, cause, [arg:id4, damage:id24], 1, univ, univ)
class(infon115, id25)
fact(infon116, inst_of, [ind:id25, class:man], 1, univ, univ)
fact(infon117, isa, [ind:id25, class:producer], 1, univ, univ)
in(infon118, id25, id22)
fact(infon119, role, [producer, id25, id22], 1, univ, univ)
class(infon133, id29)

fact(infon134, differing, [ind:id29], 1, univ, univ)
class(infon135, id30)
fact(infon136, has, [experiencer:id23, arg:id30], 1, univ, univ)
fact(infon137, inst_of, [ind:id30, class:abstract_state], 1, univ, univ)
fact(infon138, isa, [ind:id30, class:protection], 1, univ, univ)
fact(infon139, of, [arg:id29, specif:id30], 1, univ, univ)
fact(infon140, inst_of, [ind:id29, class:[measure]], 1, univ, univ)
fact(infon141, isa, [ind:id29, class:degree], 1, univ, univ)

In utterance 6 we learn that the producer should be made liable at certain conditions, and this is registered as an additional fact about producers:

[Whereas protection of the consumer requires that all the producers involved in the production process should be made liable in so far as their finished product, component part or any raw material supplied by them was defective.]
fact(infon248, liable, [nil:id25], 1, univ, univ)

In utterance 8 the text introduces an undefined set of "several persons" which are liable for a certain damage and are related to protection of the consumer in the same context. The system assigns to "injured person" the same identifier as the consumer, creates a new class of entities "person" which are liable and cause "damage", included in the same superset of producers:

[Whereas, in situations where several persons are liable for the same damage the protection of the consumer requires that the injured person should be able to claim full compensation from any one of them.]
fact(infon323, injured, [ind:id23], 1, univ, univ)
class(infon324, id73)
fact(infon325, full, [ind:id73], 1, univ, univ)
fact(infon326, inst_of, [ind:id73, class:legal], 1, univ, univ)
fact(infon327, isa, [ind:id73, class:compensation], 1, univ, univ)
fact(infon328, cause, [arg:id25, damage:id24], 1, univ, univ)
class(infon329, id74)
fact(infon330, liable, [nil:id74], 1, univ, univ)

fact(infon331, cause, [arg:id74, damage:id24], 1, univ, univ)
in(infon332, id25, id74)
fact(infon333, inst_of, [ind:id74, class:man], 1, univ, univ)
fact(infon334, isa, [ind:id74, class:person], 1, univ, univ)
class(infon335, id75)
fact(infon336, inst_of, [ind:id75, class:abstract_state], 1, univ, univ)
fact(infon337, isa, [ind:id75, class:situation], 1, univ, univ)
fact(id76, claim, [experiencer:id74, theme_aff:id73, source:id74], 1, tes(finf1_aq4), univ)

In utterance 10 liability for damage is charged on the producer: the system looks for a similar predication in the knowledge base and finds the one related to infon248, which was asserted in utterance 6 above:

[The producer shall be liable for damage caused by a defect in his product.]
fact(infon382, cause, [arg:id25, damage:id24], 1, univ, univ)
class(infon383, id84)
fact(infon384, in, [arg:id84, locative:id5], 1, univ, univ)
fact(infon385, inst_of, [ind:id84, class:legal], 1, univ, univ)
fact(infon386, isa, [ind:id84, class:defect], 1, univ, univ)
fact(infon387, cause, [causer:id84, arg:id24], 1, univ, univ)
fact(id85, be, [prop:infon248], 1, tes(f3_dd07), univ)

In utterance 11 we find "the injured person" again, and the system picks up id23 associated to consumer. Notice the computation of the meaning for "relationship" which has "causal" as modifier: the system understands it as a relation and finds infon387 which is associated to a fact in the knowledge base asserted in the previous utterance and saying that there is a "cause" relation between "defect" and "damage". This was expressed in terms of semantic roles, i.e. the defect is interpreted as a "causer" and its argument is id24, the "damage". In the new utterance, this is linguistically formulated in terms of "relationship", where there is a semantic marker "between" which expresses a relation, and the relation has two arguments:

[The injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage.]
ind(infon423, id95)
class(infon424, id96)
fact(infon425, inst_of, [ind:id96, class:abstract_state], 1, univ, univ)
fact(infon426, isa, [ind:id96, class:defect_damage], 1, univ, univ)
fact(infon427, between, [arg:id95, relation:id96], 1, univ, univ)
fact(infon428, cause, [arg:id95, relation:infon387], 1, univ, univ)
fact(infon429, inst_of, [ind:id95, class:abstract_state], 1, univ, univ)
fact(infon430, isa, [ind:id95, class:relationship], 1, univ, univ)
fact(infon433, [defect, damage, relationship], [arg:id97], 1, univ, univ)
fact(id98, prove, [actor:id23, prop:infon433], 1, tes(finf1_dd07), univ)
fact(id100, require, [actor:id23, prop:id98], 1, tes(f4_dd07), univ)

In utterance 12 the same predication is present, "be liable" and the same infon248 is picked up:

[Where, as a result of the provisions of this directive, two or more persons are liable for the same damage.]
fact(id104, be, [prop:infon248], 1, tes(f49_dd08), univ)

In utterance 13 a pronoun is introduced intersentially to corefer to the same persons, associated to the class of producers:

[They shall be liable jointly and severally, without prejudice to the provisions of national law concerning the rights of contribution or recourse.]
fact(id109, be, [prop:infon248], 1, tes(f3_dd10), univ)

In utterance 14, we find a reference to the superset of persons, the one introduced as generic entity, since the system does not find any hint in the utterance by means of which "a person" could be interpreted as belonging either to the class of consumers or to the class of producers:

[A product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account.]
fact(id116, expect, [experiencer:id22, theme_unaff:id111], 1, tes(finf1_dd10), univ)
fact(id118, entitle, [theme:id22, prop:id116], 1, tes(f6_dd10), univ)

Finally, in utterance 17, a new set of "person" is introduced which "imports a product" and according to the interpretation assigned by the utterance to this set it is included in the superset of the class of producers:

[Without prejudice to the liability of the producer any person who imports into the community a product for sale, hire or any form of distribution in the course of his business, shall be deemed to be a producer within the meaning of this directive.]
class(infon631, id132)
in(infon633, id25, id132)
fact(infon634, inst_of, [ind:id132, class:man], 1, univ, univ)
fact(infon635, isa, [ind:id132, class:person], 1, univ, univ)
fact(id138, import, [agent:id132, theme_aff:id7], 1, tes(f3_dd17), univ)

8. *Discourse structure*

Narrative has a clear temporal sequence which constitutes its backbone: legal texts lack a temporal timeline and there is neither a temporal progression nor a fixed spatiotemporal location to which events are anchored. In fact, legal texts have no specific temporal reference.

In narrative texts, an incoming clause is included within the previous discourse structure is depending on the following parameters:
- there is no change in the temporal interval, i.e. no new time focus is introduced;
- there is no change in the topic structure, i.e. no new participant is introduced which amounts to saying that either the Topic Hierarchy has no Expected Topic, but a Main Topic, and the state of the discourse is Continue; or, technically a change might be

caused by a Resume, i.e. by the resumption of some previously asserted Topic, from the Discourse Model.

When neither of these conditions are met, the same Discourse Segment is continued either by a Same_Level move or by a Down move (see Polanyi; Webber). However, legal texts lack reference to a spatiotemporal location and no Time Focus is provided, in the sense that there is no fixed external and extensional spatiotemporal location to which events or rather states and entities are bound to. On the contrary, topics hierarchy is still an important tool to organize legal texts even though it might work differently from the way in which it works in narrative texts. These differences are essentially due to the use of referential expressions which might be nominalizations or deverbal nouns which might req uire recovering the event, activity or state referred to by the base verb; deajectival nouns might require making reference to adjectives used in copulative constructions. Finally, general or generic nouns might either refer to some generic entity or to a specific class which is usually included in the general noun, at certain conditions.

dd11.[a product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account.]

same_level:level(11-14)
clause:12-15
topics:[main:id21:person, second:id25:product]
main_fact:defective([id25:product], 1, id44)
ref_int:tint(tr(f3_01), [tr(f3_01), tr(f3_10), tr(f3_10)])
temp_rel:during(tr(f3_10), tr(f3_01))
disc_rel:definition
disc_str:12-[13, 14, 15]
disc_dom:objective
p_o_view:legislator

same_level:level(12-15)
clause:12-16
topics:[main:id21:person, second:id25:product]
main_fact:provide([id25:part, id72:safety], 0, id44)
ref_int:tint(tr(f3_01), [tr(f3_01), tr(f3_10), tr(f3_10), tr(f7_10)])
temp_rel:overlap(tr(f7_10), tr(f3_01))

disc_rel:condition
disc_str:12-[13, 14, 15, 16]
disc_dom:objective
p_o_view:legislator

same_level:level(12-16)
clause:12-17
topics:[main:id21:person, second:id25:product]
main_fact:entitle([id20:person, id77:expect, id79:pr], 1, id44)
ref_int:tint(tr(f3_01), [tr(f3_01), tr(f3_10), tr(f3_10), tr(f7_10), tr(f6_10)])
temp_rel:during(tr(f6_10), tr(f3_01))
disc_rel:definition
disc_str:12-[13, 14, 15, 16, 17]
disc_dom:objective
p_o_view:legislator

9. *A Legal Text*

Council directive of july 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.
The council of the european communities has adopted this directive.
Having regard to the proposal from the commission.
Whereas approximation of the laws of the Member States concerning the liability of the producer for damage caused by the defectiveness of his products is necessary because the existing divergences may entail a differing degree of protection of the consumer against damage caused by a defective product to his health or property.
Whereas liability without fault should apply only to movables which have been industrially produced.

Whereas protection of the consumer requires that all the producers involved in the production process should be made liable in so far as their finished product, component part or any raw material supplied by them was defective.
Whereas, to the extent that liability for nuclear injury or damage is already covered in all member states by adequate special rules, it has been possible to exclude damage of this type from the scope of this directive.
Whereas, in situations where several persons are liable for the same damage, the protection of the consumer requires that the injured person should be able to claim full compensation for the damage from any one of them.
Producer means the manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part and any person who, by putting his name, trade mark or other distinguishing feature on the product presents himself as its producer.
The injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage.
This directive shall not apply to injury or damage arising from nuclear accidents and covered by international conventions ratified by the Member States.
Where, as a result of the provisions of this directive, two or more persons are liable for the same damage.
They shall be liable jointly and severally, without prejudice to the provisions of national law concerning the rights of contribution or recourse.
A product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account.
The liability of the producer arising from this directive may not, in relation to the injured person, be limited or excluded by a provision limiting his liability or exempting him from liability.
Without prejudice to the liability of the producer any person who imports into the community a product for sale, hire or any form of distribution in the course of his business shall be deemed to be a producer within the meaning of this directive.
Whereas to protect the physical well_being and property of the consumer the defectiveness of the product should be determined by reference, not to its fitness for use, but to the lack of the safety which the public at large is entitled to expect.

The liability of the producer may be reduced or disallowed when, having regard to all the circumstances, the damage is caused both by a defect in the product and by the fault of the injured person or any person for whom the injured person is responsible.
However, if within three months of receiving the said information the commission does not advise the Member State concerned that it intends submitting such a proposal to the council, the Member State may take the proposed measure immediately.
Any member state may provide that a producer's total liability for damage resulting from a death or personal injury and caused by identical items with the same defect shall be limited to an amount which may not be less than 70 million ecu.

References

Altman, G.T.M.(ed.) (1989). *Language and Cognitive Processes* 4, 3/4, Special Issue - Parsing and Interpretation.

Berwick, R. et al. (eds.) (1991). *Principle-based parsing*, Kluwer, Dordrecht.

Bianchi, D., R.Delmonte, E.Pianta (1992). "GETA_RUN - A General Text Analyzer with Reference Understanding", in Proc. 3rd Conference on Applied Natural Language Processing, Systems Demonstrations, Trento, ACL 1992, 9-10.

Bianchi, D., R.Delmonte, E.Pianta (1993). "Understanding Stories in Different Languages with GETA_RUN", Proc. EC of ACL, Utrecht, 464.

Bianchi, D., R.Delmonte, E.Pianta, S.Sartori (1993). "Reference Resolution by Contextual Reasoning from Grammatical Representations with GETA_RUN", University of Venice Working Papers in Linguistics, 3, 1, 31-68.

Clifton, C., and F. Ferreira (1989). "Ambiguity in Context", in G.Altman (ed.), *Language and Cognitive Processes*, op.cit., 77-104.

Delmonte, R.(1984). "La "syntactic closure" nella Teoria della Performance", *Quaderni Patavini di Linguistica*, 4, Padova, 101-131.

Delmonte, R. (1992). "Linguistic and Inferential Processing in Text Analysis by Computer", *Studi Linguistici Applicati*, Unipress, Padova.

Delmonte, R. (1994). "Inferences and Discourse Structure", in G.Ferrari (ed.), Proc. IV Convegno dell'AIIA, Parma, pp.11-15.

Frazier, L. (1987a). "Sentence processing", in M.Coltheart (ed.), *Attention and Performance* XII, Hillsdale, N.J., Lawrence Elbaum.

Frazier, L. (1987b). "Theories of sentence processing", in J.Garfield (ed.), *Modularity in knowledge representation and natural language understanding*, MIT Press, 291-308, Cambridge MA.

Garrod, S. and T. Sanford (1988). "Thematic Subjecthood and Cognitive Constraints on Discourse Structure", *Journal of Pragmatics* 12, 599-534.

Hobbs, J.R. (1993). "Intention, Information, and Structure in Discourse: A First Draft", ms., SRI.

Kennedy, A., et al. (1989). "Parsing Complements", in G.Altman (ed.), *Language and Cognitive Processes*, op.cit., 51-76.

Mann, W. and S.Thompson (1987). "Rhetorical Structure Theory: a Theory of Text Organization". ISI Reprint Series ISI/RS-87-190, Information Sciences Institute.

Mitchell, D.C. (1989). "Verb-Guidance and Other Lexical Effects in Parsing", in G.Altman (ed.), *Language and Cognitive Processes*, op.cit., 123-154.

Steedman, M.J. and G.T.M. Altmann (1989). "Ambiguity in Context: a Reply", in Altman (ed.), *Language and Cognitive Processes*, op.cit., 105-122.

Strzalkowski, T. (1986). "An Approach to Non-Singular Terms in Discourse", in *Proceedings of the 11th International Conference on Computational Linguistics* (COLING-86), Bonn, West Germany, 362-364.

Strzalkowski, T. and N.Cercone (1989). "Non-Singular Objects in Natural Language Discourse", *Computational Linguistics*, 15, 3, 171-186.

Vander Linden, K. (1993). "Rhetorical Relations in Instructional Text Generation", in Proc.Workshop on "Intentionality and Structure in Discourse Relations", ACL, Columbus, pp.140-143.

Vander Linden, K. (1994). "Generating Precondition Expressions in Instructional Text", in Proc. 32nd Annual Meeting of the ACL, N.Mexico.

Woods, W. (1978). "Taxonomic Lattice Structures for Situation Recognition", *American Journal of Computational Linguistics*, 33-41.

Woods, W. and R.J. Brachman (1978). *Research in Natural Language Understanding - Quarterly Technical Progress Report No.1* (BBN Report No.3742), Bolt Beranek and Newman Inc., Cambridge, MA.

Quantified Noun Phrase Structure in Bulgarian [1]

Mila Dimitrova-Vulchanova and Giuliana Giusti
University of Trondheim - University of Venice

0. *Introduction*

In this paper we address some issues in the structure of Bulgarian noun phrases, in particular the constructions that arise with the enclitic article and quantifiers. In section 1. we give a preliminary account of DP syntax in Bulgarian, including movement of the noun and other lexical elements inside the extended nominal projection. In section 2. we present a general analysis of quantifier phrases across languages that will serve as the theoretical framework to be applied to the Bulgarian data presented and discussed in sections 3. and 4.

1. *Preliminary analysis of the Bulgarian DP structure*

In recent times, noun phrase structure has become the focus of much cross-linguistic research. In particular, the existence of enclitic articles in Scandinavian and

1. We would like to thank various people and institutions for giving us support in different respects during the research on which this paper is based. First of all we are intellectually indebted to Guglielmo Cinque and Lars Hellan for much more than just discussing the ideas presented here. Much is due to the University of Venice for granting Mila a research fellowship in the spring of 1994 that gave us the opportunity to meet and work together, and to NorgesforskningsrĆd for providing Giuliana with a travel grant to present the paper at the 3rd FASL. Last but not least we thank audiences at both Venice and College Park.
For the sake of requirements by the Italian Academy, Giuliana Giusti is responsible for sections 1. and 2. and Mila Dimitrova-Vulchanova for sections 3. and 4.

in most Balkan languages has been taken as evidence to support the hypothesis that D is an independent head selecting the noun phrase; cf. Hellan (1986) and Taraldsen (1990) for Norwegian ((1)), Dobrovie-Sorin (1987) and Grosu (1988) for Romanian ((2)):

(1) a. en gutt (a boy)
b. gutt-en (the boy)

(2) a. un băiat (a boy)
b. băiat-ul (the boy)

Bulgarian is the only [2] Slavic language in what is traditionally referred to as the Balkan *Sprachbund.* In non-modified noun phrases it patterns like Norwegian and Romanian above:

(3) a. momče ([a] boy)
b. momče-to (the boy)

However, there are some crucial differences among the three languages when the noun is modified by an adjective: in Norwegian the adjective is preceded by a free form of the article, while the noun still retains what looks like the enclitic article (4), in Romanian either the noun moves to D, thus preceding the adjective in the linear order (5b), or the adjectival head functions as the base for article incorporation (5c), Bulgarian only has this latter choice (6). [3] The empirical generalization about the

---

2. We consider Modern Macedonian as comprising a variety of dialects of the Bulgarian type, especially in view of basic common syntactic properties (cf Dimitrova-Vulchanova 1995).

3. Interestingly, some Bulgarian dialects (South-Western, along the Danube border and Bulgarian dialects spoken in Romania) display N-to-D movement (i):

(i) a. deteto miniko/slamata sitna
'child-the little/hay-the fine'
b. kustata tas visoka
'house-the this high'
c. kačeata onaa golemata
'barrel-the that big-the'

Proposals have been made (Mladenov (1993)) that the presence of these patterns in some of the above dialects is largely due to a language-contact situation and, consequently, the result of influence from Rumanian. However, the pattern in c. differs from the Rumanian one in that the AP receives the

placement of the article can be stated in the following way: The article morpheme is incorporated into the first head in the DP (e.g. into the first adjective. If the adjective is modified by an adverb the article will still go on the adjective [4]) (cf. Penčev 1993 for a slightly different formulation).

(4) a. den store gutten
the big boy
b. *gutten store
c. *storen gutten

(5) a. *cel mare băiat(ul)
b. băiatul mare
'boy-the big'
c. marul băiat
'big-the boy'

(6) a. goljamo-to momče
b. *momče-to goljamo
the big boy

The variation found in (4)-(6) suggests that the bound nature of the article is no evidence *per se* for N-to-D movement, as it is impossible in Norwegian or Bulgarian, and only optional in Romanian. The trigger for noun movement, therefore, must be some other property. Although we do not go deep into this problem, we tentatively

---

definite article, too. To what extent these dialects have been affected by the Rumanian system is immaterial to our discussion. The crucial point is that the very same dialects also have intermediate N-movement as shown in (ii):

(ii) a. meso pečeno/leb dobur
'meet grilled/bread good'
b. edna babiček a durta/edna svinja diva
'one/a granny old / one/a swine wild'

4. Notice that Bulgarian neither has an indefinite article (as shown in (3a)), nor a free form of the definite article, comparable to Scandinavian *den/det* and to the Romanian adjectival article *cel*.

establish the correlation between N-to-D movement and intermediate N-movement.[5,6]

Cinque (1994) analyses the difference in the adjective-noun word order in Romance and Germanic languages in terms of partial N-movement:

(7) a. der grosse Knabe/*der Knabe grosse
b. the big boy/*the boy big
c. le grand garçon /le garçon grand
d. il grande ragazzo/il ragazzo grande

Giusti (to appear) proposes that N-to-D movement is possible only in those languages that display partial N-movement, in compliance with the Head Movement Constraint. Its impossibility in Scandinavian and Bulgarian is therefore expected. Thus the ungrammaticality of (4b) is reduced to the ungrammaticality of (8b), while the variation between (5b,c) is reduced to the variation in (9a,b).

(8) a. en stor gutt (a big boy)
b. *en gutt store (a boy big)

(9) a. un mare băiat (a big boy)
b. un băiat mare (a boy big)

Here we will refrain from discussing what the ultimate trigger for the intermediate N-movement in Romance could be. Whatever this is, it is a necessary although possibly not sufficient condition for N-to-D movement.

---

5. V-to-C movement in the Mainland Scandinavian languages is the only case we know of movement of a lexical head to a high functional projection, in a (group of) language(s) that do not display the corresponding short movement (in that case V-to-I). The crucial difference between verbs and nouns is that while the modifiers of verbs (adverbials) are of completely different nature and, as a consequance, cannot fulfill the function that triggers V-to-C movement, the modifiers of nouns, namely adjectives share with nouns the possibility of bearing nominal morphology, in our case the article, they therefore compete with the noun in the possibility of moving to a position in DP. We will turn to the hypothesis that it is the economy of derivation that requires the shortest move to fulfill the function of DP.

6. Movement of N to an intermediate functional projection has been proposed in the literature to account for word order variations in noun phrases cross linguistically, cf. Ritter (1988), Picallo (1991), Cinque (1993) among others.

Having established a relation between the absence of N-to-D movement and the absence of partial N-movement in Bulgarian and Scandinavian, there still remains an important distinction between these two (groups of) languages to be accounted for. Namely, the different strategies that are employed to realize the article, which is a bound morpheme on the adjective in Bulgarian and a free morpheme preceding the adjective in Scandinavian. We tentatively propose analysing this difference as arising from different properties of adjectival morphology.

There are strong reasons to believe that in Bulgarian, the adjective in fact inflects for "definiteness" [7] as reflected by a different form of the article depending on the morphological properties of the root it appears on ((10)).

(10) a. xora-ta/*te (people-the)
b. dobri-te/*ta xora (good-the people)

Evidence for the hypothesis of analysing the article as the internal morphology of the adjective is provided by the fact that the article appears on the adjectival head regardless of whether it has a modifier or a PP-complement. An analysis in terms of A-to-D movement, which predicts (11b), is excluded and so is an analysis of phonological encliticization of D onto an AP in SpecDP, which predicts (12b), (13b):

(11) a. mnogo xubavi-te knigi
'very nice-the books'
the very nice books
b.*xubavi-te mnogo knigi
nice-the very books

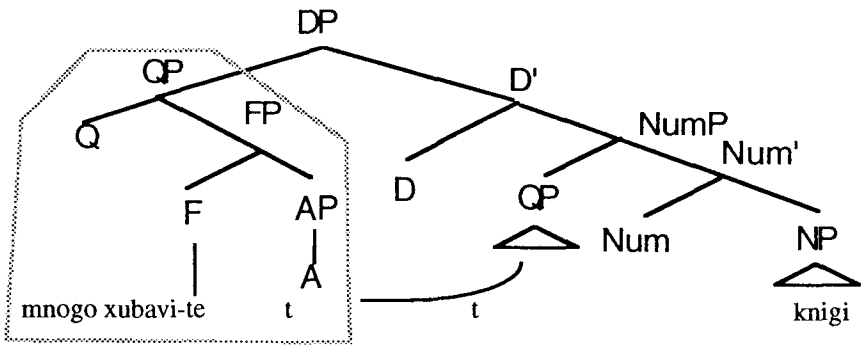
(12) a. kupeni-te včera knigi
'bought-the yesterday books'
the books bought yesterday
b. *[kupeni včera]-ta/te knigi
'bought yesterday-the books'

7. Under "definiteness" we mean the abstract features expressed by the definite article, whatever their nature and language particular realization could be.

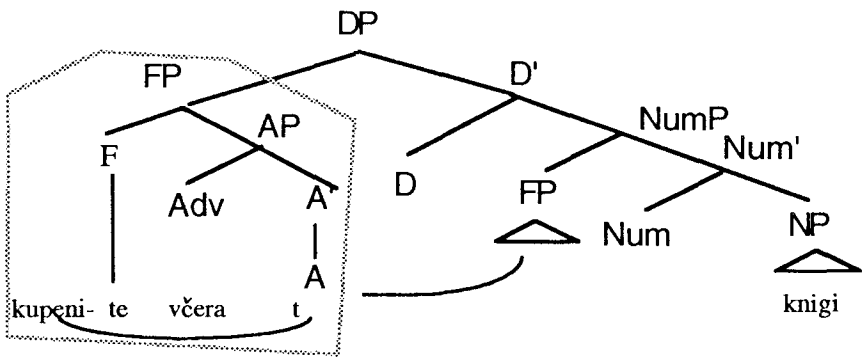
(13) a. vernij-at na žena si muž
'truthful-the to wife poss refl man'
the man truthful to his wife
b. **[veren na žena si]-ta/jat muž

The structure we propose for the modified noun phrase is (14a) for (11a), (14b) for (12a), and (14c) for (13a): [8]

(14) a.

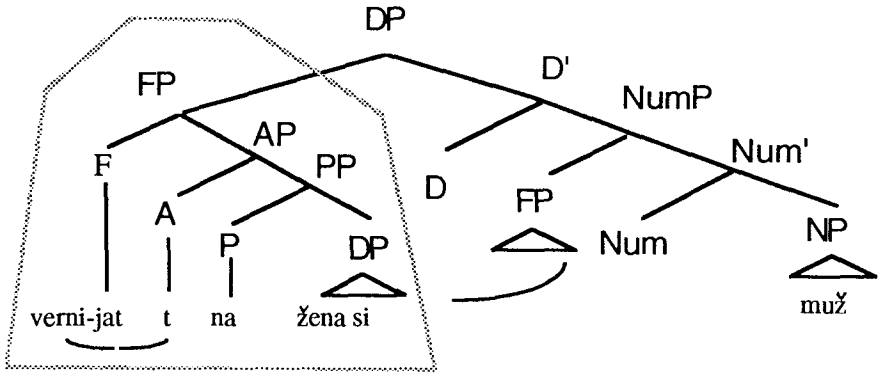


b.



8. At this point of the reasoning, the internal structure of adjectival phrases and their functional projections is irrelevant, since our analysis will be limited to the high periphery of the noun phrase. We assume that *mnogo* in (11a)=(14a) is a Q selecting an extended adjectival projection, parallel to what we are going to propose in section 2. for quantifiers selecting noun phrases.

c.



We assume that the mechanism at work here is checking the features in DP by Spec-Head Agreement of the inflected adjectival phrase moved to Spec DP and the head D. Movement of the highest adjectival phrase is just one step movement, it is therefore preferred to N-to-D movement. Due to lack of independent intermediate N-movement, N-to-D movement in Bulgarian has to take place in as many steps as there are functional heads. This is not the case in Romanian, where N is independently moved to an intermediate functional head (that we take to be Num° here for expository purposes). In Romanian, AP-to-SpecDP is in perfect competion with N-to-D, in that it requires the same number of steps. The contrast between (5b) and (6b), in this way, is reduced to the principle of economy of derivation, along the lines of Chomsky's (1992) recent proposals.

So far, we have briefly outlined a general structure for DP in Bulgarian, which is going to be the background for our analysis of quantified noun phrases in 3. and 4. below.

2. *Quantified noun phrases: a general analysis*

Before discussing the Bulgarian data we present below an independent hypothesis about the general structure of QPs across languages.

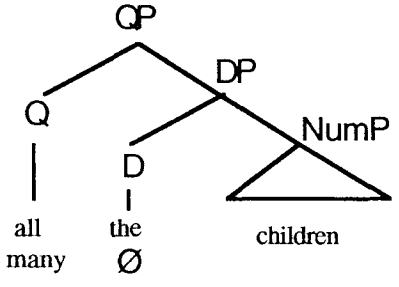
Giusti (1991) and following work, resting on cross-linguistic considerations based on contrastive analysis of some Romance and Germanic languages, suggests analysing the two occurences of the quantifier in (15) as having a different syntactic status. In (15a) the quantifier is a head selecting a DP as its complement, much in the

same way as the universal quantifier in (16a). In other words, the structure of (15a) includes and empty D position as represented in (16b). In (15b), on the contrary, the quantifier has the function of a modifier of the noun, much in the same way as the adjective in (17b).

(15) a. many children
b. the many children

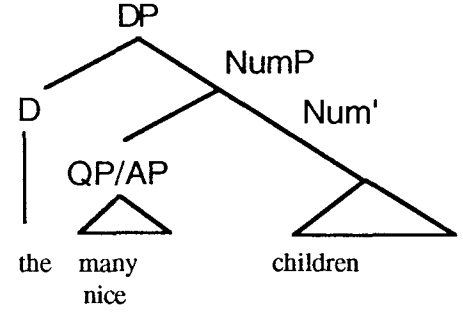
(16) a. all the children
b. many Ø children

c.



(17) a. The many children
b. the nice children

c.



The surfacing of an article in configuration (16) depends on the selectional properties of the quantifier: *many* selects a partitive DP, which must have a Ø determiner, while *all* selects a definite DP which displays a definite article in English.

This analysis was inspired by Romance data, where the definite article is obligatory after a universal Q. But it is more controversial in Germanic where the article may be missing. Consider the German examples in (18)-(19):

(18) a. all(e) die Kinder
b. all*(e) Kinder
c. die ganzen/*allen Kinder
all the children

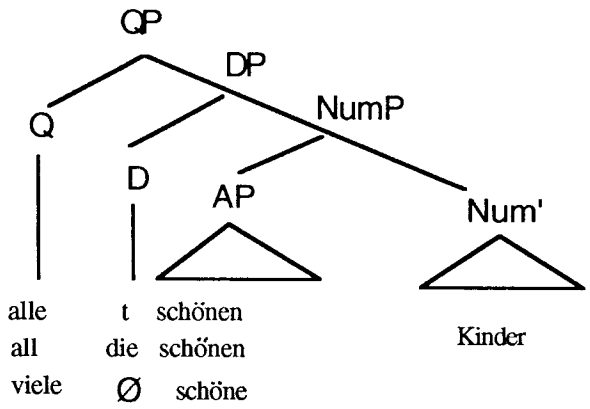
(19) a. beide (*die) Kinder
b. die beiden Kinder
both children

In (18), the quantifier *alle* is optionally inflected for nominal features in case the article is present (18a), and is obligatorily so in case the article is missing (18b). (18c) shows that *all* cannot have modifier status, as it cannot be preceded by a determiner. Instead, German has a separate lexical entry: *ganz*, which specializes for this function. In (19a) the quantifier *beide* appears in a construction like (18b) and (16) above), while in (19b) it is arguably a modifier. In fact, it follows the article and displays weak inflectional morphology, which is typical of adjectives in this position, cf. *ganzen* in (18c).

The weak vs. strong inflection of adjectives following the quantifiers in (20a) and (20c), respectively, supports the hypothesis that there is a D head between the quantifier and the noun, which overtly sufaces in (20b). Our analysis is that in (20a) the article *die* is "incorporated" into the Q, while in (20c) it is zero (=indefinite plural article in German). The structure is given in (21):

(20) a. alle/beide schöne*(n) Kinder
'all/both nice-wk/*str children'
b. all die schöne*(n) Kinder
c. viele schöne(*n) Kinder
'many nice-str/*wk children'

(21)



In (21a) the trace of the incorporated article is in the same relationship to the AP as the lexical definite article in (21b). In fact, it triggers weak morphology on the A. In (21c), on the contrary, the features on the Q cannot be taken to be the result of incorporation of D into Q, since the adjective displays strong morphology as adjectives normally do when no article is present at all. [9]

It is conceivable that the incorporation in (21a) cannot take place if the DP is raised in a floating construction such as (22a), since the trace of the article in this case would not be preceded by its antecedent incorporated in the Q left *in situ*. This is why the article is obligatory. The inflectional morphology on the quantifier can be easily explained by assuming that DP has moved through SpecQP thus triggering agreement with Q. We turn to that shortly when discussing the Bulgarian data.

(22) a. die Kinder kenne ich all*(e)/beid*(e)
'the children know I all/both'
b. Kinder kenne ich viele
'boys know I many'

9. That such an incorporation of the article is possible in German is independently shown by the existence of inflected prepositions, such as *aufs* (*auf+das* = on + art(s., neut., acc.), *im* (*in+dem* = in + art(s., m./n., dat.):

(i) a. Q+D = all-e
b. P+D = e.g. auf-s, im, etc.

3. *Quantified Noun Phrases in Bulgarian*

The QP-hypothesis outlined in section 2. above incorrectly predicts that if *vsički* is a Q as in (16), it should be followed by a complete DP (Bulgarian (23)). On the contrary (24) is what we find:

(23) a. *vsički [knigi-te]
b. *vsički-te knigi-te
c. *vsički [xubavi-te knigi]
'all good-the books'

(24) a. vsički knigi
'all books'
b. vsički-te knigi
'all-the books'
c. vsički-te xubavi knigi
'all-the good books'

An analysis of *vsički* as a high modifier of the noun (cf. the analysis of English *many* as in (17) above; for Bulgarian cf. Penčev 1993) could explain (24b,c), but leaves (24a) unaccounted for. In fact universal QPs are only found in definite DPs across languages, and definite DPs ordinarily display the article in Bulgarian. In 3.2. we will show that Bulgarian *vsički* does not depart dramatically from its counterparts in languages like Romance and Germanic (cf. the German examples from above). Let us first consider in 3.1. the more straightforward cases represented by *mnogo/malko/njakolko* ("many/ few/ a few") and cardinals.

3.1. *'Mnogo'/'malko'/'njakolko' and cardinals*

It appears that a quantifier vs. AP distiction can provide an account for the distribution of existential quantifiers. We suggest that *mnogo* in (25a) is parallel to *many* in (16b) and in (25b) is parallel to *many* in (17a).

(25) a. mnogo (novi) knigi
many new books
b. mnogo-to (novi) knigi (v bibliotekata)
'many-the new books (in library-the)'

Notice that cardinals such as *dve/dva/dvama* ("two"), *tri/trima* ("three") apparently behave like *mnogo* in either selecting an indefinite complement or functioning as a high modifier: [10]

(26) a. dve (novi) knigi
two new books
b. dvete (novi) knigi
'two-the new books'

There is, however, an interesting difference between the two classes of quantifiers. Cardinals can occur lower in the structure with respect to descriptive adjectives, while other adjectival quantifiers cannot:

(27) a. novite dve knigi
b. *novite mnogo knigi

This can be captured under an analysis of cardinals as heads in Num. Evidence for postulating this position is the agreement for [+M, +count] features on the head noun triggered by cardinals but not by other quantifiers (cf. fn. 9). Being a head, the cardinal can be bypassed by an adjectival phrase moving to Spec DP, or move to D

---

10. Cardinals exibit the peculiarity of triggering a special agreement for [count] on masculine nouns, cf. (i) and (ii). On the other hand, if the masculine noun is specified for [+human), the cardinal, instead, appears in a special form, cf. (ii) and (iii):

(i)	a.	dve/tri knigi	(two/three books)
	b.	dve-te/tri-te knigi[-M]	(the two/three books)
(ii)	a.	dva stola	(two[M] chairs)
	b.	dvata stola	(the two chairs[COUNT])
(iii)	a.	dvama/trima muže	(two/three[hum, M.] men)
	b.	dvamata/trimata muže	(the two/three[hum, M.] men[PL])

itself. On the contrary *mnogo*, being a phrase blocks the movement of a lower phrase to SpecDP.

Cardinals in Bulgarian, therefore, highlight a property of the complex syntax of quantification that was not detected in Giusti (1991).

3.2. *The universal quantifier 'vsički'*

Bulgarian behaves like Romance and Germanic with respect to quantifier floating. Furthermore floating quantifiers appear to be linked to a complete DP in higher clausal position, as is the case in Romance and Germanic and as predicted by the hypothesis.[11]

(28) a. knigi-te gi pročetox vsički-te
'books-the them cl read1sg all-the'
b. die Bücher habe ich all*(e) gelesen

---

11. Note that the construction in (28a) represents a typical topicalization configuration in Bulgarian, which involves clitic doubling of the moved constituent. It is also the exact equivalent of the German in (28b). As expected, floating quantifiers are found also in passive constructions such as the restricted (i) and the *se*-construction in (ii):

(i) ?knigi-te bjaxa pročeteni vsički-te
'books the were read all-the'

(ii) knigi-te se pročetoxa vsički-te
'books the REFL read all-the'
the books were all read

Notice also that the quantifier is found in the basic post-verbal subject position, as in (iii):

(iii) momčeta izjadoxa po edna jabulka vsičkite
'boys-the ate PO one apple all-the'
the children all ate an apple

As independently argued for in Dimitrova-Vulchanova (to appear), the landing site of topicalized constitutents cannot be unambiguously analysed as either A or A´. Therefore, clitic doubling is not to be taken as a sign for dislocation. This is also true of basic vs. derived positions for subjects. Moreover, clitic doubling is related to the aspectual features of the clause, cf. Dimitrova-Vulchanova (1992) and Dimitrova-Vulchanova and Hellan (1994).

Notice the contrast with adjectives which never appear in discontinuous constructions. Compare (28) and (29):

(29) a. pročetox xubavi(te) knigi
'(I) read nice-the books'
b. *knigite gi pročetox xubavite
'book-the CL (I) read nice-the'

It deserves mention here that a closely related South Slavic language like Serbo-Croat, which has morphological case and no article, displays free left branch extraction of adjectives and possessives ((30)).

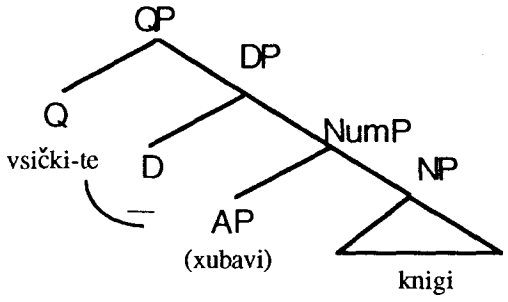
(30) a. Ivan kupuje zeleni auto
Ivan buys green car
b. Zeleni Ivan kupuje auto
'green Ivan buys car'
c. Ivan razbija tatin auto
Ivan ruins father's car
d. Tatin Ivan razbija auto
'father's Ivan ruins car'

Since this is clearly not the case in Bulgarian, the quantifier in (28a) cannot be taken as an adjective, contrary to what is suggested by Penčev (1993).

The comparison with German, instead, gives us some insight into the Bulgarian structure. The parallelism is almost perfect under the assumption that the article *-te* on the quantifier is a type of morphological agreement, much in the same way as the morphme *-e* on the German Q. In both cases this type of morphology is optional when the complement of the quantifier is in place and obligatory when the complement of Q is extracted. The difference with German is the possibility for Bulgarian to have an empty D in the complement of *vsički.* This can be related to the fact that *vsički,* contrary to *all* already bears some nominal features, namely Number and can therefore license an empty head and identify its features.

We propose that *vsičkite* in (24b,c) is the result of incorporation to the higher Q of the features of the DP, generated in D as in (31). We return shortly to the nature of these features.

(31)



The assumption that this incorporation is obligatory when the DP remains *in situ*, accommodates the ungrammaticality of (23). This mechanism can be reduced to some common principle of economy of derivation in that in Bulgarian, movement up to D or SpecDP is avoided whenever the article can appear on a higher element in the extended nominal structure.

Such an incorporation is impossible when the DP is extracted out of the QP, as was the case for German (22a)=(28b). The floating construction therefore highlights two important aspects of this construction which are otherwise obscured in the base construction. Namely that a) *vsički* is a head Q, and b) it selects a complete DP complement as its Germanic counterparts in (16).

Article doubling in (28a) is obligatory and can be analysed as agreement with the DP triggered by the movement of DP through SpecQP. Evidence that such movement takes place is provided by the possibility of the constituent [DP_i [Q t_i]] in (32b), which can actually move as such, as shown by (32d):

(32) a. *pročetox* [$_{QP}$ *vsicki-(te*$_i$) [$_{DP}$ (t_i) [$_{NP}$ *knigi*]]]
'(I) read all-the books'

b. pročetox [$_{QP}$ [$_{DP}$knigite] [$_{Q'}$ vsiciki-te [$_{DP}$ t_i]]]
'(I) read books-the all-(the)'

c. [$_{DP}$*knigite*]$_i$ *gi pročetox* [$_{QP}$ [t'$_i$] [$_{Q'}$ *vsicikite* [$_{DP}$ t_i]]]
'books-the CL (I) read all-the'

d. [$_{QP}$ [$_{DP}$knigite] [$_{Q'}$ vsiciki-te [$_{DP}$ t_i]]]$_j$ gi pročetox t_j
'books-the all-the CL (I) read'

(31) accounts for all the data presented so far. Of course, the *in situ* word order is ambiguous between the adjectival and the Q analysis. The adjectival analysis, however, does not account for the discontinuous construction.

3.2.1. *An alternative analysis*

As a matter of fact, it appears to be counterintuitive to analyse the realization of the article on *vsički* in the floating and the *in situ* construction as the result of two different processes. A possible alternative to this could be to take the article as agreement in all cases. Agreement would be obligatory in the Spec-Head configuration and in the extraction cases, and optional when the DP remains *in situ.* Optionality of agreement in the latter case could be explained in terms of the inherent definite nature of the DP selected by a universal quantifier. Optionality of agreement, however, is not enough to explain the impossibility of the examples in (23) above and (33a) below. A stipulation is needed in this analysis about the impossibility of realizing the features in DP in case the quantifier is present. Since we do not find a way of reducing this stipulation to any other more principled property of Bulgarian, we believe that the split analysis is superior to the unified one.

3.2.2. *Quantified pronouns*

Let us now substantiate the nature of the features realized by the article. In Giusti (1993) it was proposed that the article realizes Case features in German. We propose that it does so in Bulgarian as well. In fact, when the quantifier precedes a personal pronoun, which is intrinsically inflected for Case, it never displays such features. Also notice that the article is homophonous and diachronically related to the nominative form of the third person pronoun.

(33) a. [$_{QP}$ vsički (*-te) [$_{DP}$ nie/nas]]
vie/vas
te/tjax
'all(*-the) we/us'
you/youA'
they/them'

b. [$_{QP}$ [$_{DP}$ nie/nas] [$_{Q'}$ vsički (?-te)]
vie/vas all(?-the)
te/tjax

The data in (33) empirically justifies the double analysis of the article *-te* on *vsički* in the above examples. In case it is a pronoun, the complement of *vsički* cannot possibly include an article which is expected under our incorporation proposal and would not be accounted for by a unified analysis of the article as agreement with the complement. (33b) strongly suggests that the article is an instance of agreement with the complement moved into SpecQP.

4. *The interaction of quantifiers and high modifiers of the noun*

4.1. *Demonstratives*

Following Giusti (1992), we assume that demonstratives across languages are not in D but in a high Specifier and subsequently move to SpecDP [12], contrary to what has been implied in current literature on DP-structure (cf. Longobardi (1991) among many others). We apply this proposal to Bulgarian with the addition that in this language demonstratives are always found in SpecDP (either base generated there or obligatorily moved there overtly). Being intrinsically specified for definite features, *tezi* differs from an adjective in SpecDP in that it never takes the article, cf. (34a,c). However, it may, under certain conditions, co-occur with the article, as shown in (34b):

(34) a. tezi novi stolove
these new chairs

12. Giusti's claim is based on the observation that in Romanian, the demonstrative, which appears to be base generated as the leftmost modifier fo the noun (i), can be skipped by N-movement (ii), but not by AP movement (iii):

(i)	acest frumos băiat	(this nice boy)
(ii)	băiatul acesta frumos	(boy-the this nice)
(iii)	frumosul (*acesta) băiat	(nice-the this boy)

Parallel evidence is independently provided for Kiswahili by Carstens (1991).

b. tezi dva-(ta) stola
'these two-the chairs'
c. *tezi stolovete
'these chairs-the'

In (34b), the cardinal optionally takes the article. In other words, the article is optionally inserted in D° when SpecDP is occupied by a demonstrative and a cardinal is in Num°. If the article is inserted, the shortest move is Num-to-D. The article in this case inflects for the morphological features of the cardinal. The impossibility of (34c) clearly shows that the noun does not move in Bulgarian, as we have suggested above.

If *tezi* is taken to be in SpecDP, we expect the universal quantifier to precede it, as in (35a), unfortunately, what we have said so far is not sufficient to predict the possibility of (35b):

(35) a. vsički tezi knigi
b. vsički-te tezi knigi
'all-(the) these books'

(35b) would be expected under the unified agreement analysis in 3.2.1. above. Notice, however, that it does not contradict the incorporation analysis, if explained along one of the following lines: Either we take *tezi* in SpecDP to co-occur with the trace of *-te* in D left after incorporation, as in (36a); or we take *tezi* to be generated lower (in the Spec of a nominal functional projection that we generically label FP here) and stay there, in case DP already has a filled head, as in (36b):

(36) a. [Q-te$_i$ [DP tezi [t$_i$ [FP ...]]]]
b. [Q-$_i$ [DP [t$_i$ [FP tezi [F' ...]]]]]

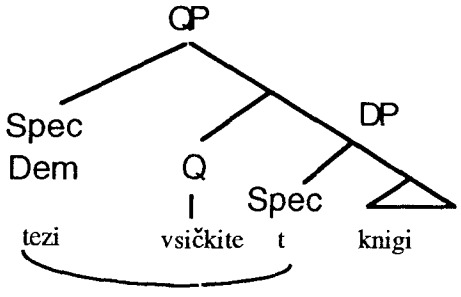
An apparent further problem to our approach is (37a). In fact, if we take the demonstrative as marking the DP-boundary, the quantifier is not expected to follow it unless it has adjectival status. However, an adjectival analysis of *vsički,* parallel to the analysis suggested above for *mnogo/njakolko* is contradicted by the obligatory

occurrence of the article on *vsički* in this construction, since ordinary adjectives preceded by a demonstrative never display an article, as shown in (37b): [13]

(37) a. tezi vsički *(-te) knigi
'these all*(-the) books'
b. tezi xubavi(*te)/njakolko(*to) knigi
'these good-(*the)/few-(*the) books'

In line with our analysis of demonstratives above, we propose that (37a) is derived by further movement of *tezi* from SpecDP to SpecQP. The structure is given in (38). The obligatory presence of the article on *vsički,* is Spec-head agreement for features:

(38)



In addition to being theoretically justifiable, the structure in (38) appears to provide for a Topic-Focus distinction between the quantifier and the demonstrative, depending on their respective linear order. In (35) the demonstrative has a focused reading, whereas in (37) it is the quantifier which falls under focus. This can be taken as an instance of syntactic structure independently serving discourse structure. Note that the Topic-Focus effect found in QP as part of the extended nominal projection is very similar to the same effect in Bulgarian clause structure.

13. Taking *vsički* to be in Num in this case will not be justified either, since it behaves differently from cardinals in the same position, cf.:

(i) novite dve knigi (new-the two books)
(ii) novite vsički knigi (new-the all books)

4.2. *Possessives*

This analysis can also capture the data in (39)-(40). In (39a) and (40a) we see a complete DP with no quantifier. In (39b-c) and (40b-c) the presence of the quantifier blocks movement inside DP. The d-examples represent the floated construction with the complete DP extracted out of QP, and in this case the extracted constituent is identical to the non quantified DP in the a-examples.

(39) a. knigi-te mi
'books-the my Dcl'
b. vsički-te mi knigi
'all-the my Dcl books'
c. *vsički (-te) knigi-te mi
d. Knigi-te mi izgorjaxa vsički *(-te)
'books-the my burned all*(-the)'

(40) a. moi-te knigi
'my-the books'
b. vsički-te moi knigi
'all-the my books'
c. *vsički moi-te knigi
d. moi-te knigi zgorjaxa vsički te
e. moi-te vsičkite knigi

This shows that incorporation of the article takes place in the adjacent position regardless of what type of DP is embedded into QP. Here we will not pursue the analysis of possessive constructions in Bulgarian. We only briefly note that we consider the constructions with the possessive pronominal adjectives and the ones with a dative possessive clitic as representing two distinct types and consequently as structurally different.

5. *Conclusion*

In this paper, we have sketched some proposals for DP structure in Bulgarian. In particular, we have argued for the following points:

a) There are two necessary conditions for N-to-D movement across languages, one is the enclitic nature of the article and the other is independent N-movement to the immediately higher nominal functional head. Neither of them is sufficient on its own. It is only their interaction that appears to be able to trigger this phenomenon.

b) The article on the prenominal adjective in Bulgarian arises in a functional projection of the adjective itself and not in D. The inflected AP is moved to SpecDP and checks the features in D.

c) Quantifiers in Bulgarian have been shown to behave in a way parallel to Romance and Germanic despite appearances. In particular, Bulgarian has highlighted the existence of cardinal insertion in Num; the possibility for SpecQP to host the complement of Q or a demonstrative.

d) Finally, the distribution of the article on *vsički* was analysed as the incorporation of Case features of the DP generated in D in case the complement of Q is *in situ* and as agreement for the same features when the complement is moved to or through SpecQP.

References

Carstens, V. M. (1991). *The Morphology and Syntax of Determiner Phrases in Kiswahili*, Ph.D. Diss. UCLA.

Cinque, G. (1994). "On the Evidence for Partial N-movement in the Romance DP" in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, R. Zanuttini (eds.) *Paths Towards Universal Grammar. Studies in Honor of Richard Kayne*, Georgetown University Press, Washington, D.C.

Chomsky (1993). "A minimalist program", in K. Hale and S. J. Keyser (eds.) *The View from Building 20* MIT Press, Cambridge, Mass.

Dimitrova-Vulchanova, M. (1993). "Clitics in Slavic", *EUROTYP Working Papers Theme Group* 8, vol. 4..

Dimitrova-Vulchanova, M. (1994). "On Subjects in Bulgarian", ms., Trondheim University.

Dimitrova-Vulchanova, M. and L. Hellan (1994). "Clitic in Bulgarian Clause Structure", ms., Trondheim University.

Dobrovie-Sorin, C. (1987). "A propos de la structure du goupe nominal en Roumain", *Rivista di Grammatica Generativa*, 12:123-152.

Giusti, G. (1991). "La sintassi dei nominali quantificati in romeno", *Rivista di Grammatica Generativa*, 16, 29-57.

Giusti, G. (1992). "Heads and Modifiers among Determiners", *University of Venice Working Papers in Linguistics* 2.I.4., to appear in Cinque, G. and Giusti, G. (eds.) *Comparative Studies in Romanian*, Linguistik Aktuell, Benjamins, Dordrecht.

Giusti, G. (1993). *La sintassi dei determinanti*, Unipress, Padova.

Giusti G. (to appear). "Enclitic Articles and Double Definiteness. A Comparative Analysis of Nominal Structure in Romance and Germanic", to appear in Lorentz, O. and T. Taraldsen *Proceedings of the Workshop in Germanic Syntax, Tromsø Norway, Nov. 1992*, special issue of *The Linguistic Review*.

Grosu, A. (1988). "On the Distribution of Genitive Phrases in Rumanian", *Linguistics* 26, 931-949.

Hellan, L. (1986). "The Headedness of NPs in Norwegian", P. Muysken abd H. van Riemsdijk (eds.) *Features and Projections*, Foris, Dordrecht.

Longobardi, G. (1991). "Proper Names and the theory of N-Movement in Syntax and Logical Form", *University of Venice Working Papers in Linguistics* 1, I, 9.

Mladenov, M. (1993). *Bulgarskite govori v Rumunija*, Bulgarian Academy of Sciences Publishing House, Sofia.

Penčev, I. (1993). Bulgarski Sintaksis. Upravlenie i Svurzvane, Plovdiv.

Picallo, C. (1991) "Nominals and Nominalization in Catalan", *Probus*, 3, 279-316.

Ritter, E. (1988). "A Head-Movement Approach to Construct State Noun Phrases", *Linguistics*, 26, 909-929.

Stojkov, S. (1993). *Bulgarska dialectologija*, Bulgarian Academy of Sciences Publishing House, Sofia.

On the Denotation and Scope of Indefinites

Carmen Dobrovie-Sorin
UA 1028-CNRS, Université Paris 7

The main goal of this paper is to establish and explain correlations between the denotation and the scope of indefinites. The first section is devoted to the widely discussed ambiguity of indefinites. Evidence will be given in favour of more fine-grained distinctions than those available in the current literature. The analysis will lead us to revise standard GB assumptions concerning the mapping between S-structure and LF: (a) QR is dispended with; (b) a rule of DR (Determiner raising) is assumed, with two possible landing sites (adjunction to the DP itself and adjunction to VP); (c) the LFs thus obtained constitute the input to Interpretive Conventions: (i) translation procedures by which open DPs are mapped into variables or quantifiers; (ii) insertion of quantifiers in their unmarked scope position, i.e. adjunction to IP. The output of these Interpretive Conventions, which will be referred to as LF', must be distinguished from LF. In Section 2, it will be shown that certain constraints concerning scope assignment follow as consequences of the constraints established in section 1, concerning the type of denotation indefinites may take.

0. *Introduction*

The main goal of this paper is to establish and explain correlations between the denotation and the scope of indefinites. The first section is devoted to the widely discussed ambiguity of indefinites. In line with Fodor and Sag (1981) and Diesing (1992), evidence will be provided in favour of the idea that it is not possible to reduce the various readings of indefinites to one semantic interpretation, out of

University of Venice
Working Papers in Linguistics
vol. 5, n.1; 1995

which the others would be obtained by means of pragmatic procedures.

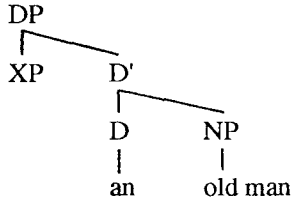
Evidence will be given in favour of more fine-grained distinctions than those available in the current literature. It will be shown that Fodor and Sag's (1981) distinction between referential and quantificational indefinites cannot be reduced to Diesing's (1992) opposition between presuppositional and existential indefinites. In addition, the recent literature has revealed the existence of still another reading, characterized as "amount/detranzitivized" (see de Hoop (1992), Dobrovie-Sorin (1992, 1993)). In sections 1.1. and 1.2.1 it will be shown that Fodor and Sag's referential reading relies on the presence of closed DPs at LF; quantificational indefinites, which correspond to open DPs, may be assigned several types of readings (existential, presuppositional or amount/detranzitivized), as a result of the application of distinct Interpretive Conventions (see sections 1.2. through 1.5.), which are partially determined by (i) the position of the indefinite DP [1] at S-structure and (ii) the position of the determiner at LF.

The analysis to be developed below will lead us to revise standard GB assumptions concerning the mapping between S-structure and LF: (a) QR is dispended with; (b) a rule of DR (Determiner raising) is assumed, with two possible landing sites (adjunction to the DP itself and adjunction to VP); as a result of DR, we obtain open DPs at LF and correlatively, the nonapplication of DR yields closed DPs at LF; (c) the LFs thus obtained constitute the input to Interpretive Conventions of several types, similar in spirit to the mechanisms proposed in Heim (1982): (i) translation procedures by which open DPs are mapped into variables or quantifiers, depending on their position [2] and on the semantic properties of their determiners; closed DPs, on the other hand, translate as constant terms; (ii) insertion of quantifiers in their unmarked scope position, i.e. adjunction to IP (see Heim's rules of Quantifier Construal and existential closure). The output of these Interpretive

---

1. The label DP, which abbreviates Determiner Phrase, is a more adequate notation for the constituents that used to be referred to as NP. The new label is meant to indicate that nominal constituents are functional projections headed by the determiner, which take a lexical projection, NP, as a complement:

(i)



2. Since QR is dispensed with, S-structure and LF positions are identical for DPs.

Conventions, which will be referred to as LF', must be distinguished from LF, because the Interpretive Conventions are not movement rules, nor binding relations, and as such they are not governed by syntactic constraints. This means that LF' does not belong to the Syntactic component (whereas LF does), but rather to the syntactic level of the Interpretive component.

In Section 2, it will be shown that certain constraints concerning scope assignment (see Liu (1990)) follow as consequences of the constraints established in section 1, concerning the type of denotation indefinites may take. The explanation to be proposed will proceed by establishing and accounting for the following generalizations: (a) strong quantifiers can take a "wide scope" reading irrespective of their S-structure/LF position; (b) referential indefinites can take a "wide scope" reading irrespective of their S-structure/LF position. It will be shown that the wide-scope reading of referential indefinites is not obtained via a rule of quantifying in (or QR), but is due to the cumulative-distributive reference property of individuals (see Link (1983), Landman (1989)); (c) presuppositional indefinites can take wide scope; (d) since it is not distributive, the existential quantifier cannot take wide scope; it can only take narrow scope or an independent reading; (e) the amount/detranzitivized reading of indefinites involves no scope-assignment for the whole DP; the determiner itself is necessarily narrow-scoped.

1. *LF Representations for Indefinites*

In section 1.1. the "referential" reading of indefinites will be isolated and defined. In section 1.2.1. an explanation will be proposed for why this reading is forced by prepositional Accusatives in Romanian. The central hypotheses on which this explanation relies are: (a) DPs function as either "referential", i.e., as constant terms, or as quantifiers, depending on whether they are "closed" or "open" at LF; (b) the status of "open" DP depends on the application of the LF rule of DR (Determiner Raising) proposed in Dobrovie-Sorin (1992, 1993). Section 1.2.2. deals with the existential reading, for which I assume Heim's (1982) mechanism of existential closure. In section 1.3. it is proposed that the amount/detranzitivized reading (see de Hoop (1992), Szabolcsi and Zwarts (1990, 1993), Dobrovie-Sorin (1992)) relies on DR landing in a VP-adjunction position. Weak-island sensitivity distinguishes

between this reading and all the others, in particular the existential reading. Sections 1.4. and 1.5. are devoted to the presuppositional reading.

1.1. *The referential and existential readings*

Let us consider the following systematic contrasts between Romanian prepositional accusatives (i.e., direct objects preceded by the dummy preposition *pe*, compare *a* in Spanish, cf. Jaeggli (1982) [3]) and unmarked direct objects.

Prepositional accusatives do not show Weak Cross Over effects

(1) a. ?* Mama ei$_i$ va ajuta [una din studentele tale]$_i$. (WCO)
her mother will help one of your students
b. Mama ei$_i$ (o_i) va ajuta **pe** [una din studentele tale]$_i$. (no WCO)

Modified cardinals cannot take the prepositional accusative

(2) a. *I-am examinat **pe** mai mult/ **pe** mai puţin de trei elevi.
[I] them-have examined 'pe' more/ 'pe' less than three students.
b. *I-am examinat **pe** între trei şi zece elevi.
[I] them-have examined 'pe' between three and ten students.
c. *I-am examinat **pe** mai mulţi băieţi decît fete.
[I] them-have examined 'pe' more boys than girls.

The donkey reading is blocked by prepositional accusatives

(3) Toţi profesorii care l'au cunoscut **pe** un student excepţional nu-l uită.
All the professors who him have known 'pe' an exceptional student don't forget him.

---

3. The analysis proposed here for Romanian does not directly extend to Spanish. It seems that *a* Accusatives favor but do not force the referential reading.

Prepositional accusatives cannot take narrow scope (see (4)b and (5)b)

(4) a. i. Fiecare profesor va examina zece elevi.
Each professor will examine ten students.
ii. Toţi judecătorii cunosc doi avocaţi.
All the judges know two lawyers.
iii. Toţi copiii au văzut un actor celebru.
All the children have seen awell-known actor
b. i. Fiecare profesor îi va examina **pe** zece elevi.
ii. Toţi judecătorii îi cunosc **pe** doi avocaţi.
iii. Toţi copiii l-au văzut **pe** un actor celebru.

(5) a. Ieri n-am examinat mulţi elevi.
Yesterday not (I) have examined many students.
b. Ieri nu i-am examinat ***pe mulţi elevi***
Yesterday not-them (I) have examined 'pe' many students.

In the examples in (4)a each professor/judge/child may examine/know/see different students/lawyers/actors. The examples in (4)b, on the other hand, talk about unique groups of students, lawyers and actors. The example in (5)a may mean "Yesterday I examined few students", whereas the example in (5)b can only mean "Many students are such that I haven't examined them yesterday".

Prepositional accusatives can be clitic-doubled

(6) a. Am întîlnit **pe** un prieten.
I met 'pe' a friend.
b. L-am întîlnit **pe** un prieten. (clitic doubling)
I him met 'pe' a friend.

The special behavior of prepositional accusatives can be directly explained if we assume that they are referential, where "referential" is defined as in (7); plural individuals are defined as in Link (1983) and Landman (1989):

(7) A referential DP is a DP that functions as a constant term: it denotes an (atomic or plural) individual.

The paradigmatic example of referential DPs being proper names, it is easy to check that indeed proper names (a) are not sensitive to WCO; (b) cannot take narrow scope (cannot be dependent on another QP); (c) clitic-doubling is allowed (in fact obligatory in present-day Romanian):

(8) a. His$_i$ mother will help John$_i$.
b. Every professor will examine John.
c. To**ţ**i profesorii îl$_i$ vor examina **pe** Ion$_i$.
All the professors him will examine 'pe' John.

Since the correlation between referentiality (defined as in (7)) and the properties illustrated in (8)a-c is quite obvious, I will not discuss it any further. The relation between the referentiality of indefinites and the absence of the donkey reading is also straightforward: as Heim (1982) has demonstrated, the donkey reading depends on the indefinite being translated as a variable, and this conflicts with referential indefinites, which function as constant terms. Somewhat more interesting is the impossibility shown in (2). The semantic distinction between non-modified and modified cardinals is well-known: non-modified cardinals may denote principal filters [4] whereas modified cardinals cannot. *A fortiori*, modified cardinals cannot be referential, i.e., they cannot denote (plural) individuals. It is on the other hand well-known that although a principal filter is not necessarily a constant term, constant terms are principal filters: proper names are principal filters. Returning to the examples in (2), their ungrammaticality is due to the fact that the presence of prepositional Accusatives - which forces referentiality - is incompatible with modified cardinals - which are not principal filters and therefore not referential -.

---

4. The notion of principal filter is defined within the theory of generalized quantifiers as follows:

> A generalized quantifier GQ denotes a principal filter if there is a set of individuals A such that for any set of individuals X, X $\in$ GQ iff A $\leq$ X.

What this definition says is that a GQ denotes a principal filter if all the sets of individuals which belong to its denotation include the set A, their generator. E.g., {John} and {x: man(x)} are included in the denotation of the GQs denoted by *John* and *every man* , respectively. No such set can be found for *no man, at least two men, exactly three men*, etc. Non-modified cardinals such as *two/three men* may denote principal filters if they are understood to denote the set of all sets that include unique sets of two men and three men, respectively.

1.2. *LF Representations*

1.2.0. *The rule of DR and the denotation of Determiners*

In 1.1. it has been established that Romanian prepositional Accusatives show certain properties characteristic of referential DPs. I will now try to propose an explanation for this observation. The main idea to be exploited is that the difference between referential and nonreferential DPs relies on an LF distinction, between "closed" and "open" DPs (Determiner Phrases).

(9) Closed DPs function as constants (they denote (singular or plural) individuals).
The determiner indicates the cardinality of the plural individual constant.

Before going into the details of our empirical analysis, let us see how (9) relates to our current understanding of Determiners. According to Higginbotham (1985, 1987), they saturate the NP-predicate, i.e., they bind the free variable related to it. It is quite natural to assume that by saturating a predicate, what we get is an individual that instantiates the predicate. And this is what (9) says. However, the semantic literature of the last decade (see in particular Keenan and Stavi (1986)) has taught us that determiners denote relations between two predicates (expressed by the NP and by the clausal predicate). These two hypotheses concerning the role of determiners can be reconciled by saying that determiners may function in two distinct ways, depending on the position they occupy at LF. In case they stay in their base position (sister of NP), determiners function as saturators, i.e., they denote functions that apply to a predicate (or the corresponding set of individuals) and return an individual. This is precisely what (9) says (a "closed DP" is a DP in which the determiner occupies its base position). But determiners may also raise at LF (see the rule of DR (Determiner Raising) proposed in Dobrovie-Sorin (1988, 1992, 1993)), and in this case they denote relations between sets. The correlation between open/closed DP at LF and their distinct semantic function seems natural: in closed DPs the determiner must apply (by the rules of functional composition) to its sister node, i.e., the NP constituent, and therefore it cannot function as a relation between sets; it can only be a saturator.

Closed DPs thus function much like proper names, which in many languages are characterized by the absence of determiners; in those languages where determiners do appear with proper names, their syntactico-semantic behavior is quite distinct from that of standard determiners (see Longobardi (1993)), and it would be interesting to show that they then behave as the determiners of closed DPs.

The rule of DR can then be viewed as being triggered by the semantic features of determiners: in the "unmarked" case, they denote relations between sets, and this forces DR to apply, because otherwise determiners can only apply to the predicate expressed by the NP. Now, since we need two sets, we must raise the determiner high enough in order to be able to apply λ-abstraction over the A-position that immediately dominates it, and thus obtain the set corresponding to the clausal predicate:

(10) DR adjoins the determiner to the first XP constituent that dominates it and occupies an A-position.

This definition differs from that in Dobrovie-Sorin (1992, 1993), insofar as I do not assume that the landing site of DR is necessarily a VP-adjunction or IP-adjunction position. Although DR may land in a VP-adjunction position (see 1.3.), it can also be a much more local rule: it may adjoin the determiner to the DP constituent itself.

1.2.1. *The referential reading*

Coming back to prepositional accusatives, we want to explain why they are necessarily interpreted as referential. Given (9) and (10), what we have to show is that the rule of DR is blocked by the presence of the preposition: hence, prepositional Accusatives are necessarily "closed" DPs at LF, and by virtue of (9) they can only function as constants, i.e., are "referential".

Let us then consider the possible LF representations of the examples in (11) and (12):

(11) Am întîlnit doi prieteni. (ambiguous)
[I] met two friends.

(12)	(I)-am întîlnit **pe** doi prieteni.	(referential)
	[I] (them)-have met 'pe' two friends.	
(11^1)	[S [S am întîlnit [DP*[doi$_i$]* [**DP** e_i prieteni]]	DR
(12^1)	◊[S [S am întîlnit [PP*[doi$_i$]* [**DP** e_i prieteni]	DR
(12^2)	[S [S am întîlnit [PP pe [**DP** *doi$_i$* prieteni]]]	no DR => close dDP

The LFs in (11^1) and (12^1) are obtained by DR. Since in (12) the object position is occupied by the PP, the determiner has to adjoin to the PP (see (10)). As indicated by the diamond, (12^1) is illicit, due to an ECP violation: since the trace of the raised determiner is not a th-role bearer, the ECP cannot be satisfied by antecedent-government; antecedent-government does not hold either, because DR crosses a barrier (the DP set in bold characters, which is not L-marked by *pe*). [5, 6] Compare the example in (11), without *pe*: here, the object DP is L-marked by the verb, and therefore DR does not cross any barrier; the resulting LF contains an open DP, which receives a nonreferential type of denotation (see 1.2.2.).

Coming back to the example in (11), it can only be assigned the representation in (12^2), which contains a closed DP. Given the interpreting convention in (9), the corresponding reading is the referential one. We thus reach the conclusion that prepositional accusatives force the referential reading. [7]

To be more precise, to say that *two friends* in (12) is referential is to say that it denotes the plural individual (in the sense of Link (1983)) which is the sum of two atomic individuals which are (the speaker's) friends: *[[DP two friends]] = friend1 + friend2* (the double brackets are to be read "denotation of"). The individual-type denotation makes it possible to check the truth values of (12) by checking whether the individual *two friends* belongs to the predicate obtained by λ-abstracting over its position. In other words, (12) is true iff [[DP two friends]] ∈ λx: I met x (or, by

---

5. The fact that the DP is not L-marked is enough to rule (12^1) out. But note that extraction out of DPs obeys an even stronger requirement, namely canonical government by the verb (see Pollock (1989)).

6. The proposed analysis of prepositional accusatives makes the following interesting prediction for a language that would have both prepositional accusatives and an S-Structure rule of DR like that found in French (*J'ai beaucoup$_i$ lu e$_i$ de livres 'I have much read of books'):* the latter should be allowed with non-prepositional accusatives, but not with prepositional accusatives. This is exactly the case of Bengali, as I was informed (p.c.) by Dasgupta (see Dasgupta (1988)).

7. Certain speakers reject examples like (12) altogether. We must assume that for such speakers, weak quantifiers are obligatorily subject to DR; for those speakers who accept (12), weak quantifiers may be left in their base position. Thus, our main point, the idea that the nonreferential reading depends on DR, is not invalidated by the acceptability judgments of those speakers who reject (12).

using another notation for sets, (12) is true iff [[DP two friends]] ∈ {x: I met x}).

What is then the denotation-type of the cardinal determiner *two* in (12)? Assuming the standard rules of type-theory, *two* is of type <<e,t>, e>, because it applies to a predicate, which is an element of type <e,t>, and it returns an individual, of type e. This is precisely the characterization of choice functions given in Reinhart (1992). While the present account is similar to Reinhart's by the type associated with the determiner, the two analyses differ with respect to the syntax-semantics mapping. For Reinhart, the (choice) functional reading of determiners relies on the possibility of translating the determiner as a choice-function. Under the present proposal, this possibility is not completely free; rather, it depends on the internal structure of DPs at LF: determiners of type <<e,t>, e> are those determiners that appear inside closed DPs at LF. In other words, the type of the determiner depends on whether it raises or not at LF.

It is easy to show that the referential reading differs from both the presuppositional and the partitive readings:

(13) a. At least two students are intelligent.
b. I examined at least two of your students.

Intelligent is an individual-level predicate, which as such forces the presuppositional reading (see Kratzer (1989), Diesing (1992)). This means that *at least two students* in (13)a has the presuppositional reading (for a characterization of this reading see 1.4. below). However, this DP is not a principal filter, and therefore it cannot be referential, as already established above. A similar observation can be made for (13)b: DPs of the form *Det of DP* are partitives, and they appear to be compatible with modified cardinals, which rule out the referential reading.

1.2.2. *The existential reading*

Let us turn now to the example in (11), with an unmarked direct object. In this case, nothing blocks DR, hence the legitimate LF shown in (11[1]):

(11) Am întîlnit doi prieteni.
[I] have met two friends.

(11[1]) [S am întîlnit [DP[*doi*$_i$] [DP e_i prieteni]] DR

Our next task is to relate the LF in (11[1]) to the existential reading, which is available for (11): "there are two friends I met". This can be done by assuming a slightly revised version of Heim's (1982) Interpretive Conventions (note that Diesing's (1992) revisions of those conventions are not necessary under the present account). The modifications do not bear on the essence of Heim's procedure, they are refinements, which take into account the empirical data described here: (a) since we deal not only with singular indefinites, but also with cardinal indefinites, the interpreting procedure should provide a translation for the cardinal determiner; (b) assuming as we do here, that closed DPs function as constant terms, the procedure of variable-insertion can only concern open DPs. Given these revisions, Heim's Conventions can be restated as in (14):

(14) a. Insert variables in the position of open DPs and
b. i. Apply existential closure (whenever the open DP/bare quantifier is not inside the restrictive term of another quantifier).
ii. Interpret cardinal determiners as cardinality markers of the variable.
iii. Interpret NP as the range of the variable.

Certain remarks are in order here, concerning plural indefinites. I will assume that, along with plural individual constants, the vocabulary of our logical representation has plural individual variables: x^n should be read "an individual variable whose cardinality is $\geq n$". To apply existential closure to a variable x^n means to insert an existential quantifier $\exists x^n$ in an adjunction position to S.

Since the condition stated inside brackets in (14)(i) is satisfied, existential closure applies in (11[1]):

(11[1])' $\exists x^2$ [I met x^2 ^ friend (x^2)]

In the general case the cardinality marker may be any number, as well as any measure phrase (mass nouns can give rise to the existential reading).

The representation in (11[1])' underlies the existential reading of (11), which can be paraphrased as "there are two friends that I met". Such a paraphrase indicates that the existence of two friends is asserted rather than presupposed (see 1.4. below).

Keenan and Stavi (1986) and Keenan (1987) assume that each occurrence of "existential determiners" (indefinites and numerals) gives rise to the existential reading. The predicate "exists" (or "is an individual"), which can be predicated of all the entities in the model, is notated "1".

(15) a. A basic determiner is called *existential* iff it is always interpreted by an existential function, where

b. A function f from properties to sets of properties is *existential* iff for all properties p, q,
$p \leq f(q)$ iff $1 \leq f(q \wedge p)$

This means that to say that f is existential is to say that f(q)'s are p's iff f (q's who are p's) exist (are individuals). This definition establishes an equivalence relation between the truth conditions that can be assigned to a sentence of the form Det NP VP (NP is q and VP is p in Keenan's formula) and those assigned to a formula of the type "a set of individuals that are both NP and VP exists (or there is a set ...") and the cardinality of the set is at least Det". This comes to say that *two students wrote to me* is truth-conditionally equivalent to *there are two students who wrote to me.* Similarly, for indefinites appearing in object positions: *I met two students* is equivalent to *there are two students whom I met.* Note now that the equivalence formula does not reflect the S-structure configuration, in which the determiner is a sister of NP; therefore the function that it denotes should apply to the denotation of NP and return the denotation of the DP. Instead, the proposed equivalence says that the determiner expresses the cardinality of the set of individuals obtained by intersecting the sets of individuals denotated by NP and VP. My rule of DR establishes a derivational relation between the two formulas: in order to function as existential, i.e., in order to be able to apply to the intersection of the two sets, determiners must raise out of their mother DP at LF. The resulting configuration is then assigned an existential reading by means of the conventions stated in (14). In case DR is blocked, determiners do not function as existential functions, but rather as choice function (see section 1.2.1. above). Keenan and Stavi's definition of existential determiners should then be revised:

(15') a. A basic determiner is called *existential* iff it may be interpreted by an existential function.

b. An existential determiner can be interpreted by an existential function only if DR applies at LF.

Under their existential reading, indefinites are not referential, and therefore they show Weak Cross Over effects, give rise to the donkey reading, can take narrow scope and cannot be clitic-doubled; also, modified cardinals may take the existential reading (see the examples in (1)-(6)).

Note, however, that given our account, WCO should never arise: since indefinites have the possibility to remain closed, i.e., they may function as referential, they could resort to that possibility and thus avoid a WCO violation. In order to explain why WCO violations do arise with indefinites, we must assume that DR is obligatory unless it is blocked by the syntax. In section 1.2.0. above it has already been assumed that DR is triggered by the unmarked semantic features of the determiner, namely the fact that it denotes a relation between sets. A determiner that is blocked inside DP loses the possibility to denote a relation between sets. The absence of DR thus appears to be a marked option, imposed by the syntax, e.g., the presence of prepositional accusatives, which block the application of DR. In other words, the absence of DR is not a possibility to which the grammar resorts "freely". Hence, whenever DR is not blocked by the syntax, DR necessarily applies, and a WCO violation arises.

1.2.3. *'At least' and 'exactly' readings*

The present account sheds new light on the discussion concerning the semantics of cardinal determiners. As is well-known, they are assumed to mean *at least n* (see in particular Keenan and Stavi (1986)); the *exactly* reading that they have in many contexts can be pragmatically derived as a consequence of Gricean principles. If the present account is correct, the relation between the *at least* and the *exactly* readings should be revised: the *exactly* reading can be forced not only by virtue of the pragmatic context, but also by the syntactic context. More precisely, the present account suggests that the *at least* reading is not a *lexical* property of the cardinal determiners; what appears to be the case is that the *at least* reading depends on the *existential* reading of cardinals; given the analysis proposed here, the *at least* reading is blocked if the determiner is blocked in its base position: since closed DPs

denote individuals, the determiners that head them indicate the cardinality of the plural individual constant (see (9)), which is necessarily of the *exactly n* type.

1.2.4. *The syntactic levels of semantic representation*

As their label indicates, Heim's Interpretive Conventions *are not LF rules*: they are neither movement nor binding rules, and as such are not subject to syntactic constraints. Note furthermore, that these conventions do not map S-structure representations onto LF representations; rather, they take LF representations as an input and map them onto another level of representation, call it LF'. [8] Although LF' must be distinguished from LF, it is still a *syntactic* level of the Interpretive Module, which feeds the *semantic rules* (functional application) of the Interpretive Module. In order to keep LF representations distinct from LF' representations, I will use the following notation: given an example (n), its LF representation will be notated (n^1) or (n^2), depending on whether DR has applies or not; the corresponding LF' representations will be notated (n^1)' or (n^2)'.

Note that although Interpretive Conventions are not syntactic rules, the syntax does constrain them indirectly, because the input of the Interpretive Conventions is syntactically defined (the existence of a certain type of LF, one which contains an open DP, depends on the syntactic context: the emergence of open DPs is due to DR, which is an LF rule, constrained by the general rules of syntax).

1.2.5. *Bare quantifiers*

The analysis of prepositional Accusatives proposed in 1.2.1 above should be completed by taking into account bare quantifiers such as *cineva* 'somebody' or *nimeni 'nobody'*, which are necessarily preceded by the preposition *pe*:

(16) Ion a examinat **pe** cineva.
John examined 'pe' somebody.

8. My LF' should be kept distinct from May's (1985) LF', which is assumed to contain representations obtained by means of pragmatic rules. If pragmantic and semantic principles are to be distinguished, my LF' is the result of semantic rules alone (the Interpretive Module that I refer to pertains to semantics rather than to pragmatics).

Given our account of the example in (12), an LF such as (16[1]) is illegitimate, on a par with (12[1]); when *pe* is present, DP constitutes a barrier for the raising of any of its subparts:

(16[1]) ◊ Ion a examinat [$_{DP}$ **cine$_i$** [$_{PP}$ pe [$_{DP}$ **e$_i$** va]]
John examined ***some$_i$*** *[pe* ***e$_i$*** *va]*

We are then bound to assume that DR does not apply to subconstituents of bare quantifiers; [9] in other words, bare quantifiers do not "open up" at LF. This seems to force upon us the conclusion that bare quantifiers are interpreted referentially, but this is incorrect. Romanian bare quantifiers have all the properties characteristic of nonreferential DPs (see the tests in 1.1. above), in particular they give rise to Weak Cross Over violations, and cannot be clitic-doubled:

(17) a. *Mama lui$_i$ a ajutat pe cineva$_i$.
His$_i$ mother helped somebody$_i$.
b. *Ion l$_i$-a examinat [$_{PP}$ pe [$_{DP}$ cineva]$_i$] (= (16) + clitic-doubling)

In order to describe correctly the behavior of bare quantifiers, we need a specific interpretive procedure. The conventions in (14) should be modified by adding the clause set in bold characters:

(14) a. Insert variables in the position of open DPs **and of bare existential quantifiers** [10] and
b. i. Apply existential closure (whenever the open DP/bare quantifier is not inside the restrictive term of another quantifier).
ii. Interpret cardinal determiners as cardinality markers of the variable.
iii. Interpret NP as the range of the variable.

---

9. I do not think that the morphological unanalyzability of *nimeni* 'nobody' and *nimic* 'nothing' (*ni* may be analyzed as a negative prefix, but *-meni* is neither a free nor a bound morpheme; *mic* may be related to the adjective *mic* 'small') can be invoked as an explanation for the fact that DR does not apply. Thus, *nobody, somebody, nothing*, etc. function as bare quantifiers (in the sense that no subconstituent should be assumed to raise at LF) even if they can be morphologically decomposed.

10. Bare quantifiers should be distinguished from indefinite pronouns such as *unul* 'litt. one-the' meaning "one" or "one of them" or *niciunul* 'litt. neither-one-the' meaning 'none' or 'none of them'. These pronouns can optionally take the prepositional Accusative, in which case they are interpreted referentially. When unmarked, they get the existential reading (see Dobrovie-Sorin (1993)).

1.3. *The amount reading*

1.3.1. *Long DR at S-structure*

Let us next consider the French examples in (18):

(18) a. J'ai beaucoup$_i$ lu [e$_i$ de livres].
I have much/many$_i$ read e$_i$ of books.
b. Combien$_i$ as-tu lu [e$_i$ de livres]?
how much/many$_i$ have you read e$_i$ of books

In (18)a the determiner *beaucoup* has raised [11] to some adjunction position to VP. [12] This position may serve as an intermediate step for the movement of *combien* in (18)b, which raises higher, presumably to an adjunction position to CP.

1.3.2. *Long DR at LF*

As proposed in Dobrovie-Sorin (1988, 1992, 1993), it is reasonable to assume that an LF counterpart of the French rule exists. [13] This rule is like the DR rule that we have assumed for deriving the existential reading insofar as in both cases the determiner raises out of its base position. The two rules differ however in their landing position (DP adjunction versus VP adjunction). To keep them distinct I will use the labels "short" DR and "long" DR, respectively. I will use distinct superscripts for the various distinct LF representations that characterize indefinites: given an example (n), the representations (n^1), (n^2) and (n^3) will notate the LFs obtained by short DR, no DR and long DR, respectively.

By applying the rule of long DR, the example in (19) would be represented as in (19^3):

---

11. Note, however, that according to Obenauer (1984-1985, 1992), *beaucoup* is base-generated in its S-structure position, attached to V, rather than moved from within the DP.

12. For our present purposes the exact position of the raised determiner is not crucial: instead of an adjunction slot we might assume a Spec,VP position as proposed by Rizzi (1990).

13. As different from Dobrovie-Sorin (1992, 1993), it is assumed here that the raised determiner does not go up to IP, but stays lower, at the VP-level.

(19) John examined ten students.

(19^3) [John [$_{VP}$**[ten$_i$]** [$_{VP}$ examined [$_{DP}$ e_i students]] (long DR)

Let us now consider the interpretation that could be assigned to (19^3), by comparing it to the existential reading. It is quite obvious that like the existential reading, the one corresponding to (19^3) is a nonreferential interpretation: no closed DP is present, therefore the constant type of denotation is barred (hence WCO violations, possibility of narrow scope reading, impossibility of clitic-doubling, etc.). However, it will be shown that (19^3) does not give rise to the existential reading either, but to a different reading, the "amount/detranzitivized" reading (see de Hoop (1992), Szabolcsi and Zwarts (1993), Frampton (1990) and Dobrovie-Sorin (1993)). Our task is twofold: (a) formulate the Interpreting Conventions that relate (19^3) to the amount/detranzitivized reading; (b) formulate tests that can distinguish between the amount and the existential readings.

The Interpreting Conventions to be proposed should rely on the differences between the LFs obtained through long and short DR, see (19^3) and (11^1), respectively. I repeat here (11^1):

(11^1) [$_S$ [$_S$ I met [$_{DP}$ *[a_i]* [$_{DP}$ e_i [$_{NP}$ friend]] (short DR)

Assuming that the determiner itself, rather than its trace, is relevant for the rules of semantic composition (i.e., functional application), the DP projection in (19^3) is not visible. Compare the representation in (11^1), where the determiner is low enough for it to apply to the NP-denotation. This observation can be summarized by saying that (11^1) contains an open DP, whereas (19^3) has only an NP; in this case the DP projection is invisible.

Since the DP projection is invisible in (19^3), the convention in (14) cannot apply, which means that an individual variable cannot be inserted in the object position. The differentiated treatment of open DPs and NPs is conceptually motivated by the already discussed function of Determiners: they are needed in order for the NP to function as an argument. Now, individual variables are arguments, so that it is reasonable to assume that they can only replace DPs, but not NPs; "bare" NPs, i.e. NPs that lack an overt determiner, can only function as predicates. [14]

14. The proposed analysis relies crucially on the idea that the traces of raised determiners are

Coming back to (19^3), the determiner is too high to apply to the predicate expressed by the NP: instead, the determiner applies to the predicate of the clause (type <e,t>), and yields another predicate. The long raised determiner is then an entity of type <<e,t>, <e,t>>, which is the characteristic type of adverbs. Most interestingly, the adverbial function of a long-raised determiner, obtained here in a deductive way, corresponds to Obenauer's (1984-1985) empirical results: by observing the empirical properties of examples such as (18), he has proposed that the determiner functions as a quantifying adverb, rather than as a determiner. Coming now to the denotation of the NP, it can only be of a predicate-type. However, a predicate cannot be interpreted in an argument position. The only possibility then is complex-predicate formation: the predicate expressed by *friend* combines with the predicate expressed by the verb. We are thus able to derive in an explanatory way a "detransitivized" type of reading, which is sometimes described in the recent literature (see in particular de Hoop (1992)). Let us then summarize this discussion by assuming the following Interpreting Conventions, which, given the preceding discussion, are not stipulative, but derive from the general rules of semantic composition:

(20) a. A long-raised Determiner functions as an adverb.
b. An NP functions as a predicate, which may combine with the clausal predicate to give rise to a complex predicate.

It thus appears that the LFs obtained by short and long DR receive distinct interpretations. Paraphrases such as those listed below can be suggested for the amount/detranzitivized reading, corresponding to long DR:

Readings corresponding to (19^3):

'John student-examined ten times.'
'Ten is the number/quantity in which John student-examined.'

To summarize, according to the analysis proposed here, determiners may take three different types of denotations, depending on the LF position they occupy at

---

invisible at LF. Other empty determiners may however be considered to be zero morphemes, and as such function on a par with overt determiners (see plural DPs in Spanish or Romanian, generic DPs in English, etc.).

LF: if they stay in their base-position they denote choice functions (i.e., are of type <<e,t>, e>); if they adjoin to the DP they denote existential functions; if they adjoin to VP they have an adverbial-type of denotation, <<e,t>, <e,t>. Under this view, the type-shifting rules (see Partee (1987)) that affect DPs depend on the position the determiners occupy at LF.

1.3.3. *Weak Islands*

Paraphrases do not constitute compelling evidence in favor of the hypothesis that two distinct readings rely on distinct LF representations. One might argue that the existential reading entails the amount reading, or the reverse. If this were the case, one LF representation would be sufficient: it would give us one of the two readings, the other being derived by entailment. In order to argue in favor of the existence of two distinct LFs, one must find syntactic contexts in which one of them is blocked by the general rules of syntax. The prediction made by the present framework is that in those contexts the corresponding reading is blocked. The case in point is provided by weak islands. Szabolcsi and Zwarts (1990, 1993), Frampton (1990) and Dobrovie-Sorin (1988, 1992, 1993) have argued that the relation between an individual variable and a quantifier is not sensitive to weak islands; only non-individual variables, e.g., the traces of long-raised determiners, are sensitive to weak islands. This means that the existential reading is not sensitive to weak islands, whereas the amount reading is.

Weak islands are usually discussed for wh-structures. However, the effect of weak islands can also be observed for quantified DPs in situ, as shown in examples (21). Since the discussion deals with the possible readings of (21)a-b, which do not show crosslinguistic variation, I use English:

(21) a. John read many books.
b. John did not read many books.

Example (21)b is compatible with two distinct types of readings, those in (22) and (23):

(22) You did not consult some books. Their number is many.
There are some books that you did not consult. Their number is many.

(23) You consulted books. Their number is not many.
You not-many consulted books.
The number of books that you consulted is not many.

In (23) the negated element is not the verb *consult*, but just the numeral. It is natural to assume that this type of reading relies on an application of DR which raises the numeral out of the DP and adjoins it to a position which is within the scope of the negation; DR could adjoin the numeral to VP:

(21[3]) b. You did not [$_{VP}$ [many$_i$ [$_{VP}$ consult [e$_i$ books]]

The idea that the reading in (23) relies on the LF in (21[3]) seems reasonable: if long DR did not apply, the numeral would be too low in the structure for it to be affected by the negation operator. Note that (21[3])b is legitimate with respect to antecedent-government because *not* does not intervene between the raised determiner and its trace.

As to the interpretation paraphrased in (22), it relies on an LF of the type in (21[1]), obtained by short DR; in this case, the determiner has not raised high enough for it to be affected by the negation. By the Interpretive Conventions in (14), (21[1]) is mapped onto the LF' representation in (21[1])', which underlies the existential reading:

(21[1]) b. [You did not consult [$_{DP}$ many$_i$ [$_{DP}$ e$_i$ books]]

(21[1])' b. $\exists x^{many}$ [You did not consult x^{many} ^ book (x^{many})]

The LF in (21[1]) is legitimate, because the trace of *many* is antecedent-governed. It is not clear that LF' representations are subject to the ECP. But in case they were, the legitimate character of (21[1])'b is straightforward: the variable in (21[1])'b is an individual variable, i.e., a variable that has a "referential" th-role (in Rizzi's (1990) sense), and therefore it need not be antecedent-governed (it is not sensitive to weak islands).

Coming back to the LF relying on long DR, one may wonder whether the raised determiner may be raised above the Negative operator:

(21[4]) b. ◊You did many$_i$ [$_{NegP}$ not [$_{VP}$ [$_{VP}$ consult [e$_i$ books]]

Assuming a minimality-based account of weak islands (see Rizzi (1990)) this LF is illegitimate, as indicated by the diamond: the determiner is not a th-marked element, and therefore its trace should be antecedent-governed, but antecedent-government is blocked by the intervening negation. [15] Compare the LFs in (21[1])b and (21[3])b: in both cases DR raises the determiner lower than the negation, and therefore antecedent-government holds (the ECP is not violated).

Given the illegitimate character of (21[4])b, we expect the corresponding reading not to be available. Such a reading is indeed not only unavailable, but in fact incoherent. The two parts of (24) present different presuppositions and therefore cannot make up a coherent discourse:

(24) You did not book-consult. Their number is many.

To sum up, weak islands provide us with a means of distinguishing between the amount and the existential readings. For QPs in situ, weak islands do not block the amount reading: rather, they take scope over the long-raised determiner (see the LF in (21[3])b). The existential quantifier, on the other hand, necessarily escapes the scope of the negation.

Note next that an interrogative counterpart of (21[3])b is not available. And more generally, the trigger of weak islands cannot take scope over a long-raised wh-determiner:

(25) a. How many books haven't you consulted?
b. How many books do you regret to have consulted?

(25)a-b cannot be interpreted as asking: "*not how many* is the number of books that you read" or "what is the number such that you regret that you read that number of books?" The absence of this kind of reading is due to the presence of an interrogative: the corresponding QPs in situ allow the long-raised determiner to stay in the scope of the Negation or of *regret* (see again (21[3])b and the corresponding LF' and interpretation). The generalization is straightforward:

---

15. Rizzi's own analysis would be somewhat different, because he does not allow adjunction; he assumes instead that raised determiners land in some Spec position. (21[4])b would be illegitimate because the Spec position of the Negation is already occupied.

(26) Wh-constituents must take widest scope.

By virtue of (26), *how many* cannot be reconstructed into the position corresponding to *many* in (21^3):

(25^3) a. ◊ you haven't [how many$_i$ [$_{VP}$ consulted [e$_i$ books]]
b. ◊ you regret [how many$_i$ [$_{VP}$ to have consulted [e$_i$ books]]

LFs such as (25^4) on the other hand obey (26), but they violate the ECP (give rise to a weak island violation):

(25^4) a. ◊ [how many$_i$ [you haven't [$_{VP}$ consulted [e$_i$ books]]
b. ◊ [how many$_i$ [you regret [$_{VP}$ to have consulted [e$_i$ books]]

Since both of the LFs in (25^3) and (25^4) are illegitimate, we obtain the generalization stated in (27):

(27) The presence of a weak-island trigger in wh-interrogatives blocks the amount reading of the wh-element.

Note on the other hand that weak islands do not block the existential reading of wh-phrases: an interrogative counterpart of the LF' in (21^1)' is legitimate, because the interrogative operator corresponding to the existential operator binds an individual variable:

(25^1)' a. for x$^{\text{how many}}$ you haven't [$_{VP}$ consulted x$^{\text{how many}}$]

Summarizing again, weak islands block the wide-scope amount reading of QPs, but do not block the existential reading. Note that the existential quantifier necessarily escapes the scope of the negation; the amount reading is the only "narrow-scoped" reading with respect to negation. To put it in other words, the amount reading has the narrowest possible scope: the determiner itself has narrow scope with respect to the negation operator, verbs like *regret*, etc. The QP itself has no scope at all, because under the amount reading, QPs do not function as quantifiers. Recall that in fact, "amount" QPs are not semantic constituents at all:

they have splitted up into a Det which functions as an adverb, and a NP, which functions as a predicate.

An even clearer argument in favor of the distinction between the existential and the amount readings can now be given. It is well-known that the intensional reading of NPs cannot be represented by existential quantification. Assuming that the intensional reading is nevertheless represented at LF, the only possibility is the LF underlying the amount reading, obtained through long DR. Given the generalization in (27), we make an interesting prediction, which appears to be correct: the intensional reading is blocked by the presence of a weak-island trigger in wh-interrogatives:

(28) a. Combien de secrétaires cherches-tu?
b. Combien de secrétaires ne cherches-tu pas?
c. Combien de secrétaires auras-tu beaucoup cherché?
d. Combien de secrétaires regrettes-tu d'avoir cherché?
e. ??How many stories$_i$ won't Diana invent e$_i$?
f. ??How many gaffes$_i$ do you regret that John made e$_i$?
g. ??How many lies$_i$ do you regret that he came up with e$_i$?
(e-f are adapted from Herburger (1992))

As predicted, the intensional reading is possible in (28)a, but not in (28)b-d, where a weak island trigger is present. The ungrammaticality notated by two question marks is due to the fact that the examples in (e)-(f) are constructed with verbs which are not compatible with the existential reading: a story does not exist before it is invented, a gaffe does not exist independently of its being made.

1.4. *The presuppositional-quantificational reading*

It is well-known that in subject position, indefinites are compatible with both the existential and the presuppositional readings, which are currently assumed to rely on the representations in (29)' and (29)", respectively (Diesing (1992)):

(29) Ten students will paint the walls.

(29)' Future $\exists x^{10}$ [x^{10} paint the walls & x^{10} is student]

(29)" ten x (x is student) [x will paint the walls]

(29)' corresponds to a reading that asserts the existence of ten students who will paint the walls: "there will be ten students who will paint the walls". Compare (29)", which presupposes the existence of ten students and asserts that they will paint the walls. One or the other of these two readings may be preferred, given a particular context. Thus, the existential reading seems to be imposed if the speaker of (29) goes on by saying *Then we'll be ready for cleaning the windows*. Alternatively, if the speaker adds *You decide who*, the presuppositional reading is more likely to have been intended. [16]

The denotation of existential indefinite DPs has already been discussed above: the NP constituent denotes a predicate that define the range of the variable by which the open DP has been replaced: the determiner indicates the cardinality of the (plural) variable. Compare (29)", where the indefinite functions as a generalized quantifier, i.e. as an element of type <<e, t>, t>: it applies to a predicate (λx: x is intelligent) and yields a truth value.The reading corresponding to (29)" may then be referred to as the "presuppositional-quantificational" interpretation of indefinites. I will use this composite term in order to avoid terminological confusion. "Presuppositional" is sometimes taken to mean "a DP that presupposes that the set denotated by its NP is not empty" (de Jong (1987)). By this definition, partitives (i.e., DPs of the type *Det of DP*) are necessarily "presuppositional". However, partitives are not necessarily interpreted as generalized quantifiers: it is not clear that in examples such as *John met two of your students*, the object is assigned an <<e,t>, t> type of denotation; an existential type of reading seems possible for the object: "there are two individuals who have been examined by John and are your students". The term "quantificational" is itself ambiguous: thus, in Fodor and Sag's terminology, quantificational indefinites are those indefinites that may give rise to a nonreferential type of reading; according to this definition, existential indefinites count as quantificational. The reader should then recall that in what follows

16. The disambiguation may also be due to lexical properties. Thus, it is well-known that the existential reading is incompatible with individual-level predicates (Carlson (1989), Kratzer (1989), Diesing (1992)); the example in (i) is therefore necessarily interpreted as presuppositional:

(i) Two students are intelligent.
(i)" two x (x is student) [x is intelligent]

"presuppositional-quantificational" indefinites are distinct from both presuppositional-partitive indefinites and quantificational indefinites in the sense of Fodor and Sag. Presuppositional-quantificational indefinites are those indefinites that are interpreted as generalized quantifiers, i.e., as constituents of type <<e, t>,t>, which apply to a clausal predicate.

It is interesting to observe that bare quantifiers also allow the presuppositional-quantificational reading. Thus, (30) has two possible readings, much like (29). Depending on whether the speaker goes on by saying *Then we'll be ready for cleaning the windows,* or *You decide who*, (30) should be represented as (30)' or as (30)":

(30) Somebody will paint the walls.

(30)' Future ∃x [x paint the walls]

(30)" an x [x will paint the walls]

The existential reading, corresponding to (30)', could be paraphrased by "there will be somebody who will paint the walls". Under the presuppositional-quantificational reading (corresponding to (30)"), the quantifier is necessarily restricted, and since bare quantifiers do not have a restrictive term, [17] the context provides one: *somebody* will be interpreted as *somebody among you*, or *somebody* among a group of people that had been already mentioned in the discourse. A possible paraphrase of this reading would be "somebody among you will have to paint the walls".

The case of bare quantifiers is crucial for an analysis of the presuppositional-quantificational reading. Thus, it has been suggested that the presuppositional-quantificational reading of indefinites is due to the *intrinsic* possibility that they have to take an <<e,t>, t> type of denotation (Diesing (1993), Delfitto (1993)). However, it is clear that *somebody* does not have this *intrinsic* possibility. I would like to argue instead that the <<e,t>,t> type of denotation is *contextually* determined for indefinites and bare quantifiers alike. My analysis will rely on the following Interpretative Convention, which will be motivated in what follows:

(31) Translate indefinite DPs and bare quantifiers as generalized

17. A discussion of why "person" cannot be taken to be the restrictive term of bare quantifiers can be found in Dobrovie-Sorin (1993).

quantifiers of the form *for Det x*, only if at S-structure they constitute the subject of a predicate.

This convention resembles Diesing's Mapping Hypothesis only partially. Like Diesing, I believe that the presuppositional-quantificational reading of indefinites cannot appear if the indefinite occupies a VP-internal position. Unlike Diesing however: (a) I assume that the indefinites that occupy the Spec,IP position may be assigned the existential reading; (b) I do not attribute any special property to the Spec,IP position itself: any position which is the sister of a clausal predicate will make the presuppositional-quantificational reading possible; (c) the relevant level of representation is S-structure: the subject-predicate configuration required by (31) must be present at S-structure; it cannot be created via LF movement.

The main idea behind (31) is that indefinites and bare quantifiers cannot be assigned a <<e,t>, t> type of denotation *intrinsically*, i.e., by virtue of their lexico-semantic features, but only *contextually*, by virtue of the fact that they apply to a predicate, i.e., to a constitutent of type <e,t>. A natural motiviation for this view may be given, by deriving it as a consequence of the fact that indefinite determiners denote symmetric functions. It is because *three* is symmetric that (32)a is truth-conditionally equivalent to (32)b:

(32) a. I met three students.
b. Three of the people I met are students.

Compare DPs that have *nonsymmetric* determiners, e.g., *most students* or *every student*: their restriction and their scope are fixed independently of the syntactic position they occupy, by virtue of their semantic properties (NP is the restriction and S the scope). Thus, the only possible representations of (33)a-b rely on generalized quantification, as shown in (33)"a-b (The LF' representations marked with " will consistently correspond to generalized quantification; those marked with ' underly the existential reading):

(33) a. John met most students.
b. John met every student.

(33)" a. most x (student (x)) [John met x]
b. all x (student (x)) [John met x]

The representations in (33)" can be obtained without using the rule of QR, by assuming instead the following interpretive conventions:

(34) a. Insert variables in the position of open DPs.
b. Interpret open DPs with nonsymmetric determiners [18] as generalized quantifiers.
c. Insert quantifiers in the scope position. [19]

Coming back to indefinite DPs, their determiner being *symmetric*, they cannot be assigned scope by virtue of their semantic properties. This means that it is not their semantic properties that allow them to function as generalized quantifiers. In order for indefinite DPs to function as quantifiers, the syntactic configuration itself must provide the basis for *nonsymmetry*. In other words, open cardinal DPs may function as quantifiers only if their scope is marked at S-structure. Since syntactic subjects take scope over their predicate, subject-predicate configurations are one kind of configuration in which indefinite DPs may function as generalized quantifiers. And this is precisely what the convention in (31) says.

There is however a conceptual problem related to the way in which (31) is formulated: it has the status of an interpretive convention, and as such it should apply on LF configurations; but instead it is defined on S-structure. It is of course easy to remedy to this shortcoming, by simply stipulating that it is LF rather than S-structure which is the relevant level at which (31) applies:

---

18. The distinction between open and closed DPs is relevant for strong DPs also:

(i) Am ascultat toţi elevii.
[I] have examined all the students.
(ii) I-am ascultat pe toţi elevii.
[I] them-have examined 'pe' all the students.

Due to the presence of the prepositional Accusative, the example in (ii) takes a referential reading (in the sense of section 1.2.1 above): *toţi elevii* denotes a plural individual, which is the sum of all the students relevant in a given context. This individual has the cumulative-distributive reference property characteristic of individuals (Link (1983), Landman (1989)). Note that *every*, as different from *all*, is necessarily of type <<e,t>,t>, which means that it cannot denote an individual.

19. The procedure stated in (34)a-c is much like Williams's (1986) scope indexing procedure, by which QPs in situ are interpreted as variables (compare (34)a) bound by a scope index on IP. The insertion procedure postulated in (34)b is just a notational variant of Scope Indexing, by which we can derive a more intuitive representation, one which is closer to the interpretation, and thus is easier to evaluate by the reader. Both of these procedures rely on the idea that scope assignment does not rely on movement, and that the unmarked scope position is S.

(31') Translate indefinite DPs and bare quantifiers as generalized quantifiers of the form *for Det x*, only if at LF they constitute the subject of a predicate.

As I will show, this restatement is possible only if QR is dispensed with. Consider (35): if QR were allowed, a presuppositional-quantificational reading could be assigned to *ten students*, by applying QR followed by (31'):

(35) I met ten students.

(35[1]) [ten students]$_i$ I met e_i.

(35[1])" for ten x (x is students) [I met x]

In order to rule out the presuppositional-quantificational reading for cardinals in object position, we must rule out the QR of indefinites (contra Diesing). [20] Since on the other hand QR is not needed for strong DPs (see the Interpretive Conventions in (34)), we can dispense with QR altogether.

To summarize, the conventions in (34)a and (34)c apply to indefinites, bare quantifiers and strong DPs alike. But while (34)b holds for strong DPs, we assume (31') for indefinites and bare quantifiers. The interpretive conventions in (14) can then be revised by adding the alternative stated in (14)c:

(14) a. Insert variables in the position of open DPs and of bare existential quantifiers and

b. i. Apply existential closure (whenever the open DP/bare quantifier is not inside the restrictive term of another quantifier).

ii. Interpret cardinal determiners as cardinality markers of the variable.

iii. Interpret NP as the range of the variable.

or

c. i. Translate indefinite open DPs and bare quantifiers as generalized

20. The reader who has already noticed that there *are* presuppositional-quantificational readings for object indefinites is kindly required to read section 1.5. below. For now, let us accept the strong (and in fact, partially incorrect) generalization that object indefinites cannot take the presuppositional-quantificational reading.

quantifiers of the form *for Det x* only if at S-structure they constitute the subject of a predicate.

ii. Insert generalized quantifiers in the scope position.

By these conventions, the example in (29) can be assigned the LF' representation in (29[1])"; (29[1]) is the LF representation obtained by short DR; (29[1])" is then obtained by applying (14)c:

(29) Ten students will paint the walls.

(29[1]) [ten$_i$ [e$_i$ students]] will paint the walls

(29[1])" ten x (student (x)) [x will paint the walls]

1.5. *A default generalized quantifier*

The perspicuous reader will have noticed that both indefinites and bare quantifiers in object position do allow a presuppositional-quantificational reading. Consider indeed the example in (36):

(36) I am sure everybody would help three homeless poor guys.

The unmarked reading is of the presuppositional-quantificational type: "I am sure that given three homeless poor guys everybody would help them". Compare the examples in (37), which favor the existential reading, "I hear there is an incredibly intelligent student whom everyone admires":

(37) a. I hear /know everyone admires an incredibly intelligent student.
b. I hear /know everyone admires somebody.

(37') a. I hear ∃x [x is incredibly intelligent and everybody admires x]

We thus seem to be confronted with a paradox. Given the argument developed in the preceding section, indefinite objects cannot be assigned a presuppositional-

quantificational reading, because (a) their determiner is symmetric and (b) the syntactic configuration does not provide us with the nonsymmetry we need. However, the indefinite object in (36) does have a presuppositional-quantificational reading.

To solve the paradox, one might be tempted to allow QR (Diesing (1992)). This is however not the only way out. I will assume instead the default procedure stated in (14)d, to be added to the conventions in (14)a-c, repeated here:

(14) a. Insert variables in the position of open DPs and of bare existential quantifiers and

b. Apply existential closure (whenever the open DP/bare quantifier is not inside the restrictive term of another quantifier).

or

c. Translate indefinite open DPs and bare quantifiers as generalized quantifiers of the form *for Det x* only if at S-structure they constitute the subject of a predicate.

or

d. Insert a default quantifier *for an* x^n.

Note that the default quantifier *for an* x^n crucially differs from the quantifier *for n x* obtained via the translation procedure in (14)c: (a) *for an* x^n binds a plural individual variable, whereas *for n x* binds an atomic individual variable; (b) correlatively, *for an* x^n is a non-distributive operator, whereas *for n x* is distributive. The distributivity feature that distinguishes between the two types of presuppositional-quantificational readings of indefinites will appear to be crucial for the analysis of scope properties (see section 2 below).

The difference between the two readings follows naturally from what has been said above concerning the semantic properties of indefinite determiners. More precisely, according to their semantic features, indefinite determiners are symmetric functions; they can function as nonsymmetric, i.e., they can be translated by *for Det x*, only if the syntactic configuration provides us with the necessary nonsymmetry. This line of reasoning is perfectly compatible with the insertion of a default operator *for an* x^{Det}, because in this case the determiner itself stays symmetric.

By applying (14)d we can derive the representation in (36)''' for the example in (36):

(36)''' I am sure for an x^3 (x^3 is a homeless poor guy) [everyone would help x^3]

1.6. *Economy of Derivation and Disambiguation*

According to the analysis proposed here, the indefinites that occupy the (Spec,IP) position can be interpreted as either presuppositional-quantificational or existential (no reconstruction, i.e. no lowering to a VP-internal position need be assumed, contra Diesing (1992)). The two readings, which rely on one and the same LF representation, are derived via distinct, but equally productive Interpretive Conventions. One or the other of the two interpretations will be selected given a particular context. Compare the difference between the referential and the existential readings in direct object positions. Since these two interpretations rely on distinct LFs, the larger context will not be able to impose a certain reading. Rather, an unadequate context will give rise to uninterpretability. Recall also that the LF corresponding to the referential reading, which contains closed DPs, is not a free option: it only arises if the movement of DR is blocked by the syntax. That is why WCO violations cannot be avoided by freely resorting to the referential reading.

Our view, according to which the existential and the presuppositional-quantificational readings are both possible in the (Spec,IP) position, seems to be unadequate for certain examples, in which the (Spec,IP) position forces the presuppositional-quantificational reading. The relevant examples are characterized in a quite straightforward way, although Diesing's (1992) own presentation obscures what I take to be their common feature:

(38) A DP whose S-structure position is higher than the position in which its morphological features are checked is interpreted as presuppositional-quantificational.

Cases in point can be found in Germanic SOV languages, in which the morphological checking of subject DPs can be done inside VP, presumably in the (Spec,VP) position (the overt manifestation of this possibility is the fact that subject DPs appear inside the VP constituent at S-structure). Since the subject is legitimate in a VP-internal position, the movement to (Spec,IP) is not forced by morphological

checking. Compare English or French, where the subject moves to (Spec,IP) for Case-checking. Correlated to this difference concerning Case-checking, the (Spec,IP) forces the presuppositional-quantificational reading in German, but not in French or English. [21] Similarly, scrambling is clearly not forced by Case-checking and correlatively scrambled DPs are interpreted as presuppositional-quantificational:

(39) dass Otto immer Bücher über Wombaths liest.

(40) dass Otto Bücher über Wombaths immer liest.

In (39) the direct object occupies a VP-internal position, whereas in (40) it has been scrambled, i.e., adjoined to IP. The different S-structure positions of the object in (39)-(40) correlate with different readings: existential and presuppositional, respectively. Within Diesing's tree-splitting Hypothesis these examples seem to be adequately accounted for, since (Spec,IP) and the scrambling positions are the first members of split configurations, and as such are necessarily associated with the presuppositional-quantificational reading. Note, however, that for English, in which (Spec,IP) is a Case-checking position, Diesing has to resort to lowering in order to account for the fact that the existential reading is allowed. The question then arises as to what blocks lowering in the examples where (Spec,IP) is higher than the Case-checking position.

Let us now come back to the analysis proposed here, where the (Spec,IP) position allows not only the presuppositional-quantificational, but also the existential reading. The fact that the presuppositional-quantificational reading is forced in the cases described in (38) and illustrated in (39)-(40) follows from Economy principles:

---

21. Delfitto (1993) observes that in Italian and French the subject of unaccusatives is necessarily interpreted as presuppositional when it occupies the (Spec,IP) position:

(i) Due linguisti sono arrivati.
(ii) Deux linguistes sont arrivés.
Two linguists arrived

Note that examples of this type are also characterized by the property described in (38): in Italian and French, unaccusatives are the only verbs that allow their subject DPs to be Case-checked in a VP-internal position:

(iii) Sono arrivati due linguisti.
(iv) Il est arrivé deux linguistes.

(41) a. Do not move unless necessary (Economy)
b. Move α is necessary for Case-checking.
c. Move α is necessary in order to allow an interpretation that is otherwise not obtained.

These principles rule out the movement of a DP to a position that is higher than its Case-checking position in case the intended reading is obtained in the Case-checking position itself. Given our account of the existential reading, it can be obtained in any syntactically legitimate position, i.e., in any Case-checking position. By (41)c, then, no movement to a higher position will be permitted if the intended reading is the existential one. The presuppositional-quantificational reading, on the other hand, can only appear if the indefinite has a C-command domain at S-structure. Since this condition is not satisfied in certain Case-checking positions, movement to a higher position will be allowed, in fact forced, if the presuppositional-quantificational reading is intended. If (Spec,IP) is a Case-checking position (as in English or French), both the presuppositional-quantificational and the existential readings are allowed because there is no lower Case-checking position, where the existential reading could appear. Thus, under the analysis proposed here, reconstruction of the subject DP into the (Spec,VP) position is not necessary in order to obtain the existential reading.

Another difference between Diesing's analysis and mine concerns the presuppositional-quantificational reading of nonscrambled objects (or of German subjects that occupy the (Spec,VP) position at S-structure). For Diesing this reading is of exactly the same type as the presuppositional-quantificational reading that arises for DPs that occupy the (Spec,IP) position, and it is obtained via the application of QR at LF. [22] Within my own analysis, QR has been dispensed with. The presuppositional-quantificational reading of direct objects is of a special type: it relies on a nondistributive operator for x^n (compare *for n x*), which is introduced via a default procedure.

22. Diesing (1992) gives no empirical evidence in favor of QR, nor in favor of reconstruction. Both QR and reconstruction *are forced* by her Mapping hypothesis, but they do not *support* it. In other words, Diesing's hypothesis is stronger than what the data allow us to assume.

1.7. *Focus-affected readings*

Consider next the following examples; (42) is borrowed from Westerstahl (1985) and (45)-(46) from Herburger (1993):

(42) a. Many Scandinavians have won the Nobel prize in literature.
b. "Many of the Nobel-prize winners were Scandinavians".

(43) a. Many Chinese applied last year.
b. "Many of the people that applied last year were Chinese"

(44) a. Many cooks applied.
b. "Many of the people that applied last year were cooks".

(45) a. Few *incompetent* cooks applied. (the italics indicate focus intonation)
b. "Few of the cooks that applied were incompetent."

(46) a. There applied few Spanish citizens *that were native speakers of Basque.*
b. Few of the Spanish citizens that applied were native speakers of Basque.

The common property of these examples is that they show a "proportional" reading of *many/few*, but the proportion is not established with respect to the NP-set; correlatively, the predicate expressed by the NP constituent does not map into the "restriction" of *many/few*, but rather in the nuclear scope. This type of reading may be triggered in different ways. In Westerstahl's example it is the lexical content that is crucial. [23] Thus, the "normal" proportional reading (where the set of reference is the NP-set, i.e., the set of Scandinavians) is not adequate, because given our knowledge about the world, the proportion of Scandinavians that got the Nobel

23. See also the following example from Partee (1988):

(i) There are more illiterate people in small rural towns than in large cities.
"The number of illiterate people living in small rural towns is more important than the number of illiterate people living in large cities"
◊ "More of the illiterate people live in small rural towns than in large cities".

prize does not constitute a relevant piece of information. This is most clearly brought out by the oddity of overtly proportional determiners:

(47) ?? 0,0001% of the Swedish people won the Nobel prize.

Conversely, the proportion of Scandinavians among Nobel-prize winners is relevant (the readers should not, of course, take seriously the percentages in (47) and (48)):

(48) 5% of the Nobel prizes were awarded to Swedes.

Herburger's (1993) examples involve focus-assignment to (part of) the DP constituent. The interesting reading, labelled "focus-affected reading" by Herburger (to be distinguished from contrastive or emphatic focus), is comparable to that described by Westerstahl (1985):

(49) Focus-affected Readings:
Semantic focus inside a DP gives rise to a f-a interpretation, where the focused predicate serves as the main predicate of the sentence and the matrix of the determiner.

It is quite clear that Diesing (1992) does not provide an adequate framework, for the following reasons:

(a) For Diesing the difference between the presuppositional-quantificational and the existential readings relies on the application of a rule that raises the DP into the restriction (i.e., to (Spec,IP)) or leaves it inside the nuclear scope of the determiner (i.e., inside VP). The problem with the f-a reading is that other elements, including the verb itself, go into the restriction. We would then need to stipulate specific raising rules for each type of reading, which would be no more than a paraphrase of the observed readings.

(b) Herburger (1993) gives evidence in favor of the idea that the correct reading cannot even be represented in terms of restricted quantification.

Focus-affected readings thus point to the need of a formalism which does not distinguish so drastically between the existential and the presuppositional-quantificational readings. The analysis proposed here seems to provide a good starting point for an adequate analysis, which is however outside the scope of this paper.

Summary of Section 1

A. Cardinal DPs in the object position may receive one of four readings, corresponding to three distinct LF representations: (i) referential (no DR); (ii) existential or (iii) nondistributive quantificational (short DR); (iv) amount/detranzitivized (long DR). Cardinal DPs in the object position cannot be interpreted as quantificational distributive.

B. Cardinal DPs are not subject to QR.

C. Cardinal DPs in the subject position may receive one of four readings, corresponding to two distinct LF representations: (i) referential (no DR); (ii) existential, (iii) nondistributive quantificational or (iv) distributive quantificational (short DR). Cardinal DPs in the subject position cannot receive the amount reading.

At LF only three distinct types of configurations are assigned. In case DR is blocked, we obtain closed DPs at LF, which underly the referential reading. The amount reading relies on long DR. All the other readings: existential, nondistributive and distributive quantificational correspond to an LF obtained by short DR. The threefold distinction between the readings relying on short DR is not represented at LF, but only at LF', as a result of the application of distinct interpretive conventions: existential closure, insertion of a default quantifier, translation of indefinites as generalized quantifiers.

2. *Scope*

The main aim of this section will be to use the results obtained above, concerning the denotation of indefinites, in order to account for certain scope restrictions that characterize them.

2.1. *Denotation and Scope*

According to Liu (1990), the following generalizations hold:

(50) a. Bare numeral QPs (BNQP: six men, many men, ..) can take scope over a C-commanding DP.
b. Modified numeral QPs (MNQP: fewer than six men, more than six men, exactly six men) cannot take scope over a C-commanding DP.
c. Modified numeral QPs (MNQP: fewer than six men, more than six men, exactly six men) can take scope over a C-commanded DP.
d. Strong QPs (every man, most men, ..) can take scope over a C-commanding DP.

For illustration, consider the examples in (51):

(51) a. Six professors examined two students. (ambiguous scope)
b. Six professors examined fewer than two students. (no O>S reading)
c. Fewer than six professors examined two students. (S>O reading)
d. Six professors examined every/most students. (ambiguous scope)

In (51)a, but not in (51)b, the object DP may take scope over the subject DP. (51)c shows that when placed in a C-commanding position, MNQPs are able to take wide scope over another DP. Their inability to take wide scope, illustrated in (51)b, is thus not a purely semantic property: it is only in certain syntactic positions that MNQPs cannot take wide scope. As observed by Liu, (50)b casts doubt on QR (or on Montague's rule of quantifying in): since in (50)b QR is not blocked in the syntax, we incorrectly predict that the object DP can take scope over the subject. The generalizations in (50)a-d suggest two possible lines of inquiry. One of them would be to let QR depend on the type of quantifier (see Beghelli (1993) who assumes QR for the distributive operator *every*). But such an approach would be highly stipulative: QR would be assumed to apply precisely to those QNPs which take wide scope.

The other possibility is to dispense with QR completely, as we have already done on independent grounds in the preceding sections. I will try to show that the generalizations in (50)a-d follow as a consequence of the types of denotation indefinites may take.

The semantic difference between BNQPs and MNQPs has already been discussed: BNQPS, but not MNQPs, may denote principal filters. It has also been established that principal filters may function as referential, unlike those DPs that

are not principal filters (see sections 1.1. and 1.2.1.). Assuming that the referential versus the nonreferential status of DPs is what accounts for the different scope properties of MNQPs and BNQPs, the generalizations in (50)a-b may be restated as follows:

(50)' Cardinal DPs can take scope over a C-commanding DP only if they are referential.

Our next task, then, is to provide an explanation for this generalization. I will proceed in three steps, by showing that :

(52) (a) existential DPs cannot take scope over a C-commanding DP;
(b) indefinite DPs translated as default quantifiers cannot take scope over a C-commanding DP;
(c) referential cardinal DPs can take scope over a C-commanding DP.

Since the three cases formulated in (52)a-c exhaust the types of denotation that can be assigned to object indefinites (see the main results of section 1), our task of explaining (52) will be completed by accounting for (52)a-c. The generalization in (50)d will also fall out straightforwardly as the demonstration proceeds.

The case of the amount-detranzitivized reading need not be considered, because the LF representation underlying that reading does not present any DP to which scope could be assigned; the determiner itself, which functions as an adverb, has the narrowest possible scope (see 1.3. above).

2.2. *Existential DPs Cannot Take Scope over a C-commanding DP*

Our first task is to show that existential DPs cannot take scope over a C-commanding DP. This result is straightforwardly obtained as a consequence of the generalization stated in (53), to be discussed below:

(53) The existential quantifier cannot take wide scope, but only narrow scope or an independent reading.

Let us assume that the example in (54) has the scope ambiguity characteristic of *Everybody loves somebody*:

(54) Everybody met two students

By the procedures introduced in previous sections, both the universal and the existential quantifiers are inserted in an IP-adjunction position. Now, examples such (54), which present two quantifiers inserted in the IP-adjunction position, raise the question of their order relative to each other. I will assume the null hypothesis, according to which the relative order is irrelevant. Using May's (1985) terminology, quantifiers adjoined to the same maximal projection form a sigma-sequence, which means that they can be interpreted with either relative scope. Note that the existence of a sigma-sequence is very much in the spirit of the present approach, which relies on the insertion of quantifiers in a scope position: since there is only one scope position accessible to several quantifiers, different orderings are possible, assuming the null hypothesis.

Given these remarks, an example such as (54) can be assigned the representations in (54)'a and (54)'b:

(54)' a. ∃y ∀x [x met y] independent reading of ∃y
b. ∀x ∃y [x met y] narrow scope ∃y

The representation in (54)'a is currently referred to as the "wide scope" reading of the existential quantifier. This is an imprecise characterization: in (54)'a the existential is not narrow-scoped; however, it is not wide-scope either, because it does not affect the denotation of the universal quantifier. A more adequate label is "independent" reading.

Note that in (54) we cannot check whether the existential can take scope over the subject, because the denotation of the universal quantifier is such that it cannot be affected by other quantifiers. Let us then examine (55), with an indefinite DP in subject position. As in (54), the existential quantifier related to the DP in object position can be inside or outside the scope of the existential quantifier related to the DP in subject position:

(55) Three professors examined two students

(55)' a. $\exists x^3\ \exists y^2$ [x^3 examined y^2] independent reading
b. $\exists y^2\ \exists x^3$ [x^3 examined y^2] independent reading

The only LF that could correspond to a wide-scope reading of the object is (55)'b. However, (55)'b cannot give us a true wide scope reading, i.e., a reading under which the denotation of the subject DP is affected by the denotation of the object DP. This impossibility is due to the fact that the existential quantifier is not distributive. To put things in more concrete terms, this means that the representation in (55d) is not equivalent to a conjunction of two existential propositions, e.g., not equivalent to *$\exists y\ \exists x^3$ [x^3 examined y] ^ $\exists y\ \exists x^3$ [x^3 examined y].*

Thus, (55)'b does not correspond to a wide-scope reading of the object DP, but rather to an "independent" reading, under which the object DP escapes the scope of the subject DP. Similarly, in (55)'a, the subject DP does not take scope over the object DP; rather, the two DPs are indepedent relative to each other.

To summarize, existentially quantified DPs may take either narrow scope or the independent reading. They cannot take wide scope over another DP. This impossibility is not due to the S-structure position of the indefinite, but is a general characteristic of the existential reading.

2.3. *Default quantifiers Cannot Take Scope over a C-commanding DP*

Consider next the example in (56), interpreted by a default quantifier, as shown in (56)''':

(56) I am sure two nuns would help three homeless poor guys.

(56)''' I am sure for an x^3 (x^3 is a homeless poor guy) $\exists y^2$ (y^2 is a nun) [y^2 would help x^3]

The result we want is that the representation in (56)''' cannot be assigned a O>S reading. This follows from (57):

(57) The default quantifier is not distributive.

Because of (57), the representation in (56)''' cannot be rewritten as a conjunction of the form *for an x ∃y² [y² would help x] ^ for an x ∃y² [y² would help x] ^ for an x ∃y² [y² would help x].*

2.4. *Referential Cardinal DPs Can Take Scope over a C-commanding DP*

Let us now consider that reading of (55) under which the object and subject indefinites are interpreted as referential and existential, respectively. These readings can be assigned by the procedures discussed in section 1. For our present purposes it is important to recall that quantifiers are not *moved* to the IP-adjunction position, but rather *inserted* there. Referential DPs are not moved either (no rule of NPR is assumed); moreover, they cannot be inserted in an IP-initial position, because they are not quantifiers. Given these remarks, the representation in (55²) underlies the reading under which the object is interpreted referentially and the subject is interpreted existentially:

(55) Three professors examined two students.

(55²) $\exists x^3$ [x^3 met [student1+ student2]]

The notation [student1+ student2] is meant to indicate the referential reading of *two students*, i.e., the fact that it denotes a plural individual which is the sum of two atomic individuals who are students. Since *two students* denotes a plural individual, the truth conditions of (55²) are identical to those of (55²)', which can be paraphrased by "the individual [student1+ student2] belongs to the set of individuals characterized by the property "y was met by an x^6 who is professor":

(55²)' [student1+ student2] $\in \lambda y \exists x^3$ [x^3 met y]
[student1+ student2] $\in$ {y : $\exists x^3$ [x^3 met y]}

The representation in (55²)'' can now be derived from (55²)', by applying the Link (1983)-Landman (1989) law of distributivity stated in (58):

(58) $P(j) \wedge P(b) \leftrightarrow P(j+b)$ on the distributive reading of P [24]

(55²)" $student1 \in \lambda y\, \exists x^3\, [x^3 \text{ met } y] \wedge student2 \in \lambda y\, \exists x^3\, [x^3 \text{ met } y]$
$student1 \in \{y : \exists x^3\, [x^3 \text{ met } y]\} \wedge student2 \in \{y : \exists x^3\, [x^3 \text{ met } y]\}$

The representation in (55²)" corresponds to a wide-scope reading: each of the two students is examined by six professors, which need not be the same; so the total number of professors may go up to twelve.

It is important to stress that under the proposed account, the wide scope of object indefinites is not the result of quantifying in (nor of QR), nor of quantifier insertion in a scope position; rather, it is due to the cumulative-distributive reference property of referential DPs. [25, 26]

Our analysis makes the following prediction for Romanian: prepositional Accusatives can be interpreted as taking wide scope, whereas unmarked accusatives are more difficult to interpret as taking scope over their subject. This is indeed a correct prediction:

(59) a. Doi profesori vor examina şase studenţi (??O>S)
b. Doi profesori îi vor examina **pe** şase studenţi (O>S)

To summarize, we have examined in turn the scope properties of all the possible denotations that can be assigned to object cardinal DPs. As the amount/detranzitivized reading relies on an LF representation which does not contain any visible DP, no scope can be assigned to "amount" objects. The existential and the default quantifiers, on the other hand, are nondistributive, and therefore cannot take wide scope. Referential DPs can take wide scope, due to the cumulative-distributive property of individuals. It thus appears that the syntactic position is not directly relevant for an explanation of scope properties: the scope-properties of existentials, default quantifiers and referential DPs have been stated

24. Cumulative reference is the phenomenon that properties of entities are inherited on their sums. Distributive reference is the phenomenon that properties of a sum distribute to its parts.

25. The idea that the wide-scope reading of referential DPs is due to their cumulative reference property can be found in Krifka (1992), whose analysis of wh-QP interactions is however technically different from ours.

26. Ben-Shalom (1992) proposes a different implementation of the hypothesis that the wide scope reading of cardinal DPs depends on referentiality: she states referentiality as a condition on the definition of a binary operators, which is responsible of the wide-scope reading. Under the proposal made here, no binary operator need be defined.

irrespective of the position they occupy. The position does determine scope properties indirectly, by determining the possible denotations of indefinites.

2.5. *Strong quantifiers may take Wide Scope independently of their S-structure position*

Strong quantifiers in object position are necessarily interpreted as quantifiers, and inserted in a scope-position (see (34)a-c). In case an indefinite DP occupies the subject position, scope ambiguity arises, as indicated in (60)'a-b:

(60) Six professors examined every/most students.

(60)' a. $\exists x^6$ every/most y [x^6 examined y] independent reading
b. every/most y $\exists x^6$ [x^6 examined y]

The example in (61), in which the S-structure order of the DPs has been reversed, is ambiguous in exactly the same way as (60). Representations parallel to those in (60)'a-b can indeed be derived, because the strong quantifier in subject position and the cardinal DP in object position can be assigned a quantificational and existential reading, respectively.

(61) Every/most professors examined 6 students.

(61)' a. $\exists x^6$ every/most y [y examined x^6] independent reading
b. every/most y $\exists x^6$ [y examined x^6] wide/narrow scope

2.6. *Generalized quantifiers may take Wide Scope*

In section 1.4. it has been established that in subject position, and only in subject position, cardinal DPs may take a nondefault quantificational reading of the type *for Det x*: The examples in (62) can then be represented as in (62)"a-b:

(62) Fewer than six professors examined two students. (ambiguous)

(62)" a. for fewer than 6x ∃y^2 [x examined y^2] (S>O)
b. ∃y^2 for fewer than 6x [x examined y^2] independent reading

Since the quantifier *fewer than six* is distributive, we directly explain why subject indefinites may take wide scope even if they are not referential (recall that modified cardinal DPs such as *fewer than six professors* cannot denote principal filters, and therefore cannot be referential).

The inverse scope reading of (62) is also marginally possible, because non-modified cardinals in object position can be assigned the referential reading, which brings us to the discussion in section 2.3. above.

It is now possible to give further evidence in favor of the idea already expressed in section 1.4., that presuppositional-partitive DPs are not necessarily assigned a quantificational-partitive reading. If they were, we would expect the object DP of the example in (63) to take scope over the subject, but this is not the case:

(63) Three professors will examine at least 2 of your students.

To summarize, the scope properties of DPs appear to depend on the type of denotation that they take, and more precisely on whether the quantifier they are associated with is distributive or not: existential and default quantifiers are not distributive, and therefore cannot take wide scope. Referential DPs may take wide scope, due to their cumulative-distributive property. Since the type of denotation of cardinal DPs depends on the position they occupy, their scope properties depend, but only indirectly, on the S-structure position they occupy. Compare strong quantifiers, which necessarily take a quantificational reading irrespective of the position they occupy. Correspondingly, their scope properties are independent of the position they occupy.

References

Abney, S. (1987). *The English Noun Phrase In Its Sentential Aspect*, PH.D., MIT, Cambridge, Mass.

Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*, The University of Chicago Press, Chicago.

Barwise, J. and R. Cooper (1981). "Generalized Quantifiers and Natural Language", *Linguistics and Philosophy* 4, 159-219.

Beghelli, F. (1993). "A Minimalist Approach to Quantifier Scope", Proceedings of NELS 23.

Ben-Shalom, D. (1993). "Object Wide Scope and Semantic Trees", manuscript, UCLA.

Carlson, G. (1989). "The Semantic Composition of English Generic Sentences", in G. Chierchia, B.H. Partee and R. Turner (eds),. *Property Theory, Type Theory, and Semantics*, Reidel, Dordrecht.

Chomsky, N. (1977). "On Wh-Movement", in P.Culicover, T.Wasow, and A.Akmajian (eds), *Formal Syntax,* Academic Press, New York, 71-132.

Chomsky, N. (1981). *Lectures on Government and Binding*, Foris Publications, Dordrecht.

Chomsky, N. (1986b). *Barriers*, Linguistic Inquiry monograph, n° 13, MIT Press.

Chomsky, N. (1989). "Some Notes on Economy of Derivation and Representation", in I. Laka & A. Mahajan, *MIT Working Papers in Linguistics, vol. 10.*

Chomsky, N. (1992). "A Minimalist Program for Linguistic Theory", MITWPL, Cambridge, Mass.

Cinque, G. (1990). *Types of A'-Dependencies*, Linguistic Inquiry Monograph, 17.

Dasgupta, P. (1988). "Bangla Quantifier Extraction, Unaccusative in Situ, and the ECP", *Linguistic Inquiry* 19, 4.

Delfitto, D. (1993). "Bare Plurals and Beyond: Some Hypotheses on the Interaction between Morphology and Logical Form", talk given at *Interfaces de la syntaxe*, colloquim held in Paris, December 1993.

Diesing, M. (1992). *Indefinites,* Cambridge, Mass., MIT Press.

Diesing, M. (1993). "NP Types and Conditions on Interpretation", talk given at the Rijksuniversiteit Utrecht, Leiden, October 1993.

Dobrovie-Sorin, C. (1988). "What Does QR Raise", *Logical Form* (Program and

Abstracts), Tilburg University.

Dobrovie-Sorin, C. (1990). "Clitic Doubling, WH Movement and Quantification in Romanian", *Linguistic Inquiry* 21, 3.

Dobrovie-Sorin, C. (1992). "LF Representations, Weak Islands and the ECP", GLOW abstract.

Dobrovie-Sorin, C. (1993,).*The Syntax of Romanian*, Mouton de Gruyter.

Enç, M. (1991). 'The Semantics of Specificity', *Linguistic Inquiry* 22:1, 1-25.

Fiengo, R., J. Huang, H. Lasnik and T. Reinhart (1989). "The Syntax of Wh-in-Situ", Proceedings of WCFLL 7.

Fodor, J. and Sag, I. (1981). "Referential and Quantificational Indefinites", *Linguistics and Philosophy* 5, 355-398.

Frampton, J. (1990). "The Fine Structure of WH Movement and the Proper Formulation of the ECP", ms., Northeastern University.

Haik, I. (1984). "Indirect Binding", *Linguistic Inquiry* 15, 185- 223.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases,* PH.D., Univ. of Mass., Amherst.

Heim, I. (1987). "Where Does the Definiteness Restriction Apply?" Evidence from the Definiteness of Variables", in E.J. Reuland and A.G.B. ter Meulen (eds.), *The Representation of (In)definiteness*, MIT Press, 21-43.

Herburger, E. (1993). "Focus and the LF of NP quantification", Proceedings of SALT 3.

Higginbotham, J. (1985,). 'On Semantics', *Linguistic Inquiry* 16:4, 547-593.

Higginbotham, J. (1987). "Indefiniteness and Predication", in Reuland, E. and Alice ter Meulen (eds.), *The Representation of (In)definiteness*, MIT Press.

de Hoop, H. (1992). *Noun Phrase Interpretation and Case Assignment*, Groningen.

Jaeggli, O. (1982). *Topics in Romance Syntax*, Foris, Dordrecht.

Jong, F.M.G. de (1987). "The Compositional Nature of (In)definiteness", in E. J. Reuland and A. G. B. ter Meulen (eds.), *The Representation of (In)definiteness*, MIT Press.

Keenan, E. (1987). "A Semantic Definition of 'Indefinite NP'", in E. Reuland and A. ter Meulen (eds.), *The Representation of (In)definiteness*, MIT Press, 286-319.

Keenan, E. and J. Stavi (1986). "A semantic characterization of natural language determiners", *Linguistics and Philosophy* 9, 253-326.

Kratzer, A. (1989). "Stage and Individual Level Predicates", in *Papers on Quantification*, Univ. of Mass, Amherst.

Kroch, A. (1990). "Amount quantification, referentiality, and wh-movement",

GLOW abstract.

Krifka, (1992). "Definite NPs aren't Quantifiers", *Linguistic Inquiry* 23, 1, 156-164.

Landman, F. (1989). "Groups", *Linguistics and Philosophy.*

Link, G. (1983). "The Logical Analysis of Plurals and Mass Terms: a Lattice-Theoretic Approach", in Bauerle, Schwarze, and von Stechow (eds.), *Meaning, Use, and the Interpretation of Language,* de Gruyter, Berlin.

Liu, F-H. (1990). *Scope Dependency in English and Chinese*, Ph.D. Diss., UCLA.

May, R. (1985). *Logical Form. Its Structure and Derivation*, MIT Press.

Milsark, G. (1977). "Toward an Explanation of Certain Peculiarities in the Existential Construction in English", *Linguistic Analysis* 3, 1-30.

Longobardi, G. (1993). "Proper Names and the Theory of N-movement in Syntax and Logical Form", to appear in *Linguistic Inquiry.*

Obenauer, H. (1984)-(1985). "On the Identification of Empty Categories", *The Linguistic Review* 4, 153-202

Obenauer, H.-G. (1992). "L'interprétation des structures *wh* et l'accord du participe passé", in H.-G. Obenauer et Anne Zribi-Hertz (eds.), *Structure de la phrase et théorie du liage*, Actes du colloque international du GDR 120 (CNRS), 9-11 mars 1990, Presses Universitaires de Vincennes.

Partee, B. H. (1988). "Many Quantifiers", *Proceedings of ESCOL.*

Partee, B. H. (1987). "Noun Phrase Interpretation and Type-Shifting Principles", in J. Groenendijk, D. de Jongh, and M. Stokhof (eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Foris, Dordrecht.

Pesetsky, D. (1987). "Wh-in-situ: Movement and Unselective Binding", in E.J. Reuland and A.G.B. ter Meulen (eds), *The Representation of (In)definites*, MIT Press, 98-130.

Pollock, J.Y. (1989). "Opacity, Genitive Subjects, and Extraction from NP in English and French", *Probus* 1.2., 151-162.

Reinhart, T. (1987). "Specifier and Operator Binding", in E.J. Reuland and A.G.B. ter Meulen (eds.), *The Representation of (In)definites*, MIT Press, 130-168.

Reinhart, T. (1992a). "Interpreting WH-in-situ", manuscript.

Reinhart, T. (1992b). "Wh-in-situ: an apparent paradox", Proceedings of the 8th Amsterdam Colloquium.

Reuland, E. and A. ter Meulen (eds.) (1987). *The Representation of (In)definiteness*, MIT Press.

Riemsdijk, H. van and E.Williams (1981). "NP-structure", *The Linguistic Review* 1, 171-217.

Rizzi, L. (1990). *Relativized Minimality*, Linguistic Inquiry Monograph 16, MIT Press.

Scha, R. (1981). "Distributive, Collective and Cumulative Quantification", in J. Groenendijk et al. (eds.), *Formal Methods in the Study of Language.* Amsterdam: Mathematical Center. [Reprinted in J. Groenendijk et al. (eds.) 1984, *Truth, Interpretation and Information.* Dordrecht: Foris]

Swart, H. de (1992). "Intervention Effects, Monotonicity and Scope", in C. Barker and D. Dowty, (eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory* (OSU WPL 40), Columbus, the Ohio State University.

Szabolcsi, A. (1992). "Weak Islands, Individuals and Scope", in C. Barker and D. Dowty (eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory* (OSU WPL 40), Columbus, the Ohio State University.

Szabolcsi, A. and F. Zwarts (1990). "Semantic Properties of Composed Functions and the Distribution of WH-phrases", in Stokhof-Torenvliet (eds.) *Proceedings of the 7th Amsterdam Colloquium*, december 19-22, 1989, ITLI, 529-555.

Szabolcsi, A. and F. Zwarts (1993). "Weak Islands and Algebraic Semantics for Scope Taking", *Natural Language Semantics.*

Westerstahl, D. (1985). "Logical Constants in Quantifier Languages", *Linguistics and Philosophy* 8, 387-413.

Williams, E.S. (1986). "A Reassignment of the Functions of LF", *Linguistic Inquiry* 17, 2, 265-301.

Adverbial Quantifiers, Negation, and Stress Rules Effects. [1]

Marie-Thérèse Vinet
Université de Sherbrooke

1. *Introduction*

The adverbial quantifiers *beaucoup* (a lot) in Standard French (henceforth SF) and *benben* (a lot) in Quebec French (henceforth QF) may appear as near synonyms when presented only with the data in (1):

(1) a. J'ai pas rencontré beaucoup/benben d'amis
I didn't meet a lot of friends
b. J'ai pas beaucoup/benben rencontré d'amis
'I didn't a lot meet of friends'

However, a more careful study of their syntactic representation reveals that only *benben* can appear sentence finally, in (2a), when it is in a triggering environment. As can be observed in (2b), *benben* must be in the scope of Neg in order to be able to appear postnominally:

---

1. This study was partially funded by SSHRC (410-93-0838) and FCAR (94ER0401). Earlier versions of this paper have been presented at UQAM, the Micro-Parametric Syntax in St-John, New-Brunswick, at the Going Romance Colloquium in Utrecht and at the University of Venice. I would like to thank the audiences at those meetings for their comments. I am also endebted to Guglielmo Cinque, Anna Cardinaletti and Maria Luisa Zubizarreta for helpful comments and suggestions. The usual disclaimers apply.

University of Venice
Working Papers in Linguistics
vol. 5, n.1; 1995

(2) a. J'ai pas rencontré d'amis benben /*J'ai pas rencontré d'amis beaucoup
'I didn't meet (of) friends a lot'
I didn't meet a lot of friends
b. *J'ai rencontré d'amis benben
'I met (of) friends a lot'

Vinet (1994) has identified a few other adverbial quantifiers of the same type in the grammar of QF, such as *J'ai pas rencontré d'amis le diable/trop trop* (I have not met friends the devil/too too much) and this situation clearly raises a certain number of theoretical questions, namely for studies on (dialectal) variation. In this paper, I would like to provide an explanation for the sharp distributional contrasts observed in (2). It is suggested that these differences in syntactic representation simply follow from different lexical specifications linked to these items and, more particularly, from general principles of Universal Grammar having to do with negative polarity, stress rules effects and principles of Economy.

Since Pollock (1989) and Chomsky (1991), word order variation concerning adverbs has been dealt with by a head movement process around the adverbial element which always remains in situ. [2] Adverbs, which usually lack agreement, do not need to check their features. [3] Hence, by the principles of Economy (Greed), they cannot move. Furthermore, as proposed by Cinque (1994, 1995), I adopt the view that adverbs are specifiers and not adjuncts as in the Emonds/Pollock hypothesis. [4] Certain adverbs, in the space delimited by past participle movement, are therefore generated in the specifier position of various functional heads. The

2. See Williams (1995) for an opposite view where it is held that word order differences between French and English do not follow from properties of V movement but are rather due to a "difference in the combinatorics of the basic elements of the language", such as the position of adverbs, of adjectives, of heads in compounds, etc.

3. Adverbial quantifiers like *beaucoup/benben* must move at LF, after Spell-Out, to check their scopal properties. Some quantifiers, *tout* and *tutto*, can move in syntax. Sportiche (1988) proposes for *tous* an analysis with NP movement and a stranded Q but Déprez (1991) has a movement rule to Spec,AGRo for the bare quantifiers *tout* and *rien*. Giusti (1993) argues for an obligatory movement of *tutto* in Spec,DP in order to account for the word order variation in the sequences *tutti i ragazzi* (all the boys) and *i molti ragazzi* (the many boys).

4. Chomsky (1993) adopts a Larsonian VP shell to account for word order with adverbials in English.

fundamental motivation behind this hypothesis is to explain the relatively rigid order among any two pairs of them without violating basic contraints on bare phrase structure in UG (Chomsky (1994), Kayne (1993)). In a minimalist approach, an adverbial cannot be adjoined to a two segment category [XP, XP] where XP has a semantic role at LF.

The paper is organized as follows. Section 2 discusses various aspects of the identification of *benben* in the lexicon. Section 3 illustrates how a "shifted" position is only possible when *benben*, a polarity sensitive item, is in the scope of a sentential Negation element such as *pas*, *plus*, *jamais* as well as the inherently negative preposition *sans*. The constituent Neg *pas*, within DP, cannot license a polarized *benben*. It is also observed that N-words like *personne* or *rien* differ in significant ways from the above negated adverbs in that they cannot license as freely NPIs like *benben* in the grammar of QF. The remainder of the paper is devoted to a discussion of stress rules effects on adverbs, as in Cinque (1993, 1995), Cardinaletti & Starke (to appear), and suggests an analysis where such items can be generated in various positions depending on stress rules, in order to explain word order variation with polarized and focussed items such as *benben*.

2. *The lexical identification of 'benben'*

Chomsky (1991, 1994) mentions that once an articulated system of features for a certain lexical element (noun, verb, etc) has been properly defined, some larger units constructed of these items will be available following more general principles of Universal Grammar. In order to correctly identify invariants in UG, it is indeed necessary to properly identify the lexical type of *benben*, its categorial status and its semantic properties.

Let us note, first, that the category labels "adverb" and "quantifier" are taxonomic artifacts in grammar. Morphologically, they often are the result of derived categories, such as nouns or adjectives. Moreover, adverbs and quantifiers can share a certain

number of grammatical properties: lack of agreement, non argumental status, mobility in the sentence, etc.

As is obvious from its morphology, *benben* is a reduplicate form whose syntactic representation varies considerably from the non reduplicate form *ben* in QF or even *bien* in SF.

Junker & Vinet (1995) have demonstrated that *benben* is not a manner adverb nor an assertion adverb. It can only be interpreted as an adverb of quantification as illustrated in (3-5) below:

(3) a. Elle chante (pas) bien/ben — MANNER ADVERB
She sings (not) well
b. Je veux bien/ben qu'elle chante — ASSERTION ADVERB
Yes I want her to sing

(4) a. *Elle chante benben (well) — MANNER ADVERB
b. *Elle chante pas benben (not well)
c. Elle chante pas benben — QUANTIFICATION
She does not sing a lot

(5) a. *J'veux benben qu'elle chante — ASSERTION ADVERB
I want very much that she sings
b. J'veux pas benben qu'elle chante — QUANTIFICATION
I don't want very much for her to sing

In a negative context, as in (4c), *benben* can be interpreted as an adverb which quantifies over multiple events denoted by the verb but this interpretation is not available in the positive context of (4a). Moreover, only the reduplicate form *benben* can appear sentence finally in the following context:

(6) a. J'ai pu de cheveux benben
'I don't have any more hair a lot'
I don't have a lot of hair any more

b. J'ai pu ben de cheveux/ *de cheveux ben
'I don't have any more a lot of hair/ *of hair a lot'
I don't have a lot of hair any more'

These facts in (6) will be further discussed in section 4 below. Note that if *ben* can also be an adverb of quantification in the grammar of QF, its distribution differs in that only *benben* is an emphatic form with prominent stress (Cinque 1993). In other words, and following the classification of pronominal forms and adverbs proposed by Cardinaletti & Starke (to appear), one could posit that *ben* is a deficient form and *benben* a strong form. Let us now turn to an analysis of the lexical categories which *benben* and *beaucoup* can be extensions of in their respective grammars.

2.1. *'Benben' vs 'beaucoup'*

The basic oppositions between *très* (very) and *beaucoup* (a lot) in French are well known. *Beaucoup* is usually compatible with verbs and nouns while *très* only appears with adjectives and adverbs. The distribution of *benben* is sometimes more closely related to the intensifier *très* than to the adverbial quantifier *beaucoup*. The paradigm in (7) below illustrates my point:

(7) a. **Adjective**:
(i) Jeanne est benben/très patiente
Jeanne is very patient
(ii) (*) Jeanne est beaucoup patiente [5]
'Jeanne is a lot patient'

5. The variation in acceptability in (7a,ii) is attributed to the fact that *beaucoup* can be compatible with an adjective in some dialects, namely in some varieties of QF (Julie Auger, p.c.). Gross (1977:156) mentions *Luc est beaucoup grand* (Luc is a lot (very) tall) and Frei (1929:151) points out the example *C'est beaucoup moche* (It is a lot (very) ugly).

b. **Adverb**:
(i) Jeanne vient benben/très souvent par icitte
Jeanne comes very often around here
(ii) *Jeanne vient beaucoup souvent par ici
'Jeanne comes a lot often around here'

c. **Past participle**:
(i) Jeanne a benben/*très dormi
'Jeanne has a lot/very slept'
(ii) Jeanne a beaucoup dormi
Jeanne has slept a lot

d. **Noun**:
(i) *Jeanne a benben/très d'amis (cf. Jeanne a benben des amis)
'Jeanne has very of friends'/ *Jeanne has a lot (of) friends*
(ii) Jeanne a beaucoup d'amis
Jeanne has a lot of friends

If quantifiers can inherit categorial features from their complements, as suggested in Cinque (1995), then *benben* could possibly be considered an extension of an adjective (*benben patiente*), an adverb (*benben souvent*), a past participle (*benben dormi*) and a noun (*benben des amis*). These categorial features would differ for *beaucoup*. This quantifier can only be an extension for a past participle (*beaucoup dormi*), a noun (*beaucoup d'amis*) and degree adverbs (*beaucoup trop mou* (much too soft)). It cannot extend an adverb (**beaucoup souvent*) nor an adjective (**beaucoup patiente*), at least in SF.

A second difference is related to the fact that, in a positive context, *benben* can appear with a partitive determiner while *beaucoup* is uniquely followed by the bare partitive preposition *de*:

(8) a. J'ai benben de la misère
b. *J'ai benben de misère
c. *J'ai beaucoup de la misère

d. J'ai beaucoup de misère
I have a lot of difficulties

This distinction could be straightforwardly explained by the fact that *beaucoup* licenses an existentially quantified N, whereas *benben* cannot. Contrary to *beaucoup* also, *benben*, just like *souvent* (often), is always optional in both positive or negative contexts, [6] illustrating its strong "adverbial" properties, as indicated in (9a,b):

(9) a. J'ai pas (benben/souvent) d'amis ici
'I don't have (a lot of/often) friends over here'
b. J'ai (benben/souvent) des amis
'I have (a lot of/often) friends'

Nevertheless, in a positive context, *benben* can only be followed by the partitive determiner *des* with a restrictive, specific reading, as in (10a). It can never appear with an indefinite *des* bearing a contrastive reading, as illustrated in (10b), from Junker & Vinet (1995):

(10) a. Je ne veux pas benben des pommes (qui sont là)
'I don't want very much (of) the apples (which are there)'
b. *Je ne veux pas benben des pommes (mais des oranges)
'I don't want very much apples (but oranges)'

These facts clearly indicate that *benben* is a polarity sensitive item whose distribution and interpretation is sensitive to negative versus affirmative contexts, just like *any* (cf. Progovac (1993)). This situation is also reminiscent of a relatively similar phenomenon with the quantifier *not much* in English which is normally in suppletive variation with *a lot*. As noted by Klima (1964:283), only *a lot* can appear in a positive sentence, as exemplified in (11) below:

---

6. A different situation is observed with the temporal adverbial quantifier *toujours* which never binds an existentially quantified N:

(i) a. *J'ai pas toujours de temps/d'amis
b. J'ai (pas) toujours du temps/des amis
'I don't always have time/friends'

(11) a. Writers don't accept suggestions much these days (Klima)
b. *Writers accept suggestions much these days
c. Writers accept suggestions a lot these days.

Moreover, it can be observed that the syntactic properties of *benben* differ whether it is in the scope of negation or not. Kayne (1981) has shown that French ($_{QP}$ (e) *de N*) structures were preceded by an empty initial QP and that *beaucoup* and *pas* could fill this position. However, *benben* and other NPI adverbs of the same type in the grammar of QF can never bind an existentially quantified N on their own, as exemplified in (12):

(12) a. J'ai pas trop trop/le diable d'amis
'I don't have too too many/the devil of friends'
b. *J'ai trop trop/le diable d'amis
'I have too too many/the devil (of) friends'

There is also a sharp contrast between the two forms *beaucoup* and *benben*, since *benben*, when it is a negative polarity item, must obligatorily be licensed and c-commanded by a sentential Neg. A sentential Neg negates a predicate or a proposition:

(13) a. Beaucoup/trop d'amis sont venus
(Too) many friends came over
b. *Benben/trop trop d'amis sont pas venus
(Too) many friends did not come over
c. ??Pas benben/trop trop d'amis sont venus
Not (too) many friends came over
d. *Pas d'amis benben/trop trop sont venus
'Not of friends a lot/too too much came over'
Not (too) many friends came over

In (13c,d) *pas* does not have sentential scope and the position of *benben/trop trop* at the end of the "subject" phrase in (13d) is therefore rejected.

Furthermore, if *beaucoup* can refer to a [+human] entity when it appears in subject position, *benben* cannot get this interpretation, given its stronger adverbial properties: *Beaucoup/*benben sont venus* (Many (guests) came over).

I therefore suggest, as developed in section 3, that only certain Neg items can be both triggers and binders for polarity sensitive adverbs like *benben* in QF.

To sum up this section on the lexical identification of *benben*, as compared to *beaucoup*, it has been demonstrated that both items present quite different lexical properties. It was indicated, for instance, that they are not extensions of the same classes of lexical categories and that only *beaucoup* can license an existentially quantified N. If both items are interpreted as measure adverbs and quantifiers, only *benben* is sensitive to polarity. It is assumed that this last property plays an important role in the syntactic representation of these adverbial forms and we now turn to a more detailed study of the licensing of such NPIs.

3. *Sentential Neg as a licenser*

The negation forms which can license such NPIs in QF are limited. There is the minimal negation quantifier *pas*, the aspectual quantifier *plus* (pronounced *pu* in QF) and *jamais*, the temporal quantifier:

(14) a. J'ai pas/pu/jamais benben/trop trop de vingt-cinq cennes
b. J'ai pas/pu/jamais de vingt-cinq cennes benben/trop trop
'I have not/no more/never (a lot) of twenty-five cents (a lot)'

It was observed, in (7) above, that *benben* can appear in a positive context as an extension of an adjective, an adverb and a past participle. It is important to note, however, that *benben* appears sentence-finally in a triggering environment only, as represented in (15c):

(15) a. benben XP / b. *XP benben / c. Neg XP benben

The examples in (16) below exemplify the situation when XP is an adjectival form:

(16) a. Elle est benben fine /*elle est fine benben
She is very nice/ she is nice very much
b. Elle est pas fine benben
She is not nice very much

The same pattern applies with adverbs and past participles. I will therefore use this distinction in syntactic representation as a test for the identification of Neg licensers for such NPIs.

For instance, the Neg licenser can never be a head such as *ne* or *non*, as in (17). *Ne* or *non* can never bind an existential quantified N anyway:

(17) a. *Je n'ai benben d'amis/ *Je n'ai d'amis benben
'I Neg have a lot of friends/ I Neg have of friends a lot'
b. *J'ai rencontré non benben d'amis, mais benben d'ennemis
'I have met not a lot of friends, but a lot of enemies'

Recall that *ne*, which is a (omittable) marker of sentential negation, can appear with N-words like *personne* or *rien* in SF:

(18) Elle (n')aime personne/rien
She likes nobody/nothing

The inherently negative preposition *sans*, which belongs to a different class of inherently negative items, can license such forms (cf. 19b) but the negative adverb *rarement* (seldom) cannot, as demonstrated in (20b):

(19) a. Il est parti sans benben/trop trop/le diable réfléchir
b. Il est parti sans réfléchir benben/trop trop/le diable
'He left without (much/the devil) thinking (much/the devil)'

(20) a. Elle invite rarement benben d'amis
b. ?*Elle invite rarement d'amis benben
'She seldom invites (a lot of) friends (a lot)'

3.1. *N-words*

The situation with N-words is not always as clear. This can be observed from the following acceptability judgments in (21-22) and (23-24) in QF:

(21) a. *Elle aime personne benben (QF)
She likes nobody a lot
b. *Elle aime rien benben (QF)
She likes nothing a lot

N-words are well-known to be a combination of nouns and quantifiers. Moreover, *personne* and *rien* are quantifiers which cannot be quantified by any positive or negative measure quantifiers, as illustrated in (22):

(22) *Elle aime benben/beaucoup/trop/pas du tout/énormément personne/rien
'She likes a lot/too much/not at all/enormously nobody/nothing'

However, in (23b) and (24b), *personne* and *rien*, contrary to their positive equivalents *tout le monde* and *quelque chose*, seem to be able to license the "shifted" position around the negative polarity item *benben*, as exemplified in (23) and (24):

(23) a. Personne a benben aimé ça
b. Personne a aimé ça benben
Noone liked it a lot
c. Tout le monde a benben aimé ça
d. *?Tout le monde a aimé ça benben
'Everybody liked a lot it'

(24) a. J'ai rien/quelque chose de benben intéressant à te montrer
I have nothing/something very interesting to show you
b. J'ai rien d'intéressant benben à te montrer
'I have nothing interesting a lot to show you'
c. *J'ai quelque chose d'intéressant benben à te montrer
'I have something interesting a lot to show you'

The reason for these different results with N-words remains mysterious for the moment. It could probably be explained by the particular categorial status of these items which cannot always act as full Neg items. I leave the question open, pending further research on N-words (cf. for instance, Déprez (1995)).

3.2. *'Pas' as a constituent Neg item*

The negated item *pas* can sometimes be used as a modifier of another degree form, as in the following sequence from SF:

(25) J'ai conclu une pas si vilaine affaire
I have driven a not so terrible bargain

It is then interpreted as a constituent Neg and not as a sentential Neg. The obligatory absence of *ne*, as illustrated in (26), shows that the scope of *pas* is not sentential:

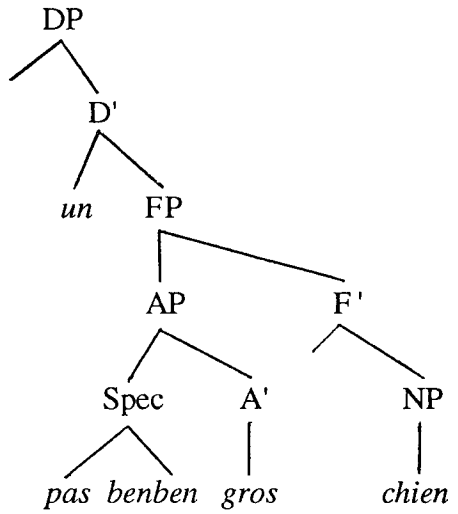
(26) Je (*n')ai conclu une pas si vilaine affaire

The same situation is observed with *pas* as a modifier of *benben* when both appear as specifiers of an adjectival form in the space delimited between D and N, as in (27) below, illustrated in the configuration (28):

(27) a. Elle avait un pas *benben/trop trop* gros chien
'She had a not too too much big dog'

b. *Elle avait un pas gros chien *benben/trop trop*
'She had a not big dog too too much'

(28)



This specifier position is only available for *pas*. It is not a slot which can be used by other sentential negated adverbs such as *plus, jamais*:

(29) *Elle avait un pus/jamais benben gros chien
'She had a no more/never very big dog'

There exists the adverbial quantifier *du tout* in SF which presents almost all of the properties of *benben* when it is a polarized adverbial item. It can indeed appear in the same positions and it can be an extension of the same lexical categories. However, *pas du tout* differs from *pas benben* in that it cannot appear as a specifier of an adjectival projection within DP:

(30) a. *Elle a un pas du tout gros chien
b. *Elle a un pas gros chien du tout
She has a not at all big dog

In other words, the constituent Neg *pas*, within DP, cannot license a polarized adverbial *benben*.

3.3. *'Benben', 'beaucoup' and the linear order of adverbial phrases*

In order to illustrate once more that *benben* is not a synonym of *beaucoup*, it can be shown that *benben*, contrary to *beaucoup*, does not occupy a fixed slot in the relatively rigid word order proposed for adverbs by Cinque (1995). The position occupied by *beaucoup* is illustrated in (31):

(31) a. J'ai pas toujours beaucoup tout bien nettoyé (SF)
'I have not always a lot all well cleaned'
b. *J'ai pas beaucoup toujours tout bien nettoyé
'I have not a lot always all well cleaned'
c. *J'ai pas toujours tout bien beaucoup nettoyé

On the contrary, *benben*, as a polarized item triggered by a sentential Neg, can occupy various positions in the predicate. These positions are given in (32) below:

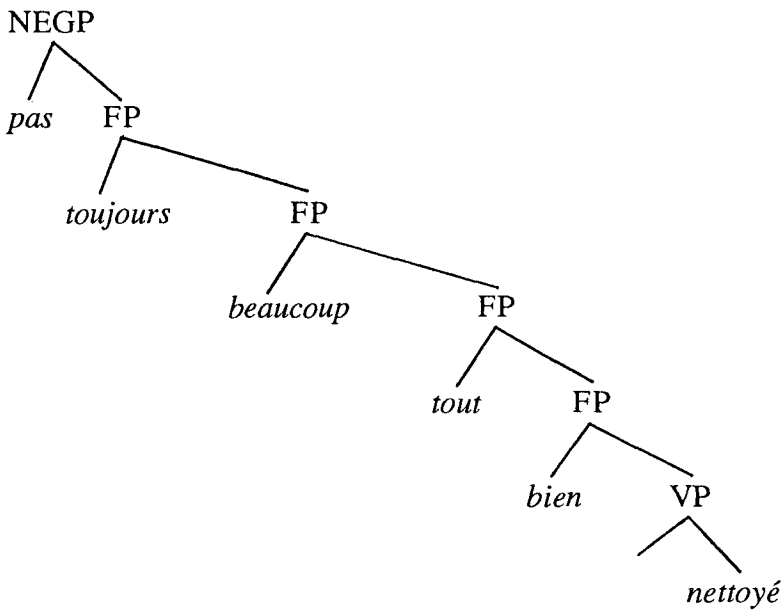
(32) a. J'ai pas benben souvent tout ben nettoyé
'I have not very often all well cleaned'
b. J'ai pas souvent benben tout ben nettoyé
'I didn't often a lot all well clean'
c.*J'ai pas benben tout ben souvent nettoyé
'I did not a lot all well often clean'

Moreover, *benben* cannot always appear in the extension of all adverbs. *Toujours*, for instance, seems semantically incompatible with it. This situation can perhaps explain why *toujours* cannot as well be modified by the synonymous degree word *très*. Compare the ill-formed **très oujours* (very always) with *très souvent/longtemps* (very often/for a very long time):

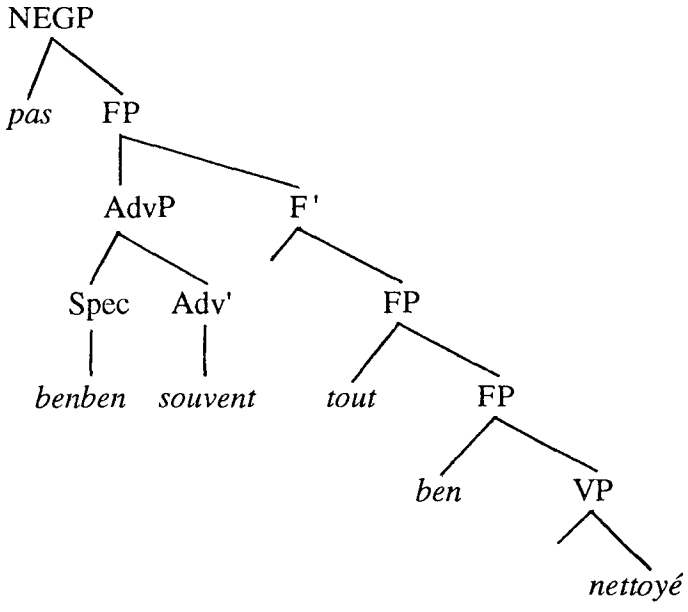
(33) a. ?*J'ai pas benben toujours tout ben nettoyé
'I didn't a lot always all well clean'
b. *J'ai pas toujours benben tout ben nettoyé
'I didn't always a lot all well clean'

The differences between *benben* and *beaucoup*, noted in (31) and (32) above, linked to the fact that only *benben* can be an extension (or a specifier) of an adverb, are represented in (34a,b):

(34) a.



b.



In other words and to sum up this last section, it was observed that *benben* is a

measure adverb of quantification which can be a polarized item. It can appear as an extension of different lexical categories (VP, NP, AP and ADVP). However, it was demonstrated that the sentence-final position is only possible when *benben* is in the scope of a sentential negation which acts as a trigger.

4. *Stress Rules Effects*

What distinguishes these adverbial quantifiers in QF which can bear prominent stress in the sentence from other adverbial quantifiers like *beaucoup* or *tout* in SF which can sometimes bear prominent stress under certain conditions?

First, it can be observed that all the adverbs which can appear sentence-finally in this category, in QF, must bear a reading that falls within the standard negative polarity environments. Adverbial quantifiers like *beaucoup* or *tout* in SF are not interpreted as negative polarity items depending on a trigger. [7] They are not polarity sensitive items. More importantly, it can be observed that these adverbs in QF are either obligatory reduplicate forms or words with at least two syllables. There is always a falling intonation on the last syllable when it appears sentence-finally. Adverbs with one syllable only, such as *ben* or *trop*, can rarely bear the greatest prominence in the sentence (cf. Cinque 1993), as illustrated in (35):

(35) a. J'ai pas ben aimé ça
b. *J'ai pas aimé ça ben
'I didn't (much) like it (much)'
c. Il a pas trop vu le changement

---

7. Note that *tout*, contrasts with *du tout* (SF) (at all), which is a negative polarity item. As already mentioned in section 3.2 above, *du tout* can also appear, just like *benben* in various positions within the scope of a sentential Neg:

(i) J'ai pas (du tout) rencontré (du tout) d'amis (du tout)
'I didn't (at all) meet (at all) (of) friends (at all)'

d. *Il a pas vu le changement trop
'He didn't (too much) see the change (too much)'

Kayne (1975:38) mentioned that certain (one syllabe) quantifiers could appear sentence-finally if they were stressed or made "heavier":

(36) a. Il a repris ??(presque) tout (Kayne (1975))
He took back almost everything
b. Elle ne va lire *(absolument) rien
She's going to read absolutely nothing

However, if as argued by Cinque (1993:245), stress prominence in a phrase is a mere reflexion of depth of embedding, how can one explain that one word syllable adverbs or non polarized quantifiers can appear sentence-finally in the following examples:

(37) Il dort$_i$ ben / beaucoup / assez / trop t_i
He sleeps well/a lot/enough/too much

In this case, the adverb or the quantifier does not bear the prominent stress of the sentence but rather a nuclear stress. The movement of the verb takes place in the computational system, before Spell-Out.

Moreover, as argued by Cardinaletti & Starke (to appear) adverbs can be classified, just like pronominal forms, as clitics or deficient forms and as strong forms. However, clitic adverbial heads do not exist in French. Deficient adverbial forms can appear sentence-finally (following a past participle) only if they are given more strenght through coordination (38a), stress effects (38b) or c-modification (38c):

(38) a. Elle a parlé ben *(pis juste assez) (QF)
She spkoke well and just enough

b. J'ai aimé ça BEAUCOUP [8]
I liked it A LOT
c. Ils ont chanté *(beaucoup) trop (SF)
'They sang much too much'

However, *benben*, contrary to *beaucoup* cannot appear sentence-finally in the following positive contexts even if it is stressed, coordinated or made heavier, as illustrated in (39):

(39) a. *J'ai aimé la tarte BENBEN
I liked the pie A LOT
b. *J'ai aimé ce film-là benben pis ben longtemps après
I liked this movie a lot and for a long time after
c. *J'ai aimé ça assez benben
'I liked it enough a lot'

These facts in (39) clearly indicate that a different mechanism is at work with these forms. It can also be observed that the positioning of *benben* sentence-finally is linked not only to intonational constraints but also to semantic constraints on the type of predicate. Junker & Vinet (1995), for instance, demonstrate that *benben* is always ruled out in both positive or negative contexts with verbs of achievement like *arriver* (40a) and in a positive context only, with verbs of activity (*courir*, *marcher*) or verbs expressing an event that lasts like *dormir*, in (40b,c):

---

8. In SF, *beaucoup* can appear sentence-finally, preceded by the functional head *de*, only when it is a modifier of a degree word, as exemplified in (i):

(i) a. Elle est plus intelligente que lui de beaucoup (SF)
'She is more intelligent than him of a lot'
b. Elle est beaucoup plus intelligente que lui
'She is a lot more intelligent than him'

This sequence is not possible with *benben* because this item cannot be an extension of a degree word, as illustrated in (ii):

(ii) a. *Elle est benben plus intelligente
'She is very more intelligent'
b. *Elle est plus intelligente que lui de benben

(40) a. *Ils arrivent (pas) benben
'They arrive (not) a lot'
b. *Ils dorment/courent/marchent benben
They sleep/run/walk a lot
c. Ils dorment/courent/marchent pas benben
They don't sleep/run/walk a lot

The sentence (40c) is acceptable because verbs like *dormir* allow a mass internal quantification (Bach 1981). In other words, this quantification expresses the fact that "sleeping", "running" and "walking" can be either a continuous or an irregular activity.

It can therefore be concluded from these facts that *benben* can bear a heavier stress in the sentence only if it is focussed through Negation. This straightforwardly explains the absence of *benben* in sentence-final position when it is not c-commanded by Neg.

4.1. *A syntactic movement prosodically motivated?*

In this section, I would like to discuss briefly the question of how to account for the sentence final position of *benben* in sentences where it is a polarized item. The relevant facts are found in the following examples:

(41) a. Elle est pas benben fière de son chien
She is not very proud of her dog
b. Elle est pas fière de son chien benben
She is not proud of her dog very much

Within both the minimalist framework of Chomsky (1993) and the Antisymmetry approach of Kayne (1993), *benben* cannot move downwards or be right-adjoined without violating major constraints. Furthermore, there is no motivation in the minimalist approach for the movement of the sequence *fière de son chien* in (41a).

This sequence of words does not need to check any of its morphological features through a movement around *benben.*

In Chomsky (1994), a strong hypothesis defines Greed as a "self-serving" principle. A category moves only to satisfy its own feature-checking needs.

A redefinition or a reinterpretation of Greed could be suggested, as in Zubizarreta (1994), who points out that some syntactic operations may be prosodically motivated.

A movement to the left could be proposed, as in (42), of a non-branching category X^0 or XP, to a Spec position, as illustrated in (43), in a way similar to the stranding of floating quantifiers (cf. Sportiche (1988)):

(42) a. Elle est pas [fière]$_i$ benben t_i de son chien
b. Elle est pas [fière de son chien]$_i$ benben t_i

(43) $_{NEGP}$[pas ... $_{XP}$[fière de son chien X'] $_{FP}$[benben t_i F']]

However, this movement in (43), as already mentioned, would not be motivated by any feature-checking operation. Moreover, such an approach does not explain why that movement is not obligatory and the stranding of *benben*, when it appears in a triggering environment, is always optional. The term "optional" is used here to indicate that an intonational pattern may appear on the last syllable of *benben* or not. When it appears in a non-final position, *benben* has a different intonational structure and this information is not contained in the lexical item. In other words, this movement would rather be motivated to satisfy the needs of another category: the stressed polarized item *benben*. Finally, another problem with the movement analysis concerns the target of movement. It is not yet clear how one could identify the target node $_{XP}$[Spec,XP] in (43).

As can be observed, such an approach is not without cost in the minimalist program. It entails, indeed, that a syntactic rule would have access before Spell-Out to some information at PF. An unwelcome step in the proposed system.

I will rather suggest that such adverbial-quantifiers can surface in various positions

in the scope of a sentential Neg item. In sentence-final positions, prosodic rules on stress will only accept items which are strong forms, as defined in Cardinaletti & Starke (to appear) and items that bear the right intonational structure. This information is made available for the computation as the item is introduced into the derivation. Deficient adverbial forms like *ben*, for instance, cannot trigger such interpretations and (44a) is ruled out because the derivation does not converge at PF and it crashes. Moreover, a strong form like *benben* in a sentence-final position must appear in a triggering environment, i.e. it must be c-commanded by a sentential Neg to bear the expected intonational pattern, as illustrated by the unacceptability of (44b) (capital letters mark the heavier stress on the last syllabe of the word):

(44) a. *Elle est pas fière ben
'She is not proud very'
b. *Elle est là souvent benBEN
'She is there often a lot'
c. Elle est pas là souvent benBEN

What then distinguishes the difference in unacceptability in the following examples?

(45) a. Elle dort ben
She sleeps well
b. *Elle est fière ben (cf. Elle est ben fière)
'She is proud a lot'
c. *Il vient souvent ben (cf. Il vient ben souvent)
'He comes often a lot'

The distinction is now obvious. In (45a), we find a morphologically-driven movement of V to the left in order to check Tense and Agreement features and if *ben* then appears in final position, it does not bear the greatest prominence in the sentence. In (45b,c), on the contrary, there is no similar movement to the left of the adjectival [9]

9. If the adjective raises overtly to AGR in (45b), the situation is still different because *ben* does not extend adjectives and verbs in the same fashion. With V-movement in (45a), *ben* is an adverbial-quantifier generated in the Spec of a functional projection, whereas in (45b) *ben* is a degree modifier which is generated in Spec,AP. It can be hypothesized that this distinction is in some way

or adverbial forms, *ben* is therefore generated in final position but it cannot be interpreted as bearing a prominent stress and, as expected, the derivation cannot converge.

5. *Conclusion*

In this paper, I have identified *benben* in the grammar of QF as a polarized adverbial form which can be an extension of various kinds of lexical categories. When it extends an adjective, an adverb, it is interpreted as a modifier or a degree word. When it extends a verb, a past participle or a noun, it is then interpreted as an adverbial quantifier. Nonetheless, *benben* can only appear sentence finally when it is licensed by a sentential Neg, defined as the triggering environment.

Finally, in order to explain this "scrambling" phenomenon, which differs from the facts observed with the distribution of *beaucoup* in SF, it was proposed that such adverbial-quantifiers can be generated in various positions in the scope of a sentential negation, and namely sentence-finally, because they are focussed and negative polarized elements and, as such, they can bear the most prominent stress in the sentence. It was also acknowledged that a heavier stress on the second syllabe of *benben* is determined by its position in the clausal configuration. When the word is used in isolation or in a positive context, the intonational pattern differs.

---

responsible for the fact that *ben* can only bear a nuclear stress with V-movement to AGR and T and not with A-movement to AGR. I leave this question open for now.

References

Bach, E. (1981). "On Time, Tense and Aspect: an Essay in English Metaphysics", in Peter Cole (eds.), *Radical Pragmatics*, Academic Press.

Cardinaletti, A. and M. Starke (to appear). "The Typology of Structural Deficiency", mimeo., (to appear in) H. van Riemsdijk (eds.), Eurotyp Final volume, ESF, Mouton de Gruyter.

Chomsky, N. (1991). "Some Notes on Economy of Derivation and Representation", *Principles and Parameters in Comparative Grammar*, ed. R. Freidin, p.417-454, Cambridge, Mass. MIT Press.

Chomsky, N. (1993). "A Minimalist Program for Linguistic Theory", K. Hale and J. Keyser (eds.), *The View from Building 20*, Cambridge, MIT Press.

Chomsky, N. (1994). "Bare Phrase Structure", *MIT Occasionnal Papers in Linguistics* 5, Cambridge, MA.

Cinque, G. (1993). "A Null Theory of Phrase and Compound Stress", *Linguistic Inquiry* 24:2, 239-298.

Cinque, G. (1995). "On the Relative Order of Certain "Lower" Adverbs in Italian and French", mimeo, University of Venice.

Déprez, V. (1991). *On the Typology of Syntactic Positions and the Nature of Chains*, Unpublished Doctoral Dissertation, MIT.

Déprez, V. (1995). "The Roots of Negative Concord in French and in French Based Creoles", (to appear in) *Parametric Approaches to Creole Syntax,* M. Degraff (ed.), MIT Press, Cambridge.

Frei, H. (1929). *La grammaire des fautes*, Geuthner, Paris.

Giusti, G. (1993). *La Sintassi dei Sintagmi Nominali Quantificati*, Ph.D. dissertation, University of Venice.

Gross, M. (1977). *Grammaire transformationnelle du français, Syntaxe du nom,* Larousse, Paris.

Junker, M-O. & M-T. Vinet (1995). "Les propriétés lexicales et syntaxiques de*benben* en français québécois", *Revue canadienne de linguistique*, 40:1.

Kayne, R. (1975). *French Syntax*, Cambridge, MIT Press.

Kayne, R. (1981). "ECP Extensions", *Linguistic Inquiry*, 12.1, 93-133.

Kayne, R. (1993). "The Antisymmetry of Syntax", mimeo., CUNY.

Klima, E. (1964). "Negation in English", *The Structure of Language. Readings in the Philosophy of Language*, J. Fodor and J. Katz (eds.) p.264-323, Prentice-Hall.

Pollock, J-Y. (1989). "Verb Movement, Universal Grammar and the Structure of IP, *Linguistic Inquiry* 20, 365-424.

Progovac, L. (1993). "Negative Polarity: Entailment and Binding", *Linguistics and Philosophy* 16, 149-180.

Sportiche, D. (1988). "A Theory of Floating Quantifiers and its Corollaries for Constituent Structure", *Linguistic Inquiry* 19, 425-449.

Vinet, M-T. (1994). "Adverbial Quantifiers and Dialect Variation in a Minimalist Framework", talk given at the 18th Annual Meeting of the Atlantic provinces Linguistic Association, University of New-Brunswick, Saint-John, mimeo., Université de Sherbrooke.

Williams, E. (1995). "A Reinterpretation of Evidence for Verb Movement in French", *Verb Movement*, D. Lightfoot and N. Horstein (eds.), Cambridge University Press.

Zubizarreta, M.L. (1994). "Word Order, Prosody, and Focus", paper presented at Going Romance, Utrecht, December, 8-10.

The series is intended to make the work of the students, faculty and visitors of the Seminario di Linguistica and of the Centro Linguistico Interfacoltà of the University of Venice available for circulation.
The series can be obtained on an exchange basis with other Working Papers.
Requests for single issues should be addressed directly to the authors.

1991:

1. G. Cinque: A Null Theory of the Phrase and Compound Stress.
2. C. Poletto: The Diachronic Development of Subject Clitics in North Eastern Italian Dialects.
3. A. M. Di Sciullo: On the Structure of Deverbal Compounds.
4. A. Cardinaletti, G. Giusti: Partitive *ne* and the QP-Hypothesis. A Case Study.
5. J. Schaeffer: The Italian Child's C-System.
6. R. Delmonte, D. Bianchi: Computing Discourse Anaphora from Grammatical Representation.
7. M. Leonetti, V. Escandell-Vidal: Secondary Predication inside DP.
8. A. Moro: The Anomaly of Copular Sentences.
9. G. Longobardi: N-movement in Syntax and in LF.

1992:

1. A. Moro: A Case Study in Linguistic Variation: The Semantics of Existential Sentences.
2. G. Cinque: The Pseudo-relative and the Acc-*ing* Constructions after Verbs of Perception.
3. G. Giusti: Heads and Modifiers among Determiners: Evidence from Romanian and German.
4. R. Delmonte: Translating with ROSIE.
5. A. Cardinaletti, M. T.Guasti: Negation in Small Clauses.
6. G. Cinque: On Leftward Movement of *tutto* in Italian.
7. C. Poletto: The Analysis of the *Passé Surcomposé*: An Hypothesis on Sentence Structure Formation.
8. F. Benucci: Prepositional Particles and the Portuguese Personal Infinitive.

1993. vol. 3, n.1:

1. G. Brugger: Generic Interpretations and Expletive Determiner.
2. D. Bianchi, R. Delmonte, E. Pianta, S. Sartori: Reference Resolution by Contextual Reasoning from Grammatical Representations with GETA-RUN.
3. M. C. Figuereido Silva: Some Notes on Null Subjects in the Brasilian Portuguese Tensed Sentences.
4. G. Giusti: Enclitic Articles and Double Definiteness: A Comparative Analysis of Nominal Structure in Romance and Germanic.
5. C. Poletto: Subject Clitic / Verb Inversion in North Eastern Italian Dialects.
6. R. Zamparelli: Pre-Nominal Modifiers, Degree Phrases and the Structure of AP.

1993. vol. 3, n. 2:

1. A. Cardinaletti: On the Internal Structure of Pronominal DPs.
2. G. Cinque: On the Evidence for Partial N-Movement in the Romance DP.
3. G. Brugger and C. Poletto: On Negation in German and Bavarian.
4. P. Crisma: On Adjective Placement in Romance and Germanic Event Nominals.
5. A. Cornilescu: Notes on the Structure of the Romanian DP and the Assignment of the Genitive Case.

1994. vol. 4, n. 1:

1. L. Brugè and G. Brugger: On the Accusative 'A' in Spanish.
2. A. Carrasco Gutiérrez and L. García Fernández: Sequence of Tenses in Spanish.
3. V. Demonte: Datives in Spanish.
4. M. L. Zubizarreta: The Grammatical Representation of Topic and Focus: implications for the Structure of the Clause.

1994. vol. 4, n. 2:

1. A. Bisetto: Italian Compounds of the *Accendigas* Type: a Case of Endocentric Formation?
2. G. Brugger and M. D'Angelo: Movement at LF triggered by Mood and Tense.
3. A. Cardinaletti and M. Starke: The Typology of Structural Deficiency. On the Three Grammatical Classes.
4. A. Moro: "Pseudo-extraction" and Problems of Binding: a Case Study in the Syntax of the Predicative Nominals.