UNIVERSITÀ
CA' FOSCARI
VENEZIA

DIPARTIMENTO DI SCIENZE DEL LINGUAGGIO

A T T I
2

Contributions to the thirtieth Incontro di Grammatica Generativa

Venice, February 26-28, 2004

Edited by
Laura Brugè
Giuliana Giusti
Nicola Munaro
Walter Schweikert
Giuseppina Turano

Contributions to the thirtieth Incontro di Grammatica Generativa, *Venice, February 26-28, 2004*

Edited by Laura Brugè, Giuliana Giusti, Nicola Munaro, Walter Schweikert, Giuseppina Turano

ISBN 88-7543-095-0

In copertina: Vittore Carpaccio, *Miracolosa guarigione dell'ossesso*, Venezia Gallerie dell'Accademia. Su gentile concessione del Ministero per i Beni e le Attività Culturali

Libreria Editrice Cafoscarina
Calle Foscari, 3259, 30123 Venezia
www.cafoscarina.it

Prima edizione Dicembre 2005

Stampato in Italia presso Selecta SpA, Milano

Contents

Preface

This volume collects some of the papers presented at the thirtieth *Incontro di Grammatica Generativa* which took place in Venice and Treviso on February 26th - 28th 2004 and was organized by the Dipartimento di Scienze del Linguaggio of the Università Ca' Foscari Venezia.

The annual *Incontro di Grammatica Generativa* started in 1975 as a gathering occasion for the small group of Italian generativists at the time and has developed in the course of the last thirty years into a conference of international importance. As students and disciples of some of these formal linguistics pioneers in Italy, we feel proud of having organized the meeting and of editing this collection.

The papers published here can be grouped into five major research topics: *first language acquisition* (Foppolo and Guasti, Franchi, Merlo and Stevenson), *the structure of the lower portion of the clause*, namely the VP shell and the Mittelfeld (Belletti, Damonte, Laenzlinger and Soare, Schweikert), *the structure of the left periphery* (Garzonio and Gracci, Lipták, Paoli, Rizzi), *the internal structure of the DP* (Cabredo Hofherr and Dobrovie-Sorin, Grohmann and Panagiotidis), *the syntax and semantics of relative clauses* (Aboh, Del Gobbo, Frascarelli and Puglielli, Sleeman).

First language acquisition

Foppolo and Guasti's contribution deals with acquisition processes related to children's ability to derive scalar implicatures, and suggests that the observed difficulties for children to derive such inferences depend on different factors such as the tasks used, the age of testing and the different scalar items. Their proposal predicts that two separate steps are involved in the lexicalization of the different scalar items and that different scales are lexicalized at different stages in the course of acquisition.

Franchi investigates the role of the input in language acquisition by comparing Italian data from a non-signing deaf subject to the data produced by normal children at the 1st stage of acquisition. The data seems to suggest that deafness can impair the linguistic information present in the input so that no triggers are available to activate the process of language acquisition.

Merlo and Stevenson approach the acquisition of argument structure and

selectional properties of lexical elements. Based on the behaviour of three verb classes – manner of motion, change of state, creation/transformation – they show how frequency facts can be used in order to properly classify a given verb. Theoretical considerations about the status of frequencies in linguistic theory and statistical experiments complete the picture.

The lower portion of the clause

Belletti's study investigates different answering strategies to a question concerning the subject of the clause. Belletti highlights three different strategies for Italian, French and English also using L2 acquisition experimental data. For null subject languages, as standard Italian, the inversion strategy makes use of the dedicated focus position in the VP periphery, with a silent preverbal subject. In languages where the null subject parameter is set negatively, two possibilities are attested: French displays a non-null subject compatible "inversion" structure, exploiting the informational content of the VP periphery, while English adopts a (DP internal) focus-in-situ strategy.

Damonte is concerned with the difference in interpretation between the variants of alternating predicates such as the English verb *load*. He argues that a structure like *I loaded the sand on the truck* corresponds to the basic structure, while the alternative structure *I loaded the truck with sand* involves additional syntactic processes. This approach predicts that if a language only has one variant, the basic one is present. This appears to be confirmed by cross-linguistic data.

Laenzlinger and Soare propose a computational system based on multiple phases in order to account for the positioning of the arguments and of the verb in the Mittelfeld. The hypothesis is that all arguments must leave the VP-shell in order to have their A-features (Case, phi-features) as well as their IS-features (topic, focus, etc.) checked/matched. It ensues that scrambling is applicable not only to OV languages like German or Japanese but also to VO languages like English and Romance languages. Variations in the information structure in Romance are responsible for the different configurations found in these languages, especially with respect to adverb intervention.

Schweikert investigates the possibilities of extending Cinque's hierarchy of adverbs to prepositional modifiers. Given the fact that the order between two PPs is less rigid than the one between adverb classes, he applies certain syntactic tests to detect base orders using statistical methods to verify the stability of the results. These reveal not only a universal order of thematic roles but also show that deviation from the base order is dependent on the distance between elements in the hierarchy.

The left periphery of the clause

Garzonio and Gracci analyse the process of acquistition of Italian interrogative strategies by Tamil speakers and observe that they generally front *wh*-words and correctly move the inflected verb past the direct object. From the fact that a topicalized DP often precedes the *wh*-item, i.e., that second language learners have access to the internal CP layering in their syntactic representations, the authors conclude that at least some principles of UG must be active during second language acquisition.

By considering Hungarian word order patterns, Lipták claims that *wh*-phrases in exclamatives target two distinct left peripheral positions, FocusP and a higher quantificational projection, labelled *many*P; both projections are associated with a scalar exclamative operator, though they have different selectional properties. The author also claims that exclamatives and interrogatives differ in the type of focus they assign to *wh*-expressions: interrogatives have a contrastive focus, while exclamatives have an evaluative scalar one, but only the former is unambiguously linked to FocusP.

Drawing on some Northern Italian varieties, early Romance texts and French acquired as a first language, Paoli points out that the category Comp can perform a variety of tasks; in the data presented, complementizers can function as subordinating elements, mood markers, topic markers and morphemes specified for [φ] features; all of these properties are shared by the inflectional domain, which shows that the information encoded in the IP field can be duplicated, albeit in a reduced way, in the higher CP field.

Rizzi considers the semantic properties linked to the subject position and observes that subjects share with topics the fact that they are selected as the point of departure in the description of the event. The identification of a scope-discourse interpretive element connected to the subject position makes it possible to formulate a Subject Criterion, satisfied by moving an argument to subject position. Under the condition of Criterial Freezing, this approach derives the unmovability of subjects in certain environments. Finally, the author observes that Italian allows indefinite, and also non-specific indefinite topics, as long as they satisfy a condition of connection to the discourse background.

The internal structure of the DP

Cabredo Hofherr and Dobrovie-Sorin analyse the incompatibility of prenominal Saxon genitives (and of other genitive constructions cross-linguistically) with a lexical determiner. They argue against the assumption of a null determiner developing a Bare Phrase Structure analysis that dispenses with it. They also argue that prenominal and postnominal Saxon genitives are to receive distinct semantic and syntactic analyses despite the

common morphological marking on the grounds of substantial differences between the two constructions.

Grohmann and Panagiotidis focus on demonstrative constructions in Modern Greek applying the Anti-Locality Hypothesis (Grohmann 2003) to the nominal domain. The class of demonstrative elements is understood to include both an overt demonstrative pronoun and a covert, phonetically empty demonstrative operator. Furthermore, the definite article which must co-occur with the demonstrative is analysed as a grammatical formative doubling it.

The syntax and semantics of relative clauses

Aboh's contribution provides evidence from Gungbe in favour of the analysis of relative clauses according to which D° selects a clause as its complement and the relative noun raises to [spec,CP]; adopting this view, the author derives the sequence Noun-[relative clause]-Det-Num attested in Gungbe relative clauses. He also accounts for the semantic contrast between relative clauses and factive constructions by suggesting that the latter do not project the outer D-system typical of relatives and should therefore be analyzed as simple CP-clauses where [spec,CP] hosts an event DP that has moved out of the embedded IP. This analysis of factive clauses is extended to Germanic and Romance, where the factive reading is achieved by inserting an expletive factive DP in [spec,CP].

Del Gobbo's contribution studies the distinction between restrictive and "descriptive" relative clauses in Mandarin Chinese and concludes that "descriptive" relative clauses cannot be compared with appositive relative clauses in languages like English and Italian. She proposes that these constructions should be analysed as restrictive modifiers of the i-level/generic type, internal to the NP-layer. Moreover, she proposes that Chinese relative clauses modifying proper nouns and pronouns should be analysed like appositive adjectives, namely as predicates of individuals (type <e,t>), and different from appositive relative clauses, which are propositions (type t).

Frascarelli and Puglielli discuss the morphosyntactic differences that characterize restrictive and appositive relative clauses in two Cushitic languages (Somali and Afar) and argue that they cannot be analysed within a uniform approach. The authors maintain a promotion analysis for restrictive relative clauses and analyse appositive relative clauses in terms of assertive sentences conjoined to their antecedents. The antecedent merges as a fully referential DP and the appositive clause, a CP, is conjoined with it through an "asymmetric conjunctive structure". The antecedent DP is then co-indexed with the head of the appositive clause, which, cross-linguistically, can be either a deleted copy or a relative pronoun occupying SpecCP.

Finally, Sleeman argues in her contribution that DP-internal subordinate clauses such as French *à*+infinitive structures occurring with a superlative antecedent are complements of the prepositional complementizer *à,* which is located in Fin°. Her claim raises interesting questions concerning extraction from the infinitival constituent and LF-fronting of the superlative element.

The following papers were also presented at the thirtieth *Incontro di Grammatica Generativa* but unfortunately not submitted for publication:

Josef Bayer and Markus Bader "P as case"
Antonietta Bisetto, Sergio Scalise and Emiliano Guevara "Head selection in compounding and derivation"
Luigi Burzio "Paradigmatic and syntagmatic relations in Italian verbal inflection"
Katja Francesca Cantone "The role of complementizer in phrase structure: evidence from code-switching in support of the *Lexical Hypothesis"*
Carlo Cecchetto, Carlo Geraci and Alessandro Zucchi "Strategies of relativization in LIS"
Raffaella Folli and Heidi Harley "On the nature of little v: causation, obligation and argument structure"
Alexander Grosu and Manfred Krifka "Argumental relative constructions with a postcopular gap"
Maria Lluïsa Hernanz "On the left periphery in Spanish: emphatic affirmative sentences with *bien*"
Miriam Lemle and Aniela Improta França "Parametrizing verbal constructions: a Distributed Morphology account of Portuguese, Italian and English"
Jaume Mateu and Maria Teresa Espinal "Argument structure and compositionality in idiomatic constructions"
Ian Roberts "Bare Head Movement"
Halldór Sigurđsson "The syntax of speech features"
Øystein A. Vangsnes "Rolling up the Scandinavian noun phrase"

We think this volume bears witness that the Venice 2004 Incontro was, as usual, a successful occasion that brought together linguists working within the framework of generative grammar, thereby resulting in a fruitful exchange of ideas and opinions on recent theoretical developments. We sincerely thank all the participants and contributors for having made this possible.

The Editors

Children's comprehension of sentences involving scalar items

Francesca Foppolo and Maria Teresa Guasti
University of Milano-Bicocca

1. Introduction

In these last years, there's been a lot of investigations concerning children's ability to derive pragmatic inferences, more specifically scalar implicatures.[1] One generalization that emerges from these studies is that children are not as prone as adults to derive such inferences. A second generalization is that children behave differently depending on the scalar items (Chierchia et al. 2001; Noveck 2001; Papafragou and Musolino (P&M) 2003; Doitchinov 2003). The aim of this paper is to evaluate the factors that may be responsible for these difficulties and to offer a proposal that accounts for both the difference between children and adults' behaviour and for children's different sensitivity with respect to different scalar items. To this purpose, data from comprehension and production data of utterances including scalar items will be discussed.

As a first step toward an understanding of the acquisition problems, we'll first provide the reader with a basic definition of scalar implicatures (SI) in the framework of a neo-gricean approach, starting from an example involving "or". Consider a trivial conversation in which the speaker utters:

(1) Angela invited Sue or Lyn at the party

Her interlocutor would normally infer (1'):

[1] Although the article has been written together, Francesca Foppolo is principally responsible for sections 1, 2.1, 3.3 and 4, and Maria Teresa Guasti for sections 2.2, 2.3, 2.4, 3.1, and 3.2.

(1') Angela invited Sue or Lyn at the party, not both

The added proposition "not both" in (1') is not part of the propositional content of the speaker's utterance, but it's an inference that the hearer derives from the fact that the speaker used "or" in her utterance[2]. When speaking, in fact, we choose to use certain items instead of others, and our choice is interpreted by our interlocutors accordingly. In particular, we try to make our contribution as informative as it's required, as suggested by the Maxim of Quantity, and to conform to the Cooperation Principle and the other conversational maxims that rule our talk exchanges (Grice 1989). Some elements, like "or" in our examples, belong to a scale that is ordered on the basis of informational strength, i.e., <or, and> where "or" is the less informative element of the scale. When the speakers chooses the weaker (less informative) element in the scale then she wants to convey that she doesn't have sufficient evidence to use the stronger (most informative) element, or that she knows that it does not apply, otherwise she would have used it. Thus, by hearing "A or B", the hearer will infer that NOT (A and B), and this inference is called "Scalar Implicature".

In various experiments, that we will review in the next session, it has been shown that children, unlike adults, may accept "A or B" and more generally a statement including a weak scalar item (like "some" or "might") in contexts in which the strongest element ("and", "all", "must") would be more appropriate, and in this respect they seem not to derive SIs. In line with current understanding of this problem, we will show that this over acceptance of underinformative utterances does not stem from children's poor pragmatic competence. At least three factors, in fact, seem to influence their low performance: the age of the subjects, the task employed and the type of scale tested, as we will discuss in the rest of this paper.

2. Comprehension: how children interpret sentences involving scalar items

Although the study of children's derivation of SIs is recent, back in the eighties several experiments have investigated, under a different perspective, children's understanding of the connective "or" and the quantifier "some".

[2] This inference can be suspended in certain contexts, as exemplified by (i.a) and (i.b), or cancelled, as exemplified by (i.c):

(i) a. If Angela invites Sue or Lyn at the party, I'll stay at home
b. If Angela invites Sue or Lyn or both at the party, I'll stay at home
c. Angela invited Sue or Lyn at the party, actually both

2.1. *Comprehension of "or"*

In Paris' study (1973) on children's understanding of the connectives "or" and "and", 8 year old children and adults were tested on sentences like "A or B" and "A and B" in four situations: when only one between A and B was true, when both were true and when both were false. While all the subjects were competent with sentences containing "and" (lowest rate percentage of children correct response = 87,5%) a difference in the performance between children and adults was evident on sentences containing "or". In particular, children accepted the sentence "A or B" when both A and B were true in 92,5% of the times, while adults did so 67,5% of the times. Interestingly, Paris also found that children had a general difficulty in understanding "or" in the situation in which only one of the two disjuncts were true (percentage of yes responses was 50% and 37,3% in these cases). Similarly, Brain and Rumain (1981) found that children interpreted "or" inclusively (A or B or both) in situations in which adults favoured the "exclusive" interpretation of the disjunction (A or B but not both). From the data presented in their study, there seems to be a developmental shift from "inclusive" to "exclusive" interpretation.

In a more recent account, Chierchia et al. (2004) tested 5 year old children on the interpretation of sentences containing "or" and "and" using the context of a bet: a puppet uttered sentences like "I bet that Batman will take a cake or/and an apple" and the child was asked to reward the puppet if he wins the bet in contexts in which Batman took only a cake, only an apple, or both. Their results suggest that children at 5 differentiate between "or" and "and" when Batman took only one of the two objects, correctly rewarding the puppet 78% of the times in case of the sentences with "or" and rewarding him 16% of the times in case of the sentence with "and". However, they also found a high rate of acceptance of the disjunctive sentence (98%) in the situation in which both disjuncts were true.

Gualmini et al. (2001) conducted another study on "or" using the technique of the Truth Value Judgment Task (TVJT) on a group of fifteen children ranging in age from 3;5 to 6;2 (mean age: 5;2). They found that seven out of the fifteen children behaved like adults, rejecting the target sentences with "or" in a situation in which a description with "and" would have been more appropriate, thus accessing the "exclusive" interpretation, while the rest of the children were stuck on the inclusive interpretation. Thus, across several experiments, children have been found to overaccept statements with the weak scalar item "or" in situations in which "and" would have been more appropriate.

2.2. *Comprehension of "some"*

A situation similar to the one observed for "or" is attested with the quantifier "some", an item that normally conveys the implicature "some but not all". Back in 1980, Smith tested preschoolers on the interpretation of quantifiers like "some" and "all" and concluded that, while children correctly interpreted these quantifiers in felicitous contexts, most of them treated "some" as compatible with "all", giving a "yes" answer to questions like "Do some elephants have trunks?" In a partial replication of this work, Noveck (2001) found that 7 year old French speaking children were more prone than adults in accepting underinformative statements like "some giraffes have a long neck" (children's mean acceptance rate = 89%, adults' = 41%). As shown by Guasti et al. (2004, in press), this result is due to the particular task (evaluate statements out of context) used by Noveck or to the difficulties in understanding the experimental question (which was simply "do you agree or not?"). Guasti and colleagues tested 7 year olds children in two different experiments. One employed the methods and material used by Noveck (translated and adapted to Italian) and the other method used the Truth Value Judgement task (TVJT) instead (Crain and Thornton 1998). In this second task, the experimenter acted out a story in front of the subjects using props and toys and at the end of the story a puppet had to describe what has happened in the story. The subject was then asked to evaluate the puppet's statement, saying if it was a good or a bad description of the story. In a typical trial, the puppet described a story in which 5 out of 5 Smurfs went for a trip by boat by saying that "some of the Smurfs went on a boat", which is a true but inappropriate description of the story. The task used turned out to be a crucial factor in the subjects' performance, as we'll briefly discuss.

Guasti and colleagues found a significant difference between children and adults in the mean acceptance rate of underinformative statements with "some" when using the task used by Noveck (87% for children and 50% for adults, $F1\ (1, 33) = 8{,}73$, $p < .005$). This difference disappeared when the TVJT was used: the mean rejection rate increased dramatically in both groups (75% for children and 83% for adults), and no age difference is found anymore ($F(1,25)=,31$, n.s.). Thus, the kind of task employed may elicit different types of response and this suggests that children's pragmatic ability in deriving implicatures is present, but is not always put to use, because the experimental goals are unclear. However, this cannot be the only factor that explains differences between children and adults. Using the same methods employed by Guasti et al. (2004, in press), Papafragou and Musolino (2003) found that 5 year old Greek speaking children reject the underinformative

statements only 12% of the time, while adults did so 92,5% of the time.[3] Taking together, the results obtained by Guasti et al. and Papafragou and Musolino suggest that there is a developmental effect to be considered, in the derivation of pragmatic inferences: 5 year old children are less prone than 7 year olds to derive SIs.

2.3. *Comprehension of other scales*

So far, we have considered two scales, one that includes <some, all> and the other that includes <or, and> and we have seen that the kind of task, as well as age, matter. When we consider other scales, the picture becomes more complex. On the one hand, it has been found that some scales are particularly difficult. One case is the modal scale, <might, must>. Doitchnov, (2003) found that, while at age 6 the scale <some, all> was not problematic for children and rejection rate of underinformative statements including "some" reached adult level (see also Guasti et al. in preparation), still at age 8 the scale involving modals was tricky, with rejections of underinformative statements with "may" being 65%, while it reached 100% with adults in the same task (see also Noveck 2001 who found low rejection rates even with 9 year old children). A second case is the perfectivity/aspectual scale, <start, finish> (Papafragou and Musolino 2003 and Papafragou 2004). In this case, 5 year old children rejected underinformative descriptions like "the girl started making the puzzle" when the puzzle was actually completed only 10% of the times.

On the other hand, there are some scales that elicit a higher percentage of rejections by children. This is the case of the degree scale, like <half, whole>, which elicited 67% of rejections from 5 year olds (Papafragou 2004), and the scale involving numerals (Papafragou and Musolino 2003), where underinformative statements are rejected 65% of the time (see note 3 in this page).

2.4. *Taking stock*

From the discussion presented in the previous sections, some general conclusions seem to emerge:

[3] In their paper, they report a variant of this experiment in which the experimental session was preceded by a training session (and other manipulations were made, but not considered, see Guasti et al., in preparation). Children's performance after the training improved significantly, but, given that the effective role of training is not clear, we'll always refer to the results of the basic version of the experiment reported in the paper.

A. different tasks (SET vs. TVJT) seem to evoke different responses, in children but also in adults[4];
B at the same age, children behave differently with respect to different scalar items, thus suggesting that different scales are accessible at different stages in the development (discrete<logical<epistemic)

It seems to be the case that children at age 7 possess the ability to derive pragmatic inferences related to some scalar items (i.e., "some"), but are prevented from displaying their pragmatic competence by some features of the experimental design. This explanation is compatible with the hypotheses that a limitation, and not a delay, in the Pragmatic components is responsible of the failure in non-naturalistic tasks. Evidence for this hypothesis is the fact that children's performance improves when they are tested with the TVJT, which constitutes a better attempt to reproduce an ordinary conversational exchange. Crucial to the dynamics of normal discourses is the fact that speakers (in this case puppet and experimenter) and hearers (subjects) share a common conversational background, which they update on the basis of what happens in the context, i.e. the events occurring during the story. This dynamics was prevented in the task in which subjects were asked to say whether they agreed or not with some statements, where no context was given and thus no sharing of the conversational background was possible. To be sure that the sharing of the conversation background is one of the factors that may affect children performance, we turned to a different experiment, this time investigating production and not comprehension. Our aim was to see whether children at age 5 display some difficulties in the production of the items involved in the scale <some, all>, as they seem to do in a comprehension task, when the conversational background is shared by the participants to the experiment.

3. Eliciting "some" and "all"

In this experiment we used an Elicitation Task. Our aim was twofold: we wanted to test if children could produce "some" and "all" in the appropriate context, and we aim at recording which of the variants to the standard "qualche" (lit. = some) they used preferably. In Italian, as in other languages, there are many different ways to express the existential quantifier. It is thus interesting to test which variant they preferably produce when they are asked to describe a situation in which some, but not all, the characters in the story were involved in some sort of action.

[4] Cf. Guasti et al. (2004) for a discussion on adult responses.

3.1. *Participants*

A group of 23 children ranging in age from 3;7 to 5;8 (M=4;6, SD=.7) participated in the study.

Children were recruited from different nursery schools in the Milan Area (Cassina de' Pecchi, Milano 2).[5]

3.2. *Materials and procedure*

Children were tested individually in a quiet room in their nursery school, after being familiarized with the experimenters and the puppet. They were presented with three stories acted out in front of them using props and toys. At the end of each story a puppet asked the children to explain what has happened. To carry out the test, two experimenters were involved, one manipulating a puppet and the other narrating the stories. Let's consider one of the stories used as an example. The child was shown all the props and toys to be used in the story, which were for example ten Smurfs, a teacher and several candies. Since the Smurfs did very well at school, the teacher decided to give a candy to all of them. After that, the puppet asked the child to describe what has happened. The target sentence for this story was a sentence of the kind: "The teacher gave a candy to all the Smurfs" or "All Smurfs got a candy". At this point, the story went on with the teacher that wanted to give something else to the Smurfs. This time she had marbles and hats. Since she had not enough of these for all Smurfs, she gives a marble to some of the them and a hat to the rest. The target sentence we wanted to elicit for this part of the story was of the kind: "Some Smurfs got a marble and some Smurfs got a hat" or "The teacher gave a marble to some of the Smurfs and a hat to the others".

Each test session was recorded and children's production data were transcribed.

3.3. *Results and discussion*

None of the subjects tested had any difficulty in producing sentences containing the universal quantifier "all". We obtained 71 such utterances overall. In Italian, the universal quantifier can be expressed by different items: *tutti* (all), *ogni* (every) and *ciascuno* (each). With the exception of a single child who produced three utterances using *ogni* (every) and *ciascuno*

[5] We would like to thank the children and the teachers of the schools. Part of the data used in section were collected with the help of Andrea Gualmini. We wish to thank him for this.

(each), in all the other utterances children used the quantifier *tutti* (all) to describe a story in which each of the participants got/did the same thing. Interestingly, to convey a distributional reading of "all", one of the children used the following sentence in which the expression *uno per uno* (lit. one by one) was used:

(2) Tutti i puffi hanno preso un secchiello, uno per uno
[All the Smurfs have taken a bucket, one by one]

Sometimes children did not use the universal quantifier at the first request, but used *tanti* (many) or the definite plural article instead. Only when they were asked again they used the quantifier *tutti*. There were 4 instances of *tanti* and 3 of the definite plural article (none of these was in fact inappropriate in the situation). Some examples are reported below:

(3) Ci sono tanti signori che hanno preso tanti biscotti
[There are many gentlemen that have taken many biscuits]

(4) Abbiamo dato dei secchi ai puffi (a tutti i puffi)
[(We) gave "some" buckets to-the-PLU Smurfs (to all the Smurfs)]

Turning now to the elicitation of "some", we observe that children produced 53 utterances containing this quantifier overall. Different lexical items were produced: *alcuni*, *certi,* and the partitive plural article *dei* (all can be considered instances of the English "some"), *qualcuno* (someone), *altri* (others), *un po'* (a bunch), *metà* (half). Sometimes children produced items that do not exist in the adult language. For example, *qualcuno* (someone) is morphologically singular, but can refer to a plurality. Children invented a morphological plural form of this item, substituting the singular morpheme *-o* with the standard plural morpheme *-i*, thus producing *qualcuni*. Similarly, *uno* (one) is the indefinite singular article (generic). Children invented *uni*, which is meant to be its plural counterpart.

In addition, some stories that were designed to elicit *alcuni* in fact elicited *tutti*. But we cannot count this as a real mistake, as it's clear if we consider an example in which this substitution was made:

some children uttered Φ: Tutti hanno il cappello e la biglia
[All have the hat and the marble]

instead of the expected Ψ: Alcuni hanno il cappello e alcuni la biglia
[Some of them have the hat and some of them the marble]

Three of the twenty-three children made this mistake two times each. The sentence in Φ was used to convey the meaning that all of the Smurfs had either the hat or the marble. This sentence would have been appropriate if the connective "or" rather than "and" was used (see Bolster et al. 1993).

Summarising our results, we can conclude that, although we found some inappropriate uses of the quantifiers, children were pretty good at producing quantified sentences. In particular, they correctly used *all* when the action described in the story involved all the relevant characters and used *some* when the action involved only a subset of the characters.

4. Discussion and conclusions

In the first part of this paper, we reported some generalizations that seem to emerge from the acquisition literature and our own work on pragmatic inferences. Children (at least up to the age of 5) have some difficulties in deriving pragmatic inferences and respond differently to different scalar items. It has been shown, however, that children's failures with scalar implicatures may depend on the particular task used, which can mask their effective competence (Papafragou and Musolino 2003; Guasti et al. 2004; Foppolo and Guasti, 2004). Moreover, children's performance depends on the age of testing and on the particular scalar item used, as discussed above.

In section 3, we reported an elicited production study in which children were prompted to produce sentences containing the quantifiers involved in the logical scale <some, all> to test if they were able to produce these quantifiers and use them when the conversational background is shared among the participants in the experiment. In fact, most of the experiments conducted so far used the TVJT to attest children's competence in comprehending these quantifiers, but no experiment was carried out to verify if children produced them, and which versions of these quantifiers they preferred among the ones available in their language. Italian was a good test for an elicitation study, given that this language provides a number of different alternatives that could be used by children. The results obtained in our study confirm the generalization emerged in previous studies, showing that children are competent with "some" and "all". If we consider, however, the performance of 5 year old children on the scalar implicatures related to *some*, we find a discrepancy between their ability to use the quantifiers involved in the scale and their readiness to derive the inferences related to

this scale. Knowing the items involved in the scale seems not to be a sufficient condition to derive the SI associated to this scale. An attempt to explain this failure as the result of a limitation of the processing resources in children is incompatible with the fact that children at a certain age are able to derive SI for some scales, but not for others.[6] For example, at the age of 5 children are quite good at computing SIs related to the number scale or the scale <half, whole>, but not the ones related to logical scales like <some, all>; at 6 they are able to derive SI related to these discrete scales and also to the <some, all> scale, but not the ones related to the epistemic terms or to perfectivity/aspectual scale, an ability that will emerge even later. On the basis of these data, we would like to suggest that, simply showing that children at a certain age know the items in the scale and how to use them in the appropriate situation is not sufficient to conclude that they can derive the implicatures; one needs also to show that they are able to connect these items to form a scale, so that the use of the weaker item in the scale prompts the activation, and the consequent suspension, of the stronger element. At this point, we can formulate an hypotheses, which we'll call the Lexical Hypotheses, that makes two predictions:

i) two separate steps are involved in the lexicalization of the scalar items, and both must be completed in order to derive the scalar implicatures. One preliminary step is the acquisition of each lexical scalar item separately, with its restrictions on use in felicitous contexts. This step is the first to be acquired. Subsequently, and at a further separate step, the scale itself needs to be lexicalized, i.e. the scalar items should be linked to form an ordered scale (a sort of paradigm). This step can be acquired as a separated step, and can take a while to be completed after the first step is acquired, depending on the kind of scale.
ii) different scales may be lexicalized at different stages in development.

If the Lexical Hypothesis is correct, then the data presented in our elicitation task, showing that children are competent with "some" in felicitous contexts, and the data reported in previous studies with the TVJT, showing that children at 5 still interpret "some" as "some even all", can be explained appealing to the steps in the process of item/scale lexicalization. At the age of 5, in fact, children have completed step 1 in the acquisition of the lexical items "some" and "all" separately, but they have not connected these items in a scale, they haven't built the whole "paradigm" related to these items yet.

[6] The fact that children at a certain age and using the same task derive implicatures more with some scalar items than others is also evidence against a Pragmatic Delay Hypothesis.

Thus, the second step, i.e. the lexicalization of the scale <some, all> is not complete, as predicted by (i). At the same age, however, children can derive the SI related to other scales, as reported by Papafragou and Musolino (2003) and Foppolo and Guasti (2004) and this is compatible with the prediction expressed in (ii), namely that different scales may be lexicalized at different stages in development.

References

Boster, Carole and Stephen Crain. 1993. "On Children's Understanding of *Every* and *Or*". *Proceedings of Early Cognition and Transition to Language*, University of Texas at Austin.

Brain, M. and B. Roumain. 1981. "Children's comprehension of "or": Evidence for a sequence of competences". *Journal of Experimental Child Psychology* 31:46-70.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, Andrea Gualmini and Luisa Meroni. 2001. "The Acquisition of Disjunction: Evidence for a Grammatical View of Scalar Implicatures". In A. H.-J. Do, L. Dominguez and A. Johansen (eds.) *Proceedings of the 25th Boston University Conference on Language Development.* 157-168. Sommerville MA, Cascadilla Press.

Chierchia, Gennaro, Maria Teresa Guasti, Andrea Gualmini, Luisa Meroni, Stephen Crain and Francesca Foppolo. 2004. "Semantic and pragmatic competence in children and adults' comprehension of *or*". In I. Noveck and D. Sperber (eds.) *Experimental Pragmatics*.

Crain, Stephen and Mark Steedman. 1985. "On not being led up the garden path: the use of context by the psychological syntax processor". In D. R. Dowty, L. Karttunen and A. Zwicky (eds.) *Natural Language Parsing: Psychological, Computational and Theoretical Perspective*. Cambridge MA, Cambridge University Press.

Crain, Stephen and Rosalind Thornton. 1998. *Investigations in Universal Grammar. A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge MA, MIT Press.

Foppolo, Francesca, Maria Teresa Guasti and Gennaro Chierchia. 2004. "Pragmatic inferences in children's comprehension of scalar items". Talk presented at *Second Lisbon Meeting on Language Acquisition*, Lisboa, 1-4 June 2004.

Foppolo, Francesca and Maria Teresa Guasti. 2004. "Scalar Implicatures in children: failures or skillful strategies?". Poster presented at *NELS 35*, Storrs, 22-24 October 2004.

Grice, Paul. 1975. “Logic and Conversation”. In P. Cole and J. Morgan (eds.) *Syntax and Semantics 3: Speech Acts*, New York, Academic Press. Also in Paul Grice. 1989. *Studies in the Way of Words*. Cambridge MA, Harvard University Press.

Guasti, Maria Teresa. 2002. *Language Acquisition. The Growth of Grammar*. Cambridge MA, MIT Press.

Guasti, Maria Teresa, Gennaro Chierchia, Stephen Crain, Francesca Foppolo, Andrea Gualmini and Luisa Meroni. 2004. “Why children and adults sometimes (but not always) compute implicatures”. To appear in *Language and Cognitive Processes,* 2004.

Hirschberg, Julia. 1985. *A Theory of Scalar Implicatures*. Ph.D. dissertation, University of Pennsylvania.

Horn, Laurence. 1972. *On the Semantic Properties of Logical Operators in English*. Ph.D. dissertation, UCLA.

Levinson, Stephen. 2000. *Presumptive meaning*. Cambridge MA, MIT Press.

Musolino, Julien. 2004. “The Semantics and Acquisition of Number Words: Integrating Linguistic and Developmental Perspectives”. *Cognition* 93-1:1-41.

Noveck, Ira. 2001. “When children are more logical than adults: Experimental investigations of scalar implicatures”. *Cognition* 78:165-188

Papafragou, Anna and Julia Musolino. 2003. “Scalar Implicatures at the Semantic-Pragmatics Interface”. *Cognition* 80:253-282.

Smith, Carol L. 1980. “Quantifier and question answering in young children”. *Journal of Experimental Child Psychology* 30:191-205.

Sperber, Dan and Deirdre Wilson. 1986/1995. *Relevance: communication and cognition*. Oxford, Basil Blackwell.

Omission versus random selection of *essere*. Comparing a deaf subject Italian data elicited through logogenia with acquisition data

Elisa Franchi
University of Venice

1. Introduction

The analysis of the role of the linguistic information available in the input for triggering the process of language acquisition is a central issue within Logogenia[1], which empirically investigates this aspect of the process by analysing the comprehension and production of the oral language[2] by non signing profoundly deaf people.

The role of the input in language acquisition is currently under debate: on the one hand, as suggested by Lightfoot (1999), a given structure - although present in the input - may not become a trigger for syntactic development if its frequency in the linguistic environment does not reach a specific threshold. On the other hand, studies on home sign systems presented in Goldin Meadow (2003) suggest the possible emergence of linguistic behaviour even in the absence of linguistic input.

In this work copular Italian data from a non signing prelingually deaf subject will be compared to data on the production of the Copula *essere* in normal developing children. Both corpora include non target-like productions and were generated in crucially different situations as for the quality and the quantity of the input received. In normal acquisition conditions, the input is generally considered "poor" (in the sense of

[1] See Radelli (2000) and Radelli (1999).

[2] In this work we will always refer to the oral language as opposed to the sign language, abstracting away from the actual modality, written or oral, of data production. In fact, all the deaf person data considered here were elicited in the written modality.

Chomsky 1980's arguments from the poverty of the stimulus) but it is obviously sufficient for triggering the process of language acquisition. On the other hand, in cases of deafness the available linguistic input is limited in quantity and altered in quality and might not allow the acquisition of the oral language.

In order to check whether both these conditions can trigger the process of language acquisition, the Child Italian corpus and the deaf subject corpus will be analysed and compared in search for regular and predictable patterns as well as for the operativity of UG constraints.

Regular and predictable patterns have indeed emerged from child language data and UG constraints appear to be fully operative from the earlier stages of language development.[3]

The very same tools adopted in the literature for detecting UG constrained patterns will be applied here to the acquisition corpus and to the deaf subject's Italian corpus. If similar patterns will emerge from the two corpora, both kinds of input will be considered sufficient to trigger the process of language acquisition. On the other hand, if regular and predictable patterns will only emerge from the acquisition data but not from the deaf subject Italian data, it will be possible to conclude that the input available in cases of deafness might not always qualify as an appropriate trigger of the process of language acquisition.

1.1. *Data collection*

The deaf subject's Italian data were compiled from a prelingually deaf adult, who has never had access to the sign language.[4] A corpus of written production was elicited through Logogenia, which aims at studying the deaf people production in the oral language from a Generative Grammar perspective.[5] Both production and comprehension data were included in the analysis.

[3] See Rizzi (2004).

[4] A non-signer allows one to compile first language Italian data. Italian is the only language to which Gabriele has ever been exposed to. Since the subject was an adult, it was possible to elicit data in the written form. The use of the written modality is indeed crucial for bypassing comprehension problems due to difficulties in lip reading and to difficulties in understanding the oral language both by the deaf subject and by the experimenter.

[5] Logogenia has identified a method especially designed to trigger deaf children acquisition of the spoken (non signed) language by exposing them to a specific and selected syntactic input presented in the written form. As for linguistic analysis, Logogenia's data offer an extremely fine grained and precise picture of the actual syntactic competence of the deaf person they are elicited from.

The child language data were collected from three monolingual normal hearing children, their ages ranging from 1 year and 7 months to 3 years and 3 months.[6] Spontaneous production data were considered.

2. The general picture

In Gabriele's corpus, perfectly grammatical sentences such as (1) can co-occur with syntactically unpredictable word sequences such as (2).[7]

(1) Io sono più vecchio.
I $BE_{1st\ sing}$ more $old_{masc.\ sing}$

(2) C'È SONO QUESTO È NOME SONO IN BARBIERE.
There is $BE_{1st\ sing}$ this is name $BE_{1st\ sing}$ in barber's shop

Sentences like (1) and sequences like (2) are almost equally frequent in Gabriele's corpus.

A very similar pattern emerges when considering word order data. In Gabriele's corpus, both sequences that respect Italian word order constraints and sequences that violate them are detected.

In cases like those reported under (3), randomly selected functional elements can freely intervene in between words, in positions that are not syntactically legitimate.[8]

(3) a. Diana che vuole è tagliare il capelli è
Diana that $WANT_{3rd\ sing}$ $BE_{3rd\ sing}$ $CUT_{inf.}$ the_{sing} $Hair_{plur}$ $BE_{3rd.sing}$
corti sì o no?
short yes or no
b. ieri, io ho visto sono la partita
yesterday I $HAVE_{1st\ sing}$ $SEE_{past.part.}$ $BE_{1st\ sing}$ the match
c. no, non è la maglia sono marrone da Elena.
no not $BE_{3rd\ sing}$ the shirt $BE_{1st\ sing}$ brown from Elena

In other cases, as shown in (4), it seems impossible to detect the operativity of word order constraints in Gabriele's production.

[6] Data collected from the CHILDES database, November 2002.
[7] The use of capitals reproduces Gabriele's spontaneous use of it in his handwriting.
[8] Randomly selected functional elements are marked in the text by underlined style.

(4) a. sì, ma proprio è quello sono verrà anche
yes but really $BE_{3rd\ sing}$ that $BE_{1st\ sing}$ $COME_{fut.\ 3rd\ sing}$ also
sono verrò
$BE_{1st\ sing}$ $COME_{fut.\ 1st\ sing}$
b. ma sono abbastanza è correre sono città anche in treno.
but $BE_{1st\ sing}$ enough $BE_{3rd\ sing}$ $RUN_{inf.}$ $BE_{1st\ sing}$ city also in train
c. sono ferrovia per treno a parte XXX.[9]
$BE_{1st\ sing}$ rail station for train to $LEAVE_{3rd\ sing}$ XXX

The sequences reported under (5) indicate that syntactically legitimate sequences can only emerge in the data by means of arbitrary selection. Syntactically legitimate sequences must be isolated in larger units of Gabriele's production which cannot be syntactically analysed as a whole.[10]

(5) a. <u>È VICINO</u> SONO XXX
[it's close]
b. C'È SONO <u>QUESTO È NOME</u> SONO IN BARBIERE
[this is name]
c. <u>SI, SONO ARRIVA A XXX</u> CHE FERMARSI SONO TROPPO GENTE È MEGLIO COME IN CAMMINARE SONO XXX
[Yes, $BE_{Aux1st\ sing.}$ $ARRIVE_{3rd\ sing.\ pres.}$]
d. SÌ, COME SONO IN PAPÀ <u>HA DETTO</u> SONO SENTE PARLARE INSIEME.
[$HAVE_{Aux.\ 3rd\ sing.}$ $SAID_{Past.\ Part.}$]

Gabriele's data, when considered as a whole unit, seem then very unpredictable and contradictory.

This absence of predictable and regular patterns is not expected in normal conditions and in fact never emerges from acquisition data.

Strong regularities emerge from normal acquisition data, which appear to be predictable within the theory. Relevant studies have shown that children never violate word order constraints. Child language data can always be syntactically analysed, even when they are not consistent with the target adult language. In fact, children's early production indicates that they have correctly set all the relevant parameters of their language at the onset of the two word stage.[11] Many relevant studies on the syntax of child language thus

[9] Every reference to places and locations is cancelled and replaced with XXX.
[10] Syntactically legitimate sequences are marked in the text by underlined style.
[11] This observation has been formally defined by Wexler as Very Early Parameter Setting (VEPS). See, amongst others, Wexler (1998).

indicate that a Full (Syntactic) Competence emerges from the very earlier stages of language acquisition.[12]

On the other hand, if all features of Gabriele's data are taken into account, it seems possible to suggest, in his case, absence of syntactic competence. A Null (Syntactic) Competence Hypothesis is not unexpected within Logogenia. Cases of Null Syntactic Competence have in fact been empirically observed and theoretically predicted as a possible consequence of the limited access to relevant linguistic input caused by deafness.

3. Systematic omission vs. random selection of functional elements

In case of Null Syntactic Competence, absence of functional structure should clearly emerge from the data. A precise set of predictions on how to detect the presence or absence of functional structure in non target productions is independently formulated in Borer and Rohrbacher (2003).[13] According to the authors, the systematic omission of functional material in a given corpus should indicate the presence of functional structure, whereas the random over use of functional material should argue for the absence of functional structure. A given functional element - the copula *essere* - was then chosen and its production analysed in both the normal acquisition corpus and in Gabriele's corpus.

3.1. *Acquisition data*

The data taken into consideration for the present analysis were drawn from the CHILDES corpus.[14] Production data were compiled from three monolingual Italian children, their ages ranging from 1;7 to 3;3. 1587 utterances were coded that could be analysed as copular constructions. The most relevant non adult like phenomenon detected was omission of copula.[15]

3.1.1. *Omission of the copula*

Examples under (6) to (8) and the data in *Table 1* show that omission of copula is a relevant phenomenon in Italian Child Language, both considering all data and selecting only omissions in contexts with a realised subject (+subject contexts).

[12] See Guasti (2002) for a detailed survey of the relevant studies.
[13] Henceforth B&R.
[14] The Childes Database, November 2002, but see also MacWinney and Snow (1985).
[15] For a more detailed discussion of the data, see Franchi (in press_a).

Copula omission in +Subject contexts (SP) and in -Subject contexts (P):

(6) Martina (1; 08.02) (SP) quello pezzo
That$_{masc.\ sing.}$ Ø$_{BE}$ piece$_{masc.\ sing}$
(P) un pezzo
Null Subject Ø$_{BE}$ Det$_{masc.\ sing.}$ piece

(7) Raffaello (1; 11.25) (SP) Pallo butto
Paolo Ø$_{BE}$ ugly$_{masc.sing.}$
(P) glossa
Null Subject Ø$_{BE}$ big$_{fem.\ sing.}$

(8) Rosa (2; 10.14) (SP) ette bee
Demonstr$_{fem.plur.}$ Ø$_{BE}$ sheep
(P) zucchero!
Null Subject Ø$_{BE}$ sugar

Table 1: Number of contexts and omission rate in +Subject contexts (SP/SCP), -Subject contexts (P/CP) and total, per child and in total.

	N SP/SCP	% SP/SCP	n. P%CP	%P %CP	n.total	% total
Martina 1;7-2;7	69/197	35 %	70/169	41 %	139/367	38 %
Raffaello 1;7-2;11	30/133	27 %	92/218	42 %	122/331	37 %
Rosa 1;7-3;3	136/382	36 %	310/506	61 %	446/888	50 %
TOTAL	235/692	**34 %**	472/893	**53 %**	707/586	**45 %**

The analysis of the realised subject contexts, which were considered genuine omission data, indicates the existence of a 1st stage of acquisition in which the omission phenomenon is even more relevant (Table 2).

The child language copular data seem then to indicate that omission corresponds to a syntactic option available in child grammar.

Table 2: Age span and omission rate in +Subject contexts in the 1st and in the 2nd stage of acquisition.

	1st stage	1st stage	2nd stage	2nd stage
Martina	1;7-1;11	49 %	2;1-2;7	17 %
Raffaello	1;7-2;4	65 %	2;5-2;11	17 %
Rosa	1;7-2;5	81 %	2;3-3;3	26 %

3.1.2 *Patterns of omission: the WH contexts*

The possible influence on the omission pattern of declarative versus *wh*-interrogative contexts was taken into account.

419 *wh*-copular contexts were isolated in the acquisition corpus, finding only two cases of copula omission, with an omission rate of 0,48% (relevant data are presented in *Table 3*).

Table 3: Number of contexts and omission rate in *wh*-contexts, per child and in total.

Copular Wh	# contexts	# omissions	% omissions
Martina (1;7-2;7)	51	1	2.00 %
Raffaello (1;7-2;11)	78	0	0.00 %
Rosa (1;7-3;3)	290	1	0,34 %
TOTAL	**419**	**2**	**0.48 %**

The data so far analysed show that Italian children have the option of omitting the copula in declarative contexts but must produce a fully inflected copular form in very specific syntactic environments.[16]

3.1.3. *Agreement data*

In the Full Competence Hypothesis children are expected to always produce the correct copular form when they choose not to omit it. Consistently with this hypothesis, children in this study are shown to use the correct morphological form of the copula 99% of the time, as shown in the second and third column of *Table 4*.

[16] A similar pattern emerges from the analysis of Auxiliary data. The presence of this "WH constraint" leads one to imagine a Truncation strategy operative in Italian Child Grammar. The possibility of generating truncated structures would account for both the omission of functional verbs and its restrictions, crucially predicting obligatory presence of a realised copula (and auxiliary) in cases of WH preposing. See Franchi (in press_b) for a discussion of this proposal.

Table 4: Agreement data

	Errors / total	%	Errors / Plural	%
Martina	1 / 367	0.27 %	1 / 26	3.85 %
Raffaello	3 / 331	0.91 %	3 / 43	6.98 %
Rosa	13 / 888	1.46 %	13 / 101	12.87 %
TOTAL	**17 / 1586**	**1.07 %**	**17 / 170**	**10.00 %**

As the fourth and fifth column in *Table 4* show, the few non target forms detected are all related with the production of the 3rd person plural form of the copula (10% of errors detected), as the examples in (9) to (11) show.

(9) Martina (2; 03.22) s'è i gatti
there is the cats

(10) Raffaello (2; 11. 09) cos'è quelle?
what is those

(11) Rosa (2; 9. 04) tuo c'è e@p macchine intanto
(in) yours there is the cars meanwhile

A detailed analysis of the plural contexts in which non target forms emerge shows that the 3rd person plural form *sono* is only replaced by the 3rd person singular form *è*. This form mainly appears in post-verbal plural subject contexts (14 cases out of 17).

As shown in Guasti and Rizzi (2002), there appear to be free variation amongst languages in expressing the morphological agreement with postverbal subjects, as the examples in (12) indicate.

(12) Patterns of agreement with postverbal subject.
a. Three girls are in the garden.
a'. There are three girls in the garden
b. Trois filles sont arrivées.
Three girls are arrived
b'. Il est arrivé trois filles.
It is arrived three girls
c. Questo, i bambini lo fanno sempre
This the kids it do_{plur} always]
c'. Questo, lo fa sempre i bambini.[17]
This it do_{sing} always the kids

[17] Examples (12a) to (12c') are drawn from Guasti and Rizzi (2002).

d. Ci sono troppi soldi
there is too much money$_{plur}$
d'. C'è troppi soldi[18]
there is too much money$_{plur}$

In a cross-linguistic perspective, child language data seem then consistent with adult language data and - therefore - consistent with UG. The agreement "errors" found in copular constructions in acquisition are not errors at all, but instances of an option made available by UG and exploited both in child and in adult languages.

3.2. *The deaf subject Italian data*

The deaf subject's Italian data was analysed, looking for either systematic omission of the copula and syntactically constrained morphological errors or for random over-use of *essere* and random distribution of morphological errors.

During the experimental sessions, 262 utterances were produced, 145 of which contained forms of *essere*.

No omission of copula is detected in Gabriele's corpus. In his production a form of *essere* is always present when a sentence is to be understood as copular.

3.2.1. *Random selection of* essere

As the few examples reported in (13) and (14) show, the coding of the different uses of *essere* was a non trivial task.

(13) a. Io sono più vecchio.
I BE$_{1st\ sing}$ more old
b. No, però [lei] sono dottoressa
No, but <she> BE$_{1st\ sing}$ doctor$_{fem}$.
c. Ieri, io sono andato a XXX
Yesterday I BE$_{1st\ sing}$ gone to XXX

(14) a. La Juve è cade, sono sconfitta ha perso!
The Juventus BE$_{3rd\ sing}$ fall$_{3rd\ sing}$ BE$_{1st\ sing}$ beaten HAVE$_{3rd\ sing}$ lost
b. C'È SONO QUESTO È NOME SONO IN BARBIERE.
There is BE$_{1st\ sing}$ this is name BE$_{1st\ sing}$ in barber's shop

[18] Langhe dialect, Piedmont.

Not all instances of *essere* could be easily assigned a copular or auxiliary reading (as in (14b), for instance). In the whole of Gabriele's production, only 70 utterances containing forms of e*ssere* out of 145 (48%) were coded as copular structures and 14 as uses of auxiliary *essere* (10%).

The remaining 65 instances of *essere* could not be assigned the syntactic role of copula or that of auxiliary. Some examples are reported under (15).[19]

(15) a. SI, MA SONO FAVORE ANCHE TI SONO
Yes but $BE_{1st\ sing.}$ favour also you_{dat} $BE_{1st\ sing.}$
LAVORO PER OPERAIO.
$work_{1st\ sing./Noun}$ for worker

b. NON È VERO, NON IL TELEFONARE È CELLULARE!
not is true not the phoning $BE_{3rd\ sing.}$ mobile phone

c. MA SONO IL CAMPIONATO DILETTANTI GIRONE "C"
but $BE_{1st\ sing.}$ the championship amateur Round C
ERA A XXX HO PERSO SONO CLASSIFICA PER
$BE_{3rd.sing.past}$ in XXX HAVE lost $BE_{1st\ sing.}$ classification for
31 PUNTI.
31 scores

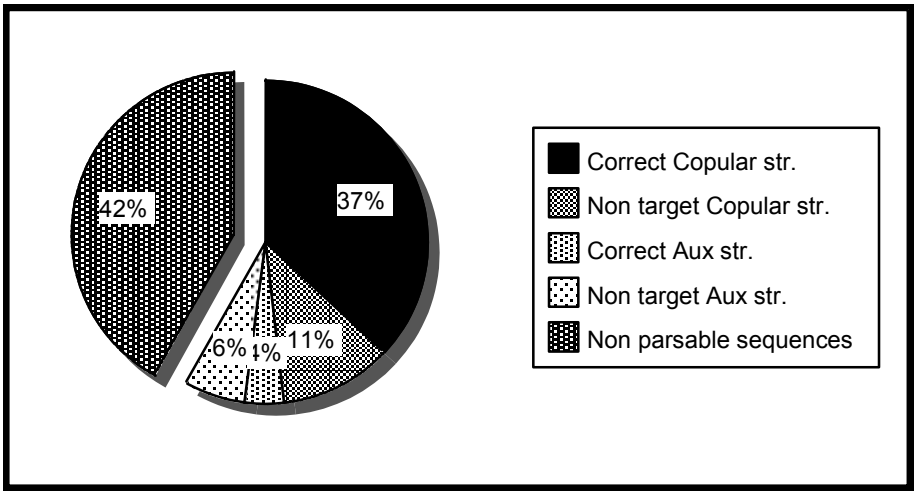
d. IERI ERO MOLTO FRESCO SONO L'ARIA
Yesterday $BE_{1st\ sing.\ past}$ very fresh $BE_{1st\ sing.}$ the air
PERCHÉ UMIDA SONO VENTO.
because damp $BE_{1st\ sing.}$ wind

e. Si, sono poco piove!!!
Yes, $BE_{1st\ sing.}$ little $RAIN_{3rd\ sing}$

As summarised in *Table 5*, 42% of the total of Gabriele's uses of *essere* were impossible to code as copular or auxiliary constructions.

[19] Each sequence presented under (15) is drawn from a different session. The over-use of forms of *essere* is not limited to a specific session and no evolution over time emerges along the 5 months of data collection.

Table 5: Proportion of correct, incorrect and non parsable forms of *essere* in Gabriele's data.



In Gabriele's Italian production data, overuse of forms of *essere* emerges instead of the systematic and highly constrained omission pattern that emerges from normal acquisition data, as summarised in *Table 6*.

Table 6: Omission versus random overuse of *essere*

Form of ESSERE	Acquisition data	Deaf data
Omission	**45 %**	1 %
Random over use	0 %	**42 %**

3.2.2. *Agreement data*

Once the 85 instances of *essere* that can actually be analysed as copular or auxiliary forms are taken into account, 16% of non target forms emerge with respect to verbal agreement.

Agreement errors do not appear to be restricted to any specific syntactic context, and mainly appear in preverbal subject contexts.

The non target forms are not limited to 3rd person plural contexts, but are distributed across different contexts, as shown in (16).

(16) a. Gli occhiali è blu [the glasses$_{plur.}$ BE$_{3rd\ sing}$ blue]
b. [io] è contento [I BE$_{3rd\ sing}$ happy$_{masc/sing}$]

c. Elena sono alta. [Elena $BE_{1st\ sing}$ $tall_{fem/sing.}$]

Table 7: The distribution of agreement errors in Gabriele's corpus.

Contexts	SONO	E'	SEI	Errors %
I p. sing.	13	2 (*)	-	13 % (2 /15)
II p. sing.	1 (*)	-	6	14 % (1 / 7)
III p. sing.	5 (*)	48	1 (*)	11 % (6 54)
III p. plur.	4	5 (*)	-	56 % (5 / 9)
TOTAL	6* /23	7* / 55	1* /6	16 % (14 /85)
%	**26.09 %**	**12.73 %**	**16.67 %**	**16.47 %**

In Gabriele's corpus, the 3rd person singular form *È* emerges in more than a half of the 3rd person plural contexts. *È* also appears in 1st person singular contexts. The 1st person singular form *Sono* appears in 3rd person singular contexts. In these contexts also the 2nd person singular *Sei* can emerge (relevant data in *Table 7*).

The data indicate spontaneous use of the two forms *Sono* and *È* only.[20] Those forms are selected in the appropriate context most of the time, but they also freely appear 16% of the time in all other syntactic contexts.
In Gabriele's corpus, agreement errors are more frequent than they are in normal acquisition data and they are not as predictable and syntactically restricted, as shown in *Table 8*.

Table 8: Syntactically constrained errors versus random selection of forms of *essere* in Child Italian and in Gabriele.

Agreement Errors	Child Italian	Gabriele
I p. sing.	0 %	13 %
II p. sing.	0 %	**43 %**
III p. sing.	0 %	11 %
III p. plur.	**10 %**	**56 %**

[20] The form *Sei* in fact emerges only in contexts of elicited production such as (i) and is therefore not considered productive in Gabriele's system:

(i) E. Sono magra? [$BE_{1st\ sing}$ $skinny_{fem/sing.}$]
G. Sì, sei magra. [Yes, $BE_{2nd\ sing}$ $skinny_{fem/sing.}$]

3.2.3. *Comprehension data*

Given the syntactically unpredictable pattern emerging from the deaf subject's production data, comprehension data were considered too.

Data on the comprehension of agreement features on the copula was elicited through minimal pairs of declarative sentences such as (17) or Yes/No questions such as (18).[21]

(17) Sono in piscina. [$BE_{1st\ sing}$ in the swimming pool]
Sei in piscina. [$BE_{2nd\ sing}$ in the swimming pool]

(18) È seduta? [$BE_{3rd\ sing}$ sitting]
Sono seduta? [$BE_{1st\ sing}$ sitting]

The comprehension tasks were designed so that only subjects able to "read" the syntactic information expressed by the verbal morphology could produce a correct answer.

Gabriele's correct answers were 37% of the total. His data thus indicate no access to the syntactic information expressed by the verbal morphology.

Gabriele's comprehension of the specific information carried by the verbal morphology was further investigated with a second task in which he was asked to identify the referent of the subject in overt pronominal subject contexts. Some examples are under (19) and (20).[22]

(19) E. Tu sei Elena? [You are Elena]
G. Sì (#) [Yes]
E. Tu sei Gabriele? [You are Gabriele]
G. Sì [Yes]

(20) C. Io sono Gabriele? [I am Gabriele]
G. No. [No]
C. Io sono Elena? [I am Elena]
G. No. [No]
C. Io sono Carol? [I am Carol]
G. No (#) [No]

[21] Gabriele was asked to write WHO was in the swimming pool after reading a sentence such as (17a/b) or asked to answer Yes or No to questions such as (18a/b). The yes/no questions were structured so that all the information was available from the extra-linguistic context, once the referent of the syntactic subject was identified.

[22] Questions in (19) were asked by Elena, questions in (20) by Carol. Answers marked with (#) are not correct in the given context.

Gabriele's performance in this task was at chance level (50% of correct responses). The same performance emerges when he is asked to identify the subject of copular forms he had produced himself, as shown in (21) and (22).[23]

(21) E. Sono magra? [$BE_{1st\ sing}$ $skinny_{fem/sing.}$]
G. Sì, sei magra. [Yes, $BE_{2nd\ sing}$ $skinny_{fem/sing.}$]
E. Chi? [Who?]
G. Carol (#) [Carol]

(22) C. Sono magra? [$BE_{1st\ sing}$ $skinny_{fem/sing.}$]
G. Sì, sei magra. [Yes, $BE_{2nd\ sing}$ $skinny_{fem/sing.}$]
E. Chi? [Who?]
G. Elena (#) [Elena]

Gabriele's ability to respect syntactic constraints on subject-verb number agreement was further investigated through grammaticality judgments on sentences such as (23) and (24).[24]

(23) a.*I capelli è lunghi. [The $hair_{plur}$ $BE_{3rd\ sing}$ long]
b. I capelli sono lunghi. [The $hair_{plur}$ $BE_{3rd\ plur}$ long]

(24) a. La penna è nuova. [The pen_{sing} $BE_{3rd\ sing}$ new]
b.*Le penne è nuova. [The pen_{plur} $BE_{3rd\ sing}$ new]

Gabriele's performance on the grammaticality judgement task was below chance level (35% of correct answers). Some of his answers are reported under (25) and (26).[25]

(25) a. *√ La penna sono sul tavolo.
The pen_{sing} $BE_{3rd\ plur}$ on the table
b. √√ Le penne sono sul tavolo.
The pen_{plur} $BE_{3rd\ plur}$ on the table
c. √√ I miei capelli sono biondi e neri.
My $hair_{plur}$ $BE_{3rdplur}$ blond and black

[23] In this task Gabriele is not able to identify the referent of the 2nd person singular form *Sei*. This fact further suggests a non productive use of this form. Answers marked with (#) are not correct in the given context.
[24] It is important to notice that every native speaker of Italian would give a straightforward answer to the grammaticality judgment task proposed here.
[25] The left most column indicates a native speaker's answer, the second indicates Gabriele's answers.

d. *√ I miei capelli è biondi e neri.
My hair$_{plur}$ BE$_{3rd\ sing}$ blond and black

(26) a. √* La penna è blu.
The pen$_{sing}$ BE$_{3rd\ sing}$ blue
b. *√ La penna sono blu.
The pen$_{sing}$ BE$_{3rd\ plur}$ blue
c. *√ Le penne è blu.
The pen$_{plur}$ BE$_{3rd\ sing}$ blue
d. √* Le penne sono blu.
The pen$_{plur}$ BE$_{3rd\ plur}$ blue

An average performance at chance level or below emerges from Gabriele's data on the perception of the information carried by the verbal morphology and from his data on the perception of agreement constraints.

4. Conclusion

Gabriele's data do not show the pattern of regular and syntactically restricted omission or use of a default form that can be found in normal acquisition data, where functional structure is supposed to be active.

Table 9: The contrast found between Child Italian data (C.I.) and the deaf person data (G.)

Forms of ESSERE	C.I.	G.
Omission	**45 %**	1 %
Random Over Use	0 %	**42 %**
Agreement errors	**1%**	**19%**

Agreement errors	C.I.	G.
I p. sing.	0%	13%
II p. sing.	0%	**43%**
III p. sing.	0%	11%
III p. plur.	**10%**	**56%**

In Gabriele's production data, the rate of correct use of *essere* copula and *essere* auxiliary is only 41%. If comprehension of copular constructions is taken into account, the correct answers rate lowers further to 37% and 35%. The pattern emerging from production and comprehension data seems then to indicate a general strategy of random selection of forms of *essere*, to which no syntactic content is assigned.

Given those facts, it seems possible to conclude that the deaf subject data concerning *essere* do not correspond to a syntactically constrained system. As proposed in Borer and Rohrbacher (2003), random over use of functional material indicates no availability of functional structure. Gabriele's

performance seems then to suggest - in the case of *essere* - a non syntactically constrained language behaviour.

The emergence of such an unexpected pattern from Gabriele's data could be correlated to the deprivation of the primary linguistic data to which he has been exposed to. The data so far discussed seem then to suggest that deafness can in some cases drastically reduce the quantity and alter the quality of the linguistic information present in the input so that no triggers are available to activate the process of language acquisition.

References

Becker, Misha Karen. 2000. *The Development of the Copula in Child English: the Lightness of Be*. Ph.D. dissertation, UCLA.

Belletti, Adriana. 1999. *Generalized Verb Movement: Aspects of Verb Syntax*, Torino, Rosenberg & Sellier.

Borer, Hagit and Bernhard Rohrbacher. 2003. "Minding the absent: Arguments for the Full Competence Hypothesis", *Language Acquisition* 10.2:123-176.

Chomsky, Noam. 1980. *Rules and Representations*, Oxford, Basil Blackwell.

Franchi, Elisa. 2003. "La Logogenia come strumento per individuare sistemi di comunicazione non sintatticamente fondati. Il caso delle strutture copulari nei sordi", paper presented at *Premier Congreso Internacional de Logogenia*, México D.F., October 2003.

Franchi, Elisa. 2004a. "Copula Omission as evidence for Truncation in Italian Child Language", paper presented at the *Second Lisbon Meeting on Language Acquisition*, Lisbon, June 2004.

Franchi, Elisa. 2004b. *Piena competenza e assenza di competenza linguistica: una distinzione messa in luce dalla logogenia.* Essere *copula e ausiliare in Italiano infantile e in un sordo profondo prelinguale non segnante*. Ph.D. dissertation, University of Siena and University of Florence.

Franchi, Elisa. In press_a. "L'acquisizione dei verbi funzionali in Italiano infantile", *Rivista di Grammatica Generativa* 29.

Franchi, Elisa. In press_b. "Patterns of copula omission in Italian Child Language". In V. Torrens and L. Escobar (eds.) *The Acquisition of Syntax in Romance Languages.* Amsterdam/Philadelphia, John Benjamins.

Franchi, Elisa and Debora Musola. 2002. "La Logogenia. Verificare e sviluppare la comprensione di informazioni sintattiche nell'intervento con sordi che non sanno l'italiano", poster presented at *III Convegno dell'Associazione Italiana Linguistica Applicata*, Perugia, February 2002.

Friedemann, Marc Ariel and Luigi Rizzi. 2000. *The Acquisition of Syntax.* Harlow, Longman.

Goldin-Meadow, Susan. 2003. *The Resilience of Language*. Psychology Press.

Guasti, Maria Teresa. 1993/94. "Verb Syntax in Italian Grammar: Finite and Non-finite Verbs", *Language Acquisition* 3:1-40.

Guasti, Maria Teresa. 2002. *Language Acquisition: the growth of grammar.* Cambridge MA, MIT Press.

Guasti, Maria Teresa and Luigi Rizzi. 2002. "Agreement and tense as distinctive syntactic positions. Evidence from acquisition". In G. Cinque (ed.) *Functional structure in DP and IP: The Cartography of Syntactic Structures*. Vol. 1, 167-194. New York, Oxford University Press.

Lightfoot, David. 1999. *The Development of Language: acquisition, change and evolution.* Malden MA, Blackwell.

Mac Whinney, Brian and Catherine Snow. 1985. "The Child Language Data Exchange System". *Journal of Child Language* 12:271-296.

Moro, Andrea. 1993. *I predicati nominali e la struttura della frase*. Padova, Unipress.

Musola, Debora. 2003. "Lo sviluppo della deissi personale in italiano (L1) attraverso la Logogenia: primi dati da due casi-studio e osservazioni metodologiche", paper presented at *Premier Congreso Internacional de Logogenia*, México D.F., October 2003.

Musola, Debora. In press. "L'acquisizione della deissi personale in italiano (L1) in condizioni di sordità preverbale: due casi studio", *Atti del XXVIII Convegno Annuale della Società Italiana di Glottologia - Acquisizione e mutamento di categorie linguistiche*.

Radelli, Bruna. 1985. *La ambigüedad. Un rasgo significativo para el análisis sintáctico*. Colección Científica, INAH, México D.F.

Radelli, Bruna. 1990. "El cuál y el cómo en la sintaxis del español". In B. Garza Cuarón and P. Levy (eds.) *Homenaje a Jorge A. Suárez*. 437-445. El Colegio de México, México D.F.

Radelli, Bruna. 1997. "Significados sintácticos" In M. Pool Westgaard (ed.) *Estudios de Lingüística Formal*. 237-256. El Colegio de México, México D.F., CELL.

Radelli, Bruna. 1998. *Nicola vuole le virgole: dialoghi con sordi. Introduzione alla Logogenia*. Padova, Decibel Zanichelli.

Radelli, Bruna. 1999. "La Logogenia en el desarrollo de los sordos", *Memorias del XV Congreso FEPAL*. 169-190. Facultad de Ciencias de la Educación de la Universidad de A Coruña, España.

Radelli, Bruna. 2000. "Una nueva aplicación de la lingüística: la Logogenia". *Dimensión Antropológica* 23:51-72. INAH, México D.F..

Rizzi, Luigi. 1994. "Root Null Subjects and Early Null Subjects". In T. Hoekstra and B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*. 151–176. Amsterdam/Philadelphia, John Benjamins.

Rizzi, Luigi. 1993/94. "Some notes on Linguistic Theory and Language Development: the case of Root Infinitives", *Language Acquisition* 3:371-395.

Rizzi, Luigi. 2000a. *Comparative Syntax and Language Acquisition*. London, Routledge.

Rizzi, Luigi. 2000b. "Remarks on Early Null Subjects". In M. A. Friedeman and L. Rizzi (eds.) *The Acquisition of Syntax.*268-292.London, Longman.

Rizzi, Luigi. 2004. "On the Grammatical Basis of Language Development: A Case Study". In G. Cinque and R. Kayne (eds.) *Handbook of Comparative Syntax*. 70-109. New York, Oxford University Press.

Rothstein, Susan. 1995. "Small Clauses and copular constructions". In Cardinaletti, A. and M. T. Guasti (eds.) *Syntax and Semantics 28: Small Clauses*. 27-48. New York, Academic Press.

Starke, Michal. 1995. "On the format for Small Clauses". In Cardinaletti, A. and M. T. Guasti (eds.) *Syntax and Semantics 28: Small Clauses.*237-269. New York, Academic Press.

Wexler, Ken. 1994. "Optional Infinitives, Head Movement and the Economy of Derivations in Child Grammar". In D. Lightfoot and D. Hornstein (eds.) *Verb Movement*. 305-350. Cambridge MA, Cambridge University Press.

Wexler, Ken. 1998. "Very Early Parameter Setting and the Unique Checking Constraint: a new explanation of the Optional Infinitives Stage", *Lingua* 106:23-79.

Structure and frequency in verb classification

Paola Merlo and Suzanne Stevenson
University of Geneva - University of Toronto

1. Introduction: the role of quantitative approaches in the formal study of language.

In this paper, we investigate the linguistic relevance of the notion of frequency in theories of lexicon organisation, in particular the definition of verb classes. Traditionally, the subject matter of linguistics has been to develop linguistic representations to describe and explain language as a cognitive process, including language acquisition and language comprehension and production. Theories that have been developed to adhere to this research plan have largely been of the symbolic, algebraic, categorical kind. Quantitative methods and corpus-based data collection have been used extensively in the study of language acquisition, language processing, historical linguistics and sociolingustics, but they have been systematically excluded from the representations and the methods used in the study of formal grammars. The only recognition of non-categorical phenomena in traditional grammars has been the notion of markedness. But that need not be. Investigations of the link between a richly structured linguistic theory and the distributional properties of language are not contradictory with the goals of generative grammar. The availability of probabilistic information has been shown to affect the learnability of a language and to enable learning with less and poorer data (Horning 1969). There is also ample evidence of frequency effects in language processing (Seidenberg, MacDonald and Saffran 2002, MacDonald, Pearlmutter and Seidenberg 1994). Therefore, an investigation of the quantitative correlates of abstract linguistic concepts might in fact enhance the goals of generative grammar by elucidating the theoretical relationship between structure and frequency.

Current development of large text repositories and syntactically annotated databases, and the exponential growth of computational and storage power, allow us to ask foundational questions on the role of frequency and quantitative data in the development of theories of grammar. Along with other researchers recently, we believe that enriching traditional structural representations with quantitative information will provide stronger data, and consequently could emit predictive hypotheses in areas that were before underspecified (Bresnan, Dingare and Manning 2001, Bod, Hay and Jannedy 2003, Keller 2000, Manning 2003). As in other empirical sciences, linguistic data can be arranged on a scale of expressiveness from: nominal – categorically discrete data that cannot be ordered on a scale, such as eye colour or subcategorisation frames; to ordinal – categorically discrete data that can be ordered on a scale, such as shades of colour; to numerically discrete – such as population size; to numerically continuous – such as body weight or probabilities. The data used currently in formal grammars is nominal. Nominal data are the least expressive as they cannot be compared or ordered, and few statistical techniques can be applied to them. Quantitative data support more elaborate theories, which take into account some non-categorical facts about language. For example, we can reach a better understanding of what the phenomenon of markedness really is. Because quantitative data is more expressive, they also support theories that have fewer a priori assumptions, without losing explanatoriness and predictiveness.

In this paper, we will show that principles of the verbal lexicon organisation – verb classes – show robust statistical regularities within and across languages, and we will hypothesize that this is because these frequencies are surface reflexes of underlying thematic regularities. If taken as a piece of data in its own right, frequency then becomes a tool for discovery of underlying abstract linguistic properties.

2. Case Study: Verb Classes

One of the most influential recent research programmes on the structure of the lexicon, Levin's (1993) work on verb classes, aims at reducing the information in a lexical entry to its primitive meaning components (see also Levin 1985, Pinker 1989). Under the hypothesis that semantic properties of verbs largely determine their syntactic behaviour, the linguistic knowledge about a verb consists in its specific set of meaning components along with general relations between each meaning component and its possible syntactic expressions.

Specifically, the behaviour that Levin suggests as key to verb classification is the notion of *diathesis alternation* – an alternation in the

expression of the arguments of a verb, such as, e.g., the causative/inchoative alternation in *The chef melted the butter*/*The butter melted*. Levin proposes a two-stage approach. First, the semantic classes to which verbs belong are revealed empirically by the diathesis alternations they participate in. For example, *cut* and *break* can occur in the middle alternation, while *hit* and *touch* cannot. On the other hand, *hit*, *touch* and *cut* can occur in the conative alternation, while *break* cannot. Second, once classes of verbs are individuated based on contrastive syntactic behaviour, one can propose substantive hypotheses on what meaning components best describe the observed classification. For example, verbs whose meaning requires a notion of contact can participate in the conative alternation, while verbs that do not imply physical contact, such as *break*, cannot. Using this method, Levin classifies 3024 English verbs in approximately 200 verb classes. Work by Merlo and Stevenson (2001), like others in a computational framework, have extended this idea by showing that *statistics* over the alternants of a verb effectively capture information about its class (Lapata 1999, McCarthy 2000, Schulte im Walde 2000, Lapata and Brew 2004).

Let's look at three main verb classes of English that participate in a transitivity alternation, as indicated. In Levin's account, they are distinguished from each other because the particular transitivity alternation they occur in is different in each case; however, the allowed alternants are identical for all of them – i.e., they can all be transitive or intransitive.

Manner of Motion					% Trans Usage
(1) a.	The rider	raced	the horse	past the barn	23
b.	The horse	raced	past the barn		77
Change of State					
(2) a.	The cook	melted	the butter		40
b.	The butter	melted			60
Creation/Transformation					
(3) a.	The contractors	built	the house		62
b.	The contractors	built	all summer		38

We can notice that, even though the alternations do not distinguish the verbs at the syntactic level, the alternants occur across the classes with very different frequencies. Manner of motion verbs are used transitively 23% of the time and are used intransitively 77% of the time, while change of state verbs are used transitively 40% of the time and intransitively 60%, and creation/transformation verbs are more frequently transitive (62%) and less

frequently intransitive (38%). These frequencies are derived by automatic counts taken from samples of 20 verbs in each class over 65 million words of Wall Street Journal text. All the differences are statistically significant. These significantly different frequencies raise several questions about the theoretical status and the generality of these frequency facts in syntax.

Question 1: Are these frequencies linguistic facts or do frequencies vary in a way that is unrelated to the abstract linguistic description? If frequencies are linguistic data, they require explanation. In particular, we need to explain why classes participate in grammatically licensed alternations so differently.

Question 2: How general are these differences in frequency distributions? Are such differences typical of all different verb classes? Moreover, is this statistical trend predictive – i.e., is the statistical trend strong enough to be definitional of the class?

Question 3: Do these differences in frequencies hold across languages? Do they reveal some commonalities across languages?

We answer these questions in the following sections in turn. The methodology is computational and experimental. Drawing on work presented in Merlo and Stevenson 2001, we first show that frequency differentials can be systematically derived from abstract properties of the verb class. We then use automatic learning techniques to explore the amount of generality of the proposed representations and of their frequency properties, showing that frequency differentials are useful in learning several new classes and across a new language.

3. Frequency, thematic roles, and animate subjects

In this section we will introduce the notions of markedness and harmonic scales to explain the connection between different lexical semantic classes and their different frequency distributions in the use of the transitive construction.

3.1. *Thematic roles and frequency*

Recall that the first question that we want to answer is whether these differences in the relative frequency of the transitive use across classes is related to other underlying abstract properties of the formal grammar. The answer to this question is positive. Drawing from previous work (Merlo and

Stevenson 2001), we will show that the difference in frequency of transitive use is related to different thematic assignments, and eventually possibly to different underlying lexical composition processes (Hale and Keyser 1993, Stevenson and Merlo 1997).

Let's look again at the examples using the verbs in question, this time indicating the thematic assignment of the participants in the event described by the verb.

Manner of Motion

(1) a.	The rider	raced	the horse	past the barn
	Causal Agent		*Agent*	
b.	The horse	raced	past the barn	
	Agent			

Change of State

(2) a.	The cook	melted	the butter
	Causal Agent		*Theme*
b.	The butter	melted	
	Theme		

Creation/Transformation

(3) a.	The contractors	built	the house
	Agent		*Theme*
b.	The contractors	built	all summer
	Agent		

Manner of motion verbs are intransitive action verbs whose transitive form, as in (1a), can be the causative counterpart of the intransitive form (1b). The type of causative alternation that manner of motion verbs participate in is the "induced action alternation" according to (Levin 1993). For our thematic analysis, we note that the subject of an intransitive activity verb is specified to be an Agent. The subject of the transitive form has the label Causal Agent, which indicates that the subject role is introduced with the causing event. In a causative alternation, the semantic argument of the subject of the intransitive surfaces as the object of the transitive (Brousseau and Ritter 1991, Hale and Keyser 1993, Levin 1993, Levin and Rappaport Hovav 1995). Since for manner of motion verbs this argument has agentive properties, the alternation yields an object in the transitive form that receives an Agent role (Cruse 1972, Stevenson and Merlo 1997).

The sentences in (2) illustrate the corresponding forms of a change of state verb, *melt*. Change of state verbs are intransitive, as in (2b); the transitive counterpart for these verbs also has a causative form, as in (2a). This is the "causative/inchoative alternation" (Levin 1993). Like manner of motion verbs, the subject of a transitive change of state verb is marked as the Causal Agent. Unlike manner of motion verbs, though, the alternating argument of this class of verbs (the subject of the intransitive form that becomes the object of the transitive) is a passive entity undergoing a change of state, and is therefore a Theme.

The sentences in (3) illustrate another class of verbs that can be both transitive and intransitive, creation or transformation verbs such as *build*. These are activity verbs that exhibit a non-causative transitivity alternation, in which the object is simply optional. The thematic assignment for these verbs is simply Agent for the subject (in both transitive and intransitive forms), and Theme for the optional object. We will call these classes MOM, COS and C/T for brevity's sake in what follows. Table 1 summarizes the difference in thematic assignments.

Table 1: Thematic assignments for classes undergoing a transitivity alternation

Class	Transitive		Intransitive
	Subject	Object	Subject
Manner of motion (MOM)	Causal Agent	Agent	Agent
Change of state (COS)	Causal Agent	Theme	Theme
Creation/ Transformation (C/T)	Agent	Theme	Agent

Can we explain the different frequency of usage of the transitive construction for these classes, based on their properties as reflected in their thematic assignment?

3.2. *Subcategorisation and frequency*

The Prague school's notion of linguistic markedness (Jakobson 1939, Trubetzkoy 1939) enables us to establish a scale of markedness of these thematic assignments and make a principled prediction about their frequency of occurrence. Typical tests to determine the unmarked element of a pair or scale are: *simplicity* – the unmarked element is simpler; *distribution* – the

unmarked member is more widely attested across languages; and *frequency* – the unmarked member is more frequent (Greenberg 1966, Moravcsik and Wirth 1983). The claim of markedness theory is that, once an element has been identified by one test as the unmarked element of a scale, then all other tests will be correlated. The three thematic assignments appear to be ranked on a scale by the simplicity and distribution tests, as we describe below. From this, we can conclude that frequency, as a third correlated test, is also predicted to be ranked by the same scale, and we can therefore explain the observed frequencies of the three thematic assignments.

First, transitive MOM and COS verbs have a causative meaning. Since there are two events involved in a causative form, we assume that transitivity by causation has a more complex representation than simple transitives, as in the C/T verbs. Moreover, transitive MOMs are slower to process than COS transitives (Filip Tanenhaus and Carlson 1998), and the former can cause garden path effects even when they are not ambiguous (Stevenson and Merlo 1997).[1] Transitive MOMs are therefore more complex than transitive COS verbs from a processing point of view. We have thus established a scale of complexity for these three classes in a transitive usage from most (MOM) to least (C/T) complex, with COS intermediate in complexity.

We further observe that the causative transitive of a manner of motion verb has an Agent thematic role in object position which is subordinated to the Causal Agent in subject position, yielding an unusual "double agentive" thematic structure. This lexical causativization (in contrast to analytic causativization) of manner of motion verbs, which are unergatives, is found in fewer languages than lexical causatives of change of state verbs, which are syntactically unaccusative. In asking native speakers about our verbs, we found that lexical causatives of MOM verbs are not attested in Italian, French, German, Portuguese, Gungbe (Kwa family), and Czech. On the other hand, the transitive causatives are possible for change of state verbs (i.e., where the object is a Theme) in all these languages. The typological distribution test thus indicates that transitive manner of motion verbs are a distributionally rarer phenomenon than transitive change of state verbs.

Since markedness is indicated by complexity and distributional rarity, from the above observations, we can conclude that manner of motion verbs have the most marked transitive argument structure, change of state verbs have an intermediately marked transitive argument structure, and

[1] We gave a processing explanation of the fact that these verbs cause a garden path, which was grounded in a specific extension of Hale and Keyser's (1993) lexical syntax proposal. We developed a specific representation for these cases which require an extra level of embedding, hence are more complex. Combined with Stevenson's competitive processing model (Stevenson 1994), we obtained the observed effects.

creation/transformation verbs have the least marked transitive argument structure of the three. Under the assumptions of markedness theory outlined above, we can then account for the observed behaviour: that manner of motion verbs are the least frequent in the transitive, change of state verbs have intermediate frequency in the transitive, and creation/transformation verbs are the most frequent in the transitive.

3.3. *Animacy and frequency*

Are there other properties of verb classes that we can expect to surface as statistical differences? Animacy is a property for which we can expect differential statistical values typical of the class, as it reflects underlying thematic assignments. Recall the pattern of thematic assignments, in Table 1 above. The only non-agentive subject occurs in the intransitive form of change of state verbs, which has a Theme subject. This fact has consequences for the frequency distribution of animate subjects in these classes: we expect COS verbs to have fewer animate subjects than the other two classes because we expect that Themes are less likely to be animate. This expectation follows from a combination of recent theories on the alignment of hierarchies and the thematic and animacy properties of these classes.

Recall current proposals for the harmonic combination of hierarchies where ">" indicates higher harmony.

> Alignment: Suppose given a binary dimension D1 with a scale X>Y on its elements {X,Y}, and another dimension D2 with a scale a>b>...>z on its elements. The harmonic alignment of D1 and D2 is the pair of Harmony scales:
>
> Hx: X/a > X/b>...X/z
>
> Hy: Y/z>....>Y/b>Y/a
>
> (Prince and Smolensky 1993, p.136)

In the case of thematic roles and of animacy we have the two prominence scales in which the relevant values combine most harmonically as Animate/AGENT > Animate/THEME:

> Animacy Hierarchy 1,2 > 3,Proper > Human > Animate > Inanimate
>
> (Silverstein, 1976)

> Thematic hierarchy AGENT > THEME

Consequently, according to the theory of markedness, it is less marked and therefore more frequent to express an Agent with an animate entity than to express a Theme with an animate entity. We can predict that change of state verbs (the only class with a Theme subject possibility) will therefore have a lower frequency of animate subjects than the other two classes.

The predictions concerning animacy use are fully borne out by an analysis of the data across the three classes under discussion. Table 2 shows the mean relative frequencies of the two linguistic properties we have considered: use of the transitive construction and of animate subjects.

Table 2: Mean Relative Frequencies of the Data for Two Linguistic Properties

Verb Class	Linguistic Property	
	Transitive Use	Animate Subject
Manner of Motion	23%	25%
Change of State	40%	7%
Creation/Transformation	62%	15%

The data is automatically collected (over 65 million words of text), and is therefore an approximation of actual usage. All the reported differences of mean relative frequencies are statistically significant at $p<.01$. We therefore confirm statistically all the predicted orders among the classes which were hypothesized based on the relationship of frequency of transitive use and of animacy, to underlying thematic assignments of the classes of verbs.[2]

4. Generalising to new linguistic entities: the machine learning approach to theory testing

Question 2 and question 3 in the introduction mention two ways of testing whether the observed relationship between abstract linguistic properties and frequency is an idiosyncrasy of the classes under examination or a truly general and predictive property. We ask: How well do these distributional properties generalise across new verbs, across new classes and across languages? In this section, we set up the generalisation test as an automatic classification problem. We use the ability to classify new instances as a method to test the generalising power of the correlation between defining properties of the lexical semantic classes and corresponding frequencies. We

[2] While we confirm all predicted orders, we also observe an unpredicted distinction between the C/T and MOM classes on the animacy feature. Investigation of the possible linguistic causes of this difference is left for future research.

test if the statistical differences observed in the previous section are strong enough to drive an automatic learner.

Formally, we say that a computer program learns from experience E with respect to some task T and performance measure P, if its performance at task T, as measured by P, improves with E. In our case, the training experience E will be provided by a database of correctly classified verbs; the task T consists in classifying verbs unseen in E into predetermined semantic classes; and the performance measure P will be defined as the percentage of verbs correctly classified. This learning paradigm is called *supervised learning*, because of the training phase, in which the algorithm is provided examples with the correct answers. During this phase the algorithm develops rules to describe all the training data in a compact way. A possible rule in our setting could be, for example: "*If animacy is less than 10% then verb is COS"*. In the testing phase, these rules are applied to additional verb data, not included in the training phase. The accuracy of classification on the test set indicates whether the rules developed in the training phase are general enough, yielding good test accuracy, or are too specific to the training set to generalise well to other data, thus yielding bad performance in the testing phase. There are numerous algorithms for learning in a supervised setting, and many regimes for training and testing such algorithms. In the following experiments, we use a decision tree induction learning algorithm, C5.0 (Quinlan, 1993), publicly available at http://www.rulequest.com, and 10-fold cross-validation repeated 10 times as the training and testing protocol.

A *decision tree* is a tree in which each branch node represents a choice between a number of alternatives, and each leaf node represents a classification or *decision*. The C4.5 class of decision tree induction algorithms use information theory to decide which choices provide the best partitioning of the input training data. This algorithm has good generalisation ability on many problems and yields highly readable output in the form of symbolic rules.

Cross-validation is a training and testing protocol in which the system randomly divides the data into n parts, and then runs the learner n times, using n-1 partitions for training and the remaining one for testing. At each run of the learner, a different partition is chosen for testing. This procedure is repeated m times with a different random division of the data, and the performance measure averaged over all $n * m$ experiments. When the number of data items in each class (in our case, verbs) is relatively limited, this methodology avoids the possible bias that could result from a single random split into training and testing items.

In order to present our verbs to the algorithm, each verb is encoded as a vector in which the frequencies of the identified linguistic properties serve as statistical features, as exemplified below.

Vector template: [verb, TRANS, ANIM, class]
Example: [open, .69, .36, COS]

Results confirm that the frequency correlates of the linguistic properties illustrated in the previous section are strong enough to support learning at a very good level of performance. In a task whose random baseline is approximately 33%, as it is a three-way choice, we reach performance of 70%. This corresponds to at least a 54% reduction in error rate over the baseline.[3] The class that is most accurately classified is the class of manner of motion verbs, indicating that its markedness is easy to spot in an automatic procedure.

An analysis of errors when the algorithm is run with access to different statistical features confirms that learning does indeed occur because of the hypothesized relation between the linguistic properties and observed frequencies, and not because of some uncontrolled artefact of the experiments. If we compare a tree in which the transitive feature is used to one in which it is not, we find that the transitive property improves the discrimination of all the classes. A tree in which animacy is not used, on the other hand, has worse identification of change of state verbs, as expected.

Thus, we can conclude that not only are the frequencies systematically related to underlying properties of a sample of observed verbs (providing descriptive statistics), but that frequency differentials are also strong enough to enable a learner to classify verbs that did not belong to the initial observed sample. These frequencies are predictive.

5. Generalising to new classes and to new languages

The classes of verbs presented in the previous section were chosen because they all undergo a transitivity alternation, and therefore their subcategorisation representation is the same. These classes, however, differ substantially and systematically in the percentage of use of the different subcategorisation frames that they license. In this section, we investigate other classes of verbs to show that they also exhibit differential frequency of use of their subcategorisation frames and the animacy of their subject, and

[3] This performance is achieved using a small number of other features related to the transitivity alternations, in addition to the TRANS and ANIM features we focus on here.

that such differentials are strong enough to support learning in these cases as well. We show moreover that this predictive differential in frequency of use extends to subcategorisation frames that involve a prepositional phrase and is not limited to the transitive-intransitive distinction. We look at Psychological State verbs, Dative/Benefactive verbs, and Locative verbs, which we exemplify below in examples (4) to (7).

Psychological State

(4) a. The rich love their money
b. The rich love too.

Dative/Benefactive

(5) a. Bill sold Tom a car
b. Bill sold a car to Tom (dative)

(6) a. Martha carved the baby a toy
b. Martha carved a toy for the baby (benefactive)

Locative

(7) a. Jack sprayed paint on the wall
b. Jack sprayed the wall with paint

Table 3: Summary of Subcategorisation Frames and Thematic Assignments

Verb Class	Alternant 1	Alternant 2
Psychological	NP V NP	NP V
	Experiencer, Stimulus	Experiencer
Dative/ Benefactive	NP V NP PP	NP V NP NP
	Agent Theme Goal/Beneficiary	Agent Goal/Beneficiary Theme
Locative	NP V NP PP	NP V NP PP
	Agent Locatum Location	Agent Location Locatum

5.1. Subcategorisation Frame

Differently from the other three classes that occur in a transitive-intransitive alternation, psychological verbs describe a non-volitional state. They can occur with an understood, generic object. Dative and benefactive verbs differ from the four previous classes of verbs because one of their arguments is a prepositional phrase or an indirect object. They describe a transfer or a benefactive action. Locative verbs have the meaning of putting/removing substances or things in/from containers or on/from surfaces. The substance

or thing that is moved is the locatum argument; the place (the container or surface) is the location argument. In each variant of the alternation one of the two arguments (either the locatum or the location) is expressed as the object of a preposition, while the other is expressed as a direct object.

The patterns of subcategorisation frames and thematic role assignments that distinguish these classes are shown in Table 3. One can notice that psychological verbs are simple transitives, without causation, and we predict therefore that they will be at least as frequent in the transitive form as the creation/transformation verbs. For the other two classes of verbs, we simply predict that use of particular prepositions will be a good predictor of the thematic roles assigned underlyingly.

5.2. *Animacy*

The notion of animacy of the subject which was developed in the previous section is relevant to all the classes of verbs in question. The subjects of psychological verbs are experiencers: they are likely to be animate since they must be able to experience a psychological state. The subject of dative/benefactive verbs is volitional, since it must have the intention that the goal or beneficiary receive the possession or the benefit of the object or the action. Thus, it is preferentially animate. Locatives are activity verbs, like creation/transformation and manner of motion, and their subject is preferentially animate. Since all the new classes preferentially have animate subjects, they are predicted to be more frequently animate than change of state verbs.

The different subcategorisation and animacy properties of the classes of verbs under consideration translate into different frequency distributions from class to class over the transitivity and animacy properties, as confirmed by the counts reported in Table 4.[4]

[4] One comment on the actual numbers in Table 4 is in order. The counts are collected automatically over a very large corpus, in this case the 100-million word British National Corpus. Counts of abstract notions such as animacy and, to a less extent, subcategorisation frame are therefore approximated. The numbers therefore are relevant only relationally and in their statistical properties, but their absolute values should not be taken to be an exact estimate of the phenomena in question. The fact that one can develop useful approximators is indeed in itself rather interesting, both from a computational point of view (Merlo and Stevenson 2001, Merlo, Stevenson, Tsang and Allaria 2002) and from the standpoint of language acquisition (Stevenson and Merlo 2001).

Table 4: Different frequencies of transitive use and animacy of classes

Class	Transitive Use	Animacy of Subject
MOM	0.09	0.35
COS	0.36	0.20
C/T	0.39	0.37
PSY	0.54	0.49
D/B	0.47	0.30
LOC	0.44	0.34

Table 5 illustrates overall accuracy and class by class results, in terms of precision and recall of the verbs in a machine learning experiment using subcategorization and animacy features. Precision is a measure of accuracy of the classification, and tells us how many of the verbs that the algorithm assigns to a given class actually belong to that class. Recall is a measure of coverage of the automatic classification and tells us how many of the verbs that actually belong to a class have been assigned to that class by the algorithm.

Table 5: Overall Performance and Class by Class Accuracy
(P=precision, R=recall)

Baseline (chance)											16.7
Best Performance using Subcategorisation and Animacy											56.7
MOM		COS		C/T		PSY		D/B		LOC	
P	R	P	R	P	R	P	R	P	R	P	R
67	40	36	80	67	40	75	60	80	80	50	40

The table shows that overall the algorithm classifies the verbs with 56.7% accuracy – that is, a 52% reduction of the error rate over the baseline. If we look at the class by class precision and recall, we observe that the D/B verbs are the best classified, because they very strongly select for the subcategorized preposition. All the other classes (except change of state verbs) have better precision than recall, in varying degrees. Change of state verbs, on the other hand, have much better recall than precision. This indicates that the algorithm has a tendency to assume that verbs are change of state, as a general rule. This is an interesting result, since the class of change of state verbs is one of the largest in Levin's classification, and does therefore constitute a very general case.

The main conclusion that we can draw from these results is that the methodology extends well to new classes, to new roles, and new subcategorisation frames. Globally, there is a reduction in error rate of 52%

over the chance baseline. This indicates that the features used for the classification are of general validity, and are not limited in application to the verb classes they were initially intended for.

We present now a final set of experiments which were developed to extend the investigation to Italian, by automatically classifying Italian verbs following the same methodology as we did for English. The goal of this experiment is to verify that the observed correlation between verb classes and different frequencies are attested across languages, and that they have the same learning power that they have in English. In order to make comparisons, we set out the experiment to be as similar as possible to the previously performed experiments on English verbs. We consider five of the six classes studied for English: the psychological, change of state, creation/transformation, manner of motion, and locative verbs. We choose the particular experimental verbs by translating, as far as possible given our translation procedure, the English experimental verbs. Moreover, we use the same features that were developed for English, to demonstrate that these properties are cross-linguistically valid, even though the Italian classes do not always allow the same alternations as their English counterparts. The training and testing regime is the same as the one described above.

The experiment is a five-way discrimination among classes that contain an equal number of verbs. Its baseline is, therefore, 20% accuracy. We obtain 50% accuracy based on the differentials of transitive use and animacy. This is a reasonably good performance, giving a 37.5% reduction in error rate.[5] Here we observe that manner of motion verbs are the best classified, while locative verbs are the worst. This result enables us to conclude that the same relationship between frequency and abstract linguistic notions related to verb classes holds for Italian, and is not, therefore, a specific property of English.

6. Conclusions

In this paper, we set out to answer three questions concerning the theoretical status of differential frequencies of some abstract syntactic properties across lexical semantic classes, their generality in a theory of lexical representation and their cross-linguistic validity. Through a set of computational experiments, we have shown that differences in frequency of transitive use

[5] There is reason to think that the lower absolute performance for Italian is a side-effect of the difficulty of estimating animacy automatically. Italian is a null subject language, with a clear preference for unexpressed subjects. Our estimate is based on expressed subjects and therefore is probably not as accurate as the estimate for English, as it suffers from sparse data. The development of estimates for understood elements is for future work.

and animate subjects in several classes of English and Italian verbs are systematically and predictably different.

Beside its direct relevance for a theory of lexical organisation and representation, this finding is also relevant for language acquisition studies. One of the fundamental questions of child language acquisition concerns the cues and mechanisms that are available to the child to learn the lexical semantics of the verbal lexicon. The notion of syntactic bootstrapping has been put forth, whereby the acquisition of a verb's meaning is constrained by the verb's linguistics contexts – its subcategorisation frames (Gleitman 1990) and its argument structure (Gillette, Gleitman, Gleitman and Lederer 1999). The current work is an attempt to suggest how the learner could induce subcategorisation and argument structure information. The learner uses statistics over usages that are systematically related to the underlying notion of subcategorisation frame and thematic roles, extending previous work by Brent (1993) and Allen (1997). We confirm the hypothesis by some very preliminary experiments. In the context of child acquisition, we use hierarchical clustering, a more realistic method where no training phase is available (Stevenson and Merlo 2001). The frequencies we discussed above give rise to three balanced clusters distinguishing the three original classes at 63% accuracy, without supervised training. These results thus suggest that frequencies systematically correlated to underlying abstract properties of verb classes can drive lexicon acquisition.

Acknowledgments

Some of the research reported in this paper was performed thanks to the generous contribution of the Swiss National Science Foundation, under grant 1114-065328 to the first author, and by the US National Science Foundation and the Natural Sciences and Engineering Research Council of Canada, under grants to the second author. We thank Eva Esteve Ferrer for her collaboration.

References

Allen, Joseph. 1997. “Probabilistic constraints in the acquisition of verb argument structure.” In *Proceedings of the Third Conference of Language Acquisition*, Edinburgh, Scotland.

Bod, Rens, Jennifer Hay and Stephanie Jannedy (eds.). 2003. *Probabilistic Linguistics*, Cambridge MA, MIT Press.

Brent, Michael. 1993. “From grammar to lexicon: Unsupervised learning of lexical syntax.” *Computational Linguistics* 19(2):243-262.

Bresnan, Joan, Shipra Dingare and Christopher Manning 2001. "Soft constraints mirror hard constraints." In *Proceedings of the LFG01 Conference*. 13-32. Stanford, CSLI Publications.

Brousseau, Anne-Marie and Elizabeth Ritter. 1991. "A non-unified analysis of agentive verbs." In *West Coast Conference on Formal Linguistics*, 53-64.

Cruse, David. 1972. "A note on English causatives." *Linguistic Inquiry*, 3(4):520-528.

Filip, Hana, Michael K. Tanenhaus and Gregory N. Carlson. 1998. "Reduced Relatives Judged Hard Refute Lexical Syntactic Analysis." Talk given at the *Eleventh Annual CUNY Conference on Human Sentence Processing*. New Brunswick, N. J.

Gillette, Jane, Henry Gleitman, Lila Gleitman, and Anne Lederer. 1999. "Human simulations of vocabulary learning." *Cognition* 73:135-176.

Gleitman, Lila. 1990. "The structural sources of verb meanings." *Language Acquisition* 1:3-55.

Greenberg, Joseph H. 1966. *Language Universals*. The Hague, Paris, Mouton.

Gruber, Jeffrey. 1965. *Studies in Lexical Relation*. Cambridge MA, MIT Press.

Hale, Ken and Jay Keyser. 1993. "On argument structure and the lexical representation of syntactic relations." In K. Hale and J. Keyser (eds.) *The View from Building 20*. 53-110. Cambridge MA, MIT Press.

Harley, Heidi. 1995. *Subjects, Events and Licensing*. Ph.D. dissertation, Massachusetts Institute of Technology.

Horning, James. 1969. *A study of grammatical inference*. PhD dissertation, Stanford University.

Jakobson, Roman. 1971. "Signe Zéro." In *Selected Writings*. 211-219. The Hague, Mouton, 2nd edition.

Keller, Frank. 2000. *Gradience in grammar*. PhD thesis, University of Edinburgh.

Lapata, Mirella. 1999. "Acquiring lexical generalizations from corpora: A case study for diathesis alternations." In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99)*, 397-404.

Lapata, Mirella and Chris Brew. 2004. "Verb Class Disambiguation using Informative Priors". *Computational Linguistics* 30.1:45-73.

Levin, Beth. 1985 "Introduction." In Levin (ed.) *Lexical Semantics in Review*. 1-62. Cambridge MA, Centre for Cognitive Science, MIT.

Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago, IL, University of Chicago Press.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity.* Cambridge MA, MIT Press.

MacDonald, Maryellen, Neal Pearlmutter, and Mark Seidenberg. 1994. "The lexical nature of syntactic ambiguity resolution." *Psychological Review* 89*:* 483-506.

Manning, Christopher. 2003. "Probabilistic Syntax." In Bod et al. (eds.). 289-341. Cambridge MA, MIT Press.

McCarthy, Diana. 2000. "Using semantic preferences to identify verbal participation in role switching alternations." In *Proceedings of ANLP-NAACL 2000*. 256-263.

Merlo, Paola and Suzanne Stevenson. 2001. "Automatic Verb Classification based on Statistical Distribution of Argument Structure", *Computational Linguistics*, 27.3:373-408.

Merlo, Paola, Suzanne Stevenson, Vivian Tsang and Gianluca Allaria. 2002. "A Multilingual Paradigm for Automatic Verb Classification", *Proceedings of the 40th Meeting of the Association for Computational Linguistics (ACL'02)*. 207-215. Philadelphia PA.

Moravcsik, Edith and Jessica Wirth. 1983. "Markedness - An Overview." In Fred Eckman and Edith Moravcsik and Jessica Wirth (eds.) *Markedness*. 1-3. New York NY, Plenum Press.

Pinker, Steven. 1989. *Learnability and Cognition: the Acquisition of Argument Structure.* Cambridge MA, MIT Press.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* Rutgers University Center for Cognitive Science, Technical Report 2.

Quinlan, J.Ross. 1992. *C4.5 : Programs for Machine Learning*. San Mateo, CA, Morgan Kaufmann, Series in Machine Learning.

Sabine Schulte im Walde (2000). "Clustering Verbs Semantically According to their Alternation Behaviour." *Proceedings of the 18th International Conference on Computational Linguistics*, Saarbrücken, Germany, August 2000.

Seidenberg, Mark S., Maryellen C. MacDonald and Jenny Saffran. 2002. "Does grammar start where statistics stop?" *Science* 298:553-554.

Silverstein, Michael. 1976. "Hierarchy of features and ergativity." In Robert Dixon (ed.), *Grammatical Categories in Australian Languages*. 112-171. Australian Institute of Aboriginal Studies, Canberra.

Stevenson, Suzanne. 1994. "Competition and Recency in a Hybrid Network Model of Syntactic Disambiguation" *Journal of Psycholinguistic Research* 23.4:295-322.

Stevenson, Suzanne and Paola Merlo. 1997. "Lexical structure and processing complexity." *Language and Cognitive Processes*, Vol. 12.1-2, 349-399.

Stevenson, Suzanne and Paola Merlo. 2001. "Indicator-based Learning of Argument Structure: Computational Experiments". Proceedings of the CUNY Human Sentence Processing Conference, Philadelphia PA.

Trubetzkoy, Nicolaj S. 1939. "Grundzüge der Phonologie." Prague, *Travaux du Cercle Linguistique de Prague*.

Answering with a "cleft": The role of the null subject parameter and the VP periphery

Adriana Belletti
University of Siena

1. Introduction

Why do different languages often make use of different strategies to answer the very same question concerning the subject of the clause? To what kind of analysis do the different strategies correspond? Are these strategies as different as they appear to be? These are some of the questions which I will address in this article. Without attempting at a systematic and wide overview of several languages, I will mainly focus my attention on the comparison of two quite distinct strategies that two closely related languages such as Italian and French appear to adopt in this area. Some considerations and hypotheses will also be made on the answering strategy adopted by English (and, possibly, German) in the same domain, which is different from both the one adopted by Italian and the one adopted by French. Thus, Italian, French and English appear to characteristically adopt three different strategies to answer the very same question. I assume that this provides a relatively rich spectrum of the possibilities made available through UG.

Consider the following question/answer pair in Italian:

(1) a. Chi è partito / ha parlato?
who has left/who has spoken
b. È partito / ha parlato Gianni.
has left/has spoken Gianni

There is no doubt that, if a clausal answer is provided, the respective order of the verb and the DP lexical subject is VS, with the subject in the postverbal

position[1] linearly following the verb. Consider now the equivalent question/answer pair in French:

(2) a. Qui a parlé?
b. C'est Jean.

The preferred (clausal) answer in French is (2b), namely a (reduced) cleft.

In what follows - section 2 - I will develop a proposal for an analysis of the French structures in (2b) and for the Italian-French contrast in (1), (2) which will try to uncover significant properties shared by the seemingly unrelated constructions utilized by the two languages in the same contexts (1b), (2b). In section 3 the relation between the null subject property and the availability of subject inversion will be taken up and different answering strategies concerning a "new information subject" will be discussed; a tentative proposal will also be suggested for English, in comparison with French and Italian. Section 4 will take up the "question" side of the matter, and a strategy parallel to the one at work in the (reduced) cleft answer studied here will be detected, through the analysis of *wh-in-situ* proposed in Kato (2003).

The background assumptions which will lead the analysis are spelled out in the following section.

1.1. *Assumptions: Italian "subject inversion"*

As proposed in previous work following "cartographic" guidelines[2], I assume that the low part of the clause contains a VP periphery along the lines in (3):

(3) [CP … [TP ….. [**TopP Top[FocP Foc[TopP Top**… VP]]]]]

The VP periphery is characteristically activated/made use of in so called "subject inversion" structures, where the postverbal subject typically functions as the focus of new information as in (1b) above.[3]

[1] The issue is discussed in detail in previous work of mine, e.g. Belletti (2004). On the possibly more natural answer which does not imply the repetition of the verb, see the considerations in 3.2 below.

[2] In particular Belletti (2004) and references cited there. See Cinque (2002), Rizzi (2004) for illustration of the overall cartographic approach.

[3] In appropriate discourse conditions it can also function as a topic, as discussed in Belletti (2001, 2004). See also Cecchetto (2000) on the role of this area of the clause in Right and (partly) Left dislocation.

In a null subject language like Italian, a sentence containing a postverbal subject, focus of new information as in (1b), should correspond to a (schematic) representation along the lines in (4):

(4) [CP…[TP *pro*…è…partito/ha parlato…[TopP [FocP Gianni [TopP [VP…]]]]]

I assume (much as in traditional accounts) that the relevant preverbal subject position (Cardinaletti 2004) is occupied by a non overt null *pro*. I also assume (differently form traditional accounts, Belletti (forthcoming)) a doubling derivation of "subject inversion" structures, with a *referential pro* moved from an original "big DP" filling the (relevant, Cardinaletti 2004) subject position, and the lexical subject stranded in the low focus (or topic, in different discourse conditions, footnote 3) position.

Hence, according to this analysis, the possible VS order (with no overt element in preverbal subject position) of subject inversion structures is brought about by two independent factors:

i. the null subject nature of the language;
ii. activation of the clause internal VP periphery.

The former should be considered as a necessary but not sufficient property driving "subject inversion"/VS with a silent preverbal subject; the latter is also necessary.

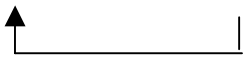
2. French

To the extent that (2a) in French is the same question as (1a) in Italian, an answer like (2b) in French should be considered the equivalent answer as (1b) in Italian. The same relevant discourse-pragmatic conditions are met in both (2b) and (1b). Let us propose that this has a direct correlate within the computational system, a fairly natural assumption, given cartographic guidelines. Let us thus assume that the computation involved in the (reduced) cleft in (2b) shares important similarities with the computation involved in the subject inversion structure (1b). I would like to propose that the two grammatical computations are indeed the same in crucial respects, modulo independent (parametric) differences between the two languages. For those aspects related to the informational content, the (reduced) cleft in French can be considered as an "inversion" structure in disguise. The proposal can be phrased as in (5):

(5) (Reduced) Clefts exploit the (low, clause internal) VP periphery.

In particular, assume that the new information subject fills the same VP-peripheral new information Focus position, exactly as the postverbal subject in (1b), (4) in Italian. This ensures the desired interpretation of the subject. However, in order for the VP periphery to be made use of in a way compatible with the negative setting of the Null subject parameter, French has to resort to some strategy: insertion of a "dummy" copula is the adopted option. The copula makes available the VP periphery which can ultimately be made use of to host the new information postverbal subject. The subject in turn is originally merged as the subject of the small clause complement of the copula[4] whose predicate, I will assume, is a relative clause containing the same verb present in the question. The (relevant) preverbal subject position of the main clause is filled by the dummy subject pronoun *ce* in compliance with the requirements of the French setting of the null subject parameter, the thematic properties of the copula, and the EPP. The reduced version of the cleft is obtained through deletion of the relative clause predicate.[5] The essential steps of the derivation are schematized in (6) (V-raising of the copula is also assumed):

(6) Ce…. [Top [Foc [Top [$_{VP}$ être [sc Jean [$_{CP}$ ~~qui a parlé~~]]]]]]



The fundamental insight of the proposal in (5)-(6) is that answering with a (reduced) cleft to a question on the subject is the strategy adopted in French to properly activate the VP periphery in a way compatible with the non null subject nature of the language.[6]

[4] Moro (1997); Burzio (1984).

[5] Speakers' judgements vary as to the extent to which deletion of the relative clause predicate is required (yielding a sentence like (2b)), or just preferred.

[6] Note that, interestingly, in a language which disposes of a (new information) Focus marker, such as Somali, the focus particle has been analyzed as derived from an original copula + 3rd person (clitic) pronoun, i.e. the structure of a cleft, as assumed here (from Fascarelli and Puglielli (to appear)):

(i) *ak + y+ aa(>*ayaa*)
be 3sgm pres

2.1. *No "subject inversion" answers in French*

Consider the exchange in (7), parallel to (1) in Italian:

(7) a. Qui a parlé?
b.*A parlé Jean.

(7b) is ruled out by the non-null subject nature of French, which I interpret as the impossibility of licensing a *pro* in the relevant preverbal subject position, neither a referential one, as in the "doubling analysis" of inversion structures in null-subject Italian referred to above (Belletti (forthcoming)), nor an expletive one.[7] On the other hand, (8), containing the overt pronominal expletive *il*, is also excluded:

(8) * Il a parlé la maman.
it has spoken the mother

(8) should be considered underivable. A doubling derivation involving a (overt) pronominal expletive could not be available, as expletives are not freely available in the original "big DP",[8] nor are they freely added to the initial numeration.

On the other hand, the doubling answer in (9a) should be considered impossible for the same reason(s) the parallel answer is impossible in Italian[9]:

[7] As in the possibly limited option in other languages (e.g. Brazilian Portuguese could be an example). See the discussion in Belletti (forthcoming).

[8] Possibly, only with indefinite DPs in the typical case. Cfr. French *il* constructions and the possibly limited option available in BP referred to in the preceding footnote; in both instances the core cases involve unaccusative verbs and no verbal agreement with the postverbal lexical subject. See Belletti (forthcoming) for the idea that the expletive could be seen as a non-agreeing pronominal within the "big DP". See Nicolis (2005) for related conclusions based on data in Icelandic TEC.

[9] In Belletti (forthcoming) the impossibility is attributed to the violation of a "discourse" constraint. As is discussed there, the constraint is assumed not to affect the doubling construction when a null *pro* is involved, thus allowing for the doubling derivation of subject inversion; the constraint does not affect this style of derivation if all elements of the original big DP are pronominal either, as discussed in the work quoted:

(i) a. Je viendrai moi.
I will come me.
b. Lui verrà lui.
he will come he.

(9) a.*Elle a parlé la maman
b.*Lei ha parlato la mamma
she has spoken the mother

Much as in Italian, however, if the lexical noun phrase is associated with a topic/known interpretation, and only in this case, the same word by word sequence becomes possible:

(10) a. Lei ha parlato, la mamma.
b. Elle a parlé, la maman.
she has spoken, the mother.

I assume that the lexical subject noun phrase fills the low topic position in the VP periphery in both cases in (10) and the pronominal referential pronoun fills the (relevant) preverbal subject position.

Note, incidentally, that, if this analysis is correct, it further illustrates the possible activation of the VP peripheral area of the clause in both languages.

2.2. *"Clefts" in L2 Italian of L1-French speakers*

Circumstantial evidence in favor of the idea developed in (5)-(6) above, according to which answering with a cleft has exactly the same informational content as subject inversion in Italian in the same contexts, comes from the behavior manifested in the L2 Italian of non advanced French speakers. In an elicitation task of VS structures of the type in (1b) in Italian, originally conceived by Belletti and Leonini (2004) and administered to a group of (non advanced) L2 speakers of Italian, L2 speakers whose L1 was French systematically produced a very high percentage of cleft sentences in place of VS. Whereas a native control group produced up to 98% of postverbal subjects/VS, the group of L1-French speakers produced only 21% of the same word order in their L2 Italian answers. The by far most widely adopted answering strategy by this group, up to 70% of the cases, was a cleft sentence. Hence, in their L2 Italian these speakers manifest Transfer of what would be the preferred answering strategy in their L1 in the same context. Beside the intrinsic interest of these findings for L2 acquisition issues[10], they are also of special interest in the present context. Indeed, since the L2 speakers with L1 French typically answered with a cleft

[10] On which see the discussion in Belletti and Leonini (2004). Overall, the reader is referred to this article for detailed discussion of the experimental design, which made use of videotapes, and of some of the material reconsidered here in the perspective of the present study.

to exactly the same questions, referred to the same situations to which the native control group typically answered with VS (and zero preverbal subject), this strongly indicates that the two answering strategies should be considered much closer to each other than meets the eye. The contrast between (1) and (2) above already suggested that. The L2 experimental data provide a particular and controlled confirmation of this impression. Given the experimental setting, we are sure that the pragmatic situations where the different answers have been produced were exactly the same indeed.

3. Null Subjects, Inversion and Clefts

Pursuing with the considerations prompted by the L2 acquisition data concerning the L1-French speakers, in Belletti and Leonini (2004) it is noted that referential null subjects were utilized to a much larger extent by these L2 non advanced speakers of Italian than VS structures (in the elicited production test). Although at this stage the percentage of null subjects (70%) is lower than the one produced by the native control group in the same testing situation (95%) still it is significantly higher than VS (21%) and the difference with the control group is less sharp for null subjects, 70% vs. 95%, than it is for VS, 21% vs. 98%. This indicates a dissociation between the fundamental property of the null subject parameter, the licensing of a null referential pronominal subject, and the currently related subject inversion property. This suggests that the formal and interpretive mechanisms licensing a null subject can be (re)set on the Italian value of the parameter, but the discourse related property tends to remain on the L1 strategy. This strategy in French, I am proposing, is actually very close and makes a use of the VP periphery in a way compatible with the non-null subject value of the parameter in this language: a (reduced) cleft sentence analyzed along the lines in (6).

3.1. *Comparing with English (and German) in L2 Italian*

In the same testing situation L1-German speakers showed a similar pattern in their L2 Italian with VS produced in just 27% of the cases and Null subjects in up to 55% of the cases. In place of VS, these speakers produced the order SV, i.e. the order appropriate in their L1. Similar findings, in testing situations with a comparable pragmatics, have been reported for L1-English L2 speakers of Italian and in attrition situations.[11] Interestingly, this pattern

[11] Bennati (2002); Tsimpli et alii (2003); see also Lonzano (2004).

appears to be persistent: also at the near native level[12], L1-English L2 speakers of Italian, in the same testing situation, continue to prefer the order SV in Italian in place of VS, whereas use of null subjects is close to native.[13] As noted by Belletti, Bennati and Sorace (in progress), the intonation of the produced SV sentences in Italian is peculiar, and reproduces the intonation of the parallel answers of the L1 English:

(11) a. Who came?
b. *John* came.

See 3.3 below for a sketch of analysis of (11b).
These findings again confirm the dissociation between the two properties currently related through the null subject parameter.

I would like to speculate here that such a dissociation is not the reflex of a grammatical property. Rather it comes as a consequence of the resetting of the null subject parameter as far as availability of null referential subjects is concerned, and of the persistency of a discourse "prominent" strategy. The answering strategy of the L1 appears to remain prominent also in the L2, despite the potential "grammatical" availability of the L2 prominent strategy, which requires a positive (re)setting of the null subject parameter, as a first necessary condition. The resetting has been undertaken by the L2 speakers. As to what pertains to the discourse aspects of the answer, the following seems to hold. In the case of the French speakers discussed in 3, the VP periphery is also activated much as it is Italian, but it continues to be activated in the "French way". In the case of the English speakers (and, probably, German as well), a possibly different strategy, implying a different structural analysis valid in English, is kept active in the L2 Italian. A speculative proposal on what this strategy and its structural analysis could be will be the topic of section 3.3.

We devote the following section 3.2. to some further empirical observations on the status of the different possible answering strategies to the same question on the subject, e.g. inversion/VS and (reduced) cleft, and we take up the French/Italian comparison again. In this connection, we will develop some considerations on the status of the different strategies from the point of view of "economy".

[12] Belletti, Bennati, Sorace (in progress); Tsimpli et al. (2003).
[13] Combined with a persistent slightly higher percentage of use of overt subject pronouns, Belletti, Bennati, Sorace (in progress) for discussion; see also Sorace (2000) for findings of the same nature in attrition.

3.2. *On "economy": inversion vs. cleft; Italian vs. French*

If answering with a cleft can be analyzed along the lines illustrated in (6) for French, involving the VP periphery in a way analogous to subject inversion in Italian, answering with a cleft can be seen as a way of mimicking subject inversion. As we put it above, the (reduced) cleft is in essence an inversion structure in disguise. If this reasoning is correct, the question then arises as to why answering with a cleft is not also a real option in Italian as well. Note that no grammatical reason would rule out the reduced cleft in Italian.

(12) a. Chi ha parlato? (a.' Qui a parlé?)
b. ?? Sono/è io/Gianni (che ho/ha parlato) (b.' C'est moi/Jean)
am-I (It's me)

But (1b) is the by far preferred answer in Italian, when V is pronounced. I would like to propose that this is ultimately due to economy reasons along the following lines: since the inversion derivation in (4) is directly available in a null subject language like Italian, it is adopted as it involves less structure and, consequently, less computation, than a (reduced) cleft. In this sense, it then qualifies as a more economical option. The adoption of VS then constitutes an "economy-driven" preferred answering strategy.

As no "grammatical" reason rules out the (reduced) cleft answer in Italian, we should expect it to show up in some contexts. Indeed, if a cleft is prompted by the question, answering with a (reduced) cleft becomes a totally natural strategy in Italian as well, as expected. Consider in this perspective, the exchanges in (13):

(13) a. Chi è stato che ha parlato/a parlare?
who has been who has talked/to talk
'Who is the one who talked.'
b. È?(stato) Gianni.
(it)-has-been-Gianni.
c. Sono? (stato) io.
(it) has been me.

(13b,c) are naturally amenable to the analysis in (14), parallel to the analysis in (6) for French, modulo presence of a null (expletive) subject *pro* in the preverbal subject position, an option available in Italian, due to the positive setting of the null subject parameter in this language:

(14) pro È stato/Sono stato…[$_{Foc}$Gianni/io..[$_{VP}$ [~~se che ha/ho parlato/a parlare~~]]]

Note, as a side remark,[14] that, as indicated by the "?" in (13b,c), use of past tense over present tense is much preferred in the reduced cleft answer in Italian, differently from French.[15] Note furthermore that nominative Case on the postverbal focalized subject can be available in Italian with the same mechanism at work in inversion/VS structures in general.[16] As for French, a clear indication that the postverbal focalized subject carries accusative Case is provided by those cases where it is a strong pronoun, for which a special Case morphology exists in French. We can assume that the same Case is carried by a lexical subject in the same position and that this is due to the fact that no comparable procedure as the one at work in Italian is available in French in structures involving expletive *ce*; hence, in these structures nominative does not reach the postverbal subject which thus displays default accusative. It should be noted that this is the only Case form which could be used for the pronominal subject in these constructions anyway, as no nominative strong form for subject pronouns exists in the French paradigm; on the other hand, the focalized status of the postcopular/postverbal subject in the (reduced) cleft necessarily requires use of a strong form for the pronominal subject. Hence, default accusative is a necessary choice in the case of a postverbal focalized strong pronominal subject in French.

Consider Italian again. The following question/answer pairs can be analyzed along similar lines, with different degrees of reduction involved in the answer:

(15) a. Chi è (- alla porta)?
who is (- at the door)
b. Sono/è io/Gianni (- ~~alla porta~~).
c. Chi è(- che parla)?
Who is (-who is talking)
d. Sono/è io/Gianni(- ~~che parlo/a~~).
Am I/is Gianni
'It's me/it's Gianni'

And, with further reduction:

[14] For which an in depth analysis would take us too far afield so it will not be pursued here.

[15] For some speakers the requirement is very strong, for others less so, but all speakers tend to prefer the past option in Italian as opposed to French. It might be that this preferential judgement is less "innocent" than one might think at first sight. For the time being I will make the (possibly simplifying) assumption that the two languages share the same analysis in the domain of clefts in all respects, leaving a more detailed study for future research.

[16] See Belletti (forthcoming) for a recent proposal linking availability of postverbal nominative Case to the "doubling" analysis of inversion/VS structures.

(16) a. Chi è (- alla porta)?
b. ~~Sono/è~~ io/Gianni (- ~~alla porta~~).
c. Chi è (- che parla)?
d. ~~Sono/è~~ io/Gianni (- ~~che parlo/a~~).
I/Gianni
"Me/Gianni"

According to the analysis proposed in (16b,d), based on (14), when the answer only contains the questioned constituent, it corresponds to a reduced sentence where the questioned constituent, the subject, fills the new information focus position in the VP periphery.[17] In the cases in (16b,d) the reduction is supposed to take place within a (reduced) cleft, assuming a parallelism to hold between the question and the relative answer. Thus, the reduction in (16b,d) constitutes in essence a step further in the reduction already at work in (15b,d).

It is undoubtedly the case that the mostly reduced answers along the lines in (16b,d) are often the preferred (/most used?) answers by speakers in current conversational exchanges; certainly, they are considered extremely natural by everybody. We might speculate that, once again, economy is at stake here. Some economy conversational principle possibly drives this intuition (and use?): complete reduction of the content corresponding to the "parallel" question allows one to say the least possible in the answer, and at the same time to be the most informative and relevant in some Gricean-style terms.[18] Be it as it may, what is of special relevance here is that partly reduced answers are also possible and, in the case of questions on the

[17] The proposal shares part of the insight of the analysis presented in Brunetti (2003): reduced answers of the type in (16) are "portions" of clauses, with the subject in focus. But there is an important difference between the two approaches: here the internal focus position is assumed to be involved and a crucial relation is established with (reduced) cleft sentences, which is not the case in Brunetti (2003). No remnant movement and no involvement of the clause external peripheral focus are hypothesized here, as it is done in Brunetti's proposal. I am assuming that the external, left peripheral focus position is dedicated to contrastive/corrective focalization in standard Italian, as discussed in detail in Belletti (2004).

[18] There is no assumption here that reduced answers in Italian should necessarily involve reduction of a cleft in general. This is a natural assumption in cases like (15)-(16), but reduction would probably obtain in different structures as well, e.g. in i.cwhere the possible analysis of the answer in i.b to the question in i.a could involve reduction of the portion of clause above the VP periphery:

(i) a. Chi ha parlato?
b. Io/Gianni
c. [Ho/Ha parlato [$_{\text{Foc}}$ io/Gianni …..[$_{\text{VP}}$ …]]]

subject, they characteristically display a postcopular or, more generally, a postverbal subject, which, we assume, fills the new information position in the VP periphery.

In conclusion, we have proposed that there is no substantial difference between French and Italian in the domain of question-answer pairs on the subject, contrary to first appearance: in both languages the VP periphery is involved in the same contexts. The apparent differences between the two languages can be interpreted as due to independent factors: adoption of different preferred strategies compatible with the different parametric setting of the null subject parameter. We have proposed that preferred strategies tend to persist, as indicated by the L2 Italian data considered. We have further speculated that ultimately economy reasons condition the prevalence of a given answering strategy, e.g. VS in Italian, where the relevant grammatical principles are met (null subject). Answering with a (reduced) cleft, however, is an available strategy in both languages, as one would expect.

3.3. *Speculations on English*

When *be* (a (reduced) cleft?) is prompted by the question, English, much as Italian, allows for an answer utilizing a (reduced) cleft. This is illustrated in the possible analysis (17b) of the answer to (17a), which parallels (6) above, with the postcopular subject sitting in the Focus position in the VP periphery:

(17) a. Who is knocking at the door?
b. It's[$_{\text{Foc}}$ me/John … [$_{\text{VP}}$[$_{\text{SC}}$ ~~-knocking at the door~~]]..]

However, a (reduced) cleft is not the preferred answering strategy in English to a question on the subject not containing *be*. Much as in Italian where inversion/VS is preferred over a (reduced) cleft when V is pronounced in the answer to a question on the subject, the order SV is typically displayed in English, with a peculiar stress on S. The question-answer pair in (11) repeated below, illustrates the point. Note that (11b) differs from (18), where, I assume, the subject is contrastively focused in the clause external left periphery:

(11) a. Who came?
b. *John* came.

(18) JOHN came (not Bill).

Use of a (reduced) cleft appears to be more limited in English than it is in French in the most typical case. Answering with a cleft, however, seems to be more readily available in English than it is in Italian, as English speakers find exchanges like the following relatively possible[19]:

(19) a. Who came?
b. It is/was John/me.

On the other hand, it should be noted that for some French speakers the following exchange, totally parallel to the English (11), can also be admitted to some extent (although it is clearly not preferred):

(20) a. Qui a parlé?
b. (*)*Jean* a parlé.

This is not surprising as no principled grammatical reasons should rule out the (favored) English strategy, to a possible analysis of which we turn momentarily, as a possible option in French.[20] On the contrary, it should be noted that no (favored) Italian strategy displaying the order VS could be available in English, for principled grammatical reasons: inversion structures of the Italian kind are excluded in English along similar lines discussed for French, ultimately due to the negative setting of the null subject parameter in English[21].

Let us now turn to a speculation on what the English strategy should amount to. I would like to tentatively suggest that it is a case of "focus *in*

[19] As already observed above for French, reduction/deletion of the cleft is preferred; i.e. *It was John who came* is perceived as somewhat odd.

[20]Indeed, some Italian speakers do not totally rule out an equivalent exchange in Italian as well:

(i) a. Chi ha parlato?
b. (*) *Gianni* ha parlato.

However, everybody agrees that VS is the by far preferred option in standard Italian when V is pronounced in the answer and the pragmatics of a pure question of information is ensured by the context. Cfr. also the consistent behavior of the native control group in the experimental results discussed in 2.2., 3.1. Once again, as no grammatical, formal principle is violated in (ib), it is not surprising that the sentence can be not completely ruled out by some speakers who resort to the answering strategy which is in fact the preferred one in English and as such manifests a UG option in this area.

[21] *There* constructions should be amenable to a partly different analysis along the lines in Moro (1997); I will not take up the issue here.

situ" which could be interpreted as involving DP-internal focalization of the same type overtly manifested in (21). The peculiar intonation necessarily associated with the preverbal subject in these contexts (11b), should be seen as the reflex of a "null" DP internally focalized *himself*, i.e. ultimately an "activated" DP-internal focus position:

(21) John himself came

Whatever the better analysis should turn out to be for answers like (11b), focalization in the answer must be obtained through an answering strategy compatible with the negative setting of the null subject parameter of English. Focus *in situ* would have this property. On the possible reasons why a (reduced) cleft answer to questions on the subject happens to be overwhelmingly active only in French (and not in English), see the tentative discussion in 4 below. Of course, the reason should identify a matter of preference, as both the English like "focus *in situ*" strategy and the French like (reduced) cleft strategy are compatible with the negative setting of the null subject parameter in these languages.

Note that the order SV, with the peculiar stress on S just described for English, cannot be amenable to the same analysis recently proposed by Cruschina (2004) for the same word order in the same discourse conditions also found in some Sicilian varieties (Cruschina 2004, ex.(5)):

(22) a. Cu partì?
Who left
b. *Salvo* partì.
Salvo left

As Cruschina (2004, ex. (3)) shows, the clause initial location of the new information focus is not limited to a subject: cfr. (23a,b) for a question-answer pair concerning the direct object and the contrast with (24a,b) which shows that the same possibility is not readily available in English:

(23) a. Chi scrivisti airi?
what wrote (you) yesterday
b. *N'articulu* scrissi
an article (I) wrote

(24) a. What have you written yesterday?
b. %*A paper* I have written
c. I have written a paper

It seems that, in the most typical case, the clause initial, left peripheral focus position is not appropriate for purely new information focus constituents in English, much as it happens in standard Italian[22]. Some extra feature must be involved, such as contrast, correction, etc. This is not the case in the Sicilian variety discussed by Cruschina (2004), where the left peripheral focus position can equally host both contrastive/corrective and new information focus. This in turn implies that the peculiarly stressed preverbal subject in the English answer in (11b) cannot be taken to fill the same left peripheral focus position as contrastive focus; whence our (tentative) proposal that answers like (11b) are an instance of (DP-internal) "focus *in situ*".[23]

4. *Wh-in-situ* in the VP periphery: the "question" side

Let us now turn to what we might call "the question" side of the matter. It is tempting to claim that the low VP periphery be made use of also in questions. This proposal has been put forth in Kato (2003) who has proposed that wh-in-situ exploits the focus position in the VP periphery in the way illustrated in (25b) for Brazilian Portuguese (BP):

(25) a. Você viu quem?
you saw who
b. $[_{CP} Q [_{IP}$ você viu $[_{FP}$ $\underline{\text{quem}}_q$ $[\varnothing_{+wh}[_{vP} t_i t_v[_{VP}$ $t_v t_q$]]]]]]
you saw who

[22] As I discuss at some length in the previous work of mine quoted above. In the well formed (24c) the direct object should fill the VP peripheral new information focus position, similarly to Italian in the same context. See (30) below.

[23] Thus, languages vary as to what kind of discourse related features they associate with the different focus positions assumed, the clause external left peripheral one and the clause internal one, in the VP periphery. English and Italian are on a par in this case, with the left peripheral focus (typically) limited to contrast/correction and the clause internal focus in the VP periphery, dedicated to new information focus. In Sicilian, the clause internal VP periphery is not "visible" as far as focalization is concerned, with both contrastive and new information focus located in the left periphery, as Cruschina (2004) shows. We might speculate that French could be taken to be the mirror image of Sicilian: in French the clause external left periphery is not activated for focalization of any kind (**JEAN j'ai rencontré*/JOHN I have met), while the clause internal VP periphery is, if our analysis of cleft sentences with a new information subject is on the right track. Since a cleft can also be contrastively used in French, the same clause internal focus position could be made available to contrastive interpretation as well, thus giving rise to the mirror image of Sicilian in this respect. Thanks to G. Bocci for raising this point. See also Bocci (2004) for relevant discussion on clause external, left peripheral focalization in standard Italian.

It is natural to assume that, if the analysis in (25b) holds for BP, the same analysis should also extend to other *wh-in-situ* languages which manifest similar properties in the syntax of *wh-in-situ* constructions; e.g. French:

(26) a. T'as fait quoi?
you have done what
b. Il a vu qui?
he has seen who

In (25b) [Q] and [wh] are the features assumed to be always involved in *wh*-questions: the [Q] feature is located in CP where it expresses the force of the clause, the [wh] feature is located in the focus position of the VP periphery. In the spirit of Miyagawa (2001), the two features can be assumed to be both located in the CP area in languages not allowing *wh-in-situ* (e.g. Italian, English, etc., with movement targeting the external focus position, Rizzi 1997). In Japanese, both positions are overtly realized, the external C[Q] (Force) and the internal [+wh]:

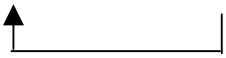
(27) Biru-wa John-ninani-o aguemashita-ka/no?
Bill-topic J-dat what-acc gave-*Q* (*Q*=question)

The only difference between (27) and (25) is to be recognized in the non-overt realization of the [Q] feature in the latter.

This analysis of *wh-in-situ* suggests the following speculative remarks. We have begun our discussion by noticing the widespread use of (reduced) clefts in answering questions on the subject in French. In the course of the analysis, we have mentioned that SV, with the English type peculiar stress on S, is also available, although to a lesser extent, in French as well. We might now ask why (reduced) clefts should take such a clear priority in French. I would like to suggest that this is favored by the combination of two factors: (wide) availability of cleft sentences to express new information focus and (wide) availability of the *wh-in-situ* strategy in question formation, analyzed along the lines in (25b). In both the cleft sentence and the *wh-in-situ* computation, the VP periphery is crucially involved in a parallel fashion. See (28), where both the *wh*-phrase in the question and the subject *Jean/moi* in the answer should fill the same VP peripheral focus position combining (25b) and (6) above:

(28) a. C'est qui qui a parlé?
who is it who has spoken

a.' [$_{CP}Q$ [$_{IP}$ c' est[$_{FP}$[$\varnothing_{+wh}$[$_{vP}$ qui [qui a parlé]]]]]]



b. C'est Jean/moi
it is Jean/me

In the case of questions on the subject, a (reduced) cleft appears to be a most natural answer in French, and the (reduced) cleft takes priority also in the lack of a strict parallelism with the question, i.e. also if question is not formulated with a cleft and *wh-in-situ* (as it was the case in (2a)). However, when the question concerns the object, native speakers do not find a (reduced) cleft to be such a natural answer in French anymore, unless the question is also expressed through a cleft with an *in-situ wh*-object. The contrast in (29) illustrates the point:

(29) a. C'est quoi que t'as lu?
is it what that you have read

b. C'est un roman
it is a novel

c. Qu'est ce que t'as lu?/ Qu'as-tu lu ?
what have you read

d.*?C'est un livre
it is a book

Indeed, French appears to behave in this case in the same way as other languages such as Italian and English. It thus seems that, in the case of the subject, the (reduced) cleft is the only way available in French to make the discourse appropriate use of the VP periphery, in a way compatible with the negative setting of the null subject parameter. Somehow this strategy is less costly than an English like (DP internal) "focus *in situ*" strategy, possibly due to the wide exploitation of clefts in French in general. In the case the object, however, French, similarly to English and Italian limits the use of an answer with a (reduced) cleft to those cases where parallelism with the question, also containing a cleft and *wh-in-situ*, is manifested[24]. Otherwise, for reasons likely to be due to computational economy, I assume that the new information object is directly located in the dedicated new information focus position in the VP periphery, yielding SVO, if V is pronounced, in all

[24] We are claiming that the parallelism in the question-answer pair is neater in French than it is in English or Italian as these latter languages are not wh-in-situ languages as French is.

three languages. Consider, in this respect, the completely analogous behavior of French, Italian, and English in question answer pairs involving the object illustrated in (30):

(30) a. Qu'est ce que t'as lu?/ Qu'as-tu lu ?
what have you read
b. J'ai lu un roman
I have read a novel
c. Cosa hai letto?
what have you read
d. Ho letto un romanzo
(I have read) a novel
e. What have you read?
f. I have read a novel

The lack of contrast in (30) stands in sharp contrast with the very different behavior manifested by the three languages in question answer pairs on the subject, which started out our discussion.

5. Summary and some conclusions

We can now highlight some of the main points of our discussion and draw the, somewhat preliminary, conclusions, listed below.

a) Languages appear to differ as to the strategy they preferably adopt in answering questions on the subject. All other things being equal, null subject languages, as standard Italian, can exploit the inversion strategy making use of the dedicated focus position in the VP periphery, with a silent preverbal subject[25]. In languages where the null subject parameter is set negatively, one of two different strategies appear to be utilized to realize a new information subject: i. creating a non-null subject compatible "inversion" structure, exploiting the informational content of the VP periphery; ii. adopting a (DP internal) focus-in-situ strategy. French and English illustrate the two options. The quite typical (reduced) cleft answer, widely adopted in French, has been interpreted here as a kind of "inversion in disguise" sharing important properties with subject inversion/VS structures of the Italian type.

[25] But we have seen that inversion is not the only option for a null subject language, as it is suggested by the Sicilian case where the clause external left periphery is also compatible with simple new information focus (Cruschina 2004).

b) It has also been suggested that (wide) availability of *wh-in-situ* and cleft sentences in French might be at the source of the prominent status of clefts in answering questions on the subject. For reasons ultimately due to computational economy, answers containing a new information object, on the other hand,tend to limit use of a (reduced) cleft to those cases where strict parallelism with the question holds, where the question also contains a cleft and a *in-situ wh*-object.
c) Answering with a (reduced) cleft to questions on the subject, appears to be a wider option than might appear at first sight, shared by the different languages considered. The VP periphery is systematically made use of in a way which parallels use of this area of the clause in subject inversion/VS structures.
d) The empirical evidence which is able to reveal deep similarities across languages can be (and should be, we might add) of different nature. In our discussion, we have utilized L2 acquisition experimental data and we have seen that they provide a peculiar magnifier, able to capture deep similarities at the discourse/pragmatics interface: answering to a question of information must be essentially the same process across languages. We have tried to make explicit how this can be expressed within the clause structure assumed here.

References

Belletti, Adriana. 2004. "Aspects of the low IP area". In L. Rizzi (ed.) *The structure of CP and IP. The Cartography of Syntactic Structures, Vol. 3*. 16-51. New York/Oxford, Oxford University Press.

Belletti, Adriana. Forthcoming. "Extended Doubling and the VP periphery". To appear in *Probus* 17.1.

Belletti, Adriana and Chiara Leonini. 2003. "Subject inversion in L2 Italian". In S. Foster-Cohen, M. Sharwood Smith, A. Sorace, M. Ota (eds.) *Eurosla Yearbook*. 95-118. Amsterdam/Philadelphia, John Benjamins.

Belletti, Adriana, Elisa Bennati and Antonella Sorace. In progress. "Revisiting the Null subject parameter from an L2 developmental perspective: Pronominal subjects and postverbal subjects in adult L2 near native Italian".

Bennati, Elisa. 2002. *Soggetti nulli e soggetti postverbali nell'italiano L2 di parlanti adulti di madrelingua inglese*. Tesi di Laurea. University of Siena.

Bocci, Giuliano. 2004. *Aspetti teorici e analisi sperimentale della focalizzazione contrastiva nella periferia sinistra della frase: sintassi, fonologia e fonetica*. Tesi di Laurea, University of Siena.

Brunetti, Lisa. 2003. *A Unification of Focus*. Ph.D.dissertation. University of Florence.

Burzio, Luigi. 1986. *Italian Syntax*. Dordrecht, Reidel.

Cardinaletti, Anna. 2004. "Towards a Cartography of Subject Positions". In L. Rizzi (ed.) *The structure of CP and IP. The Cartography of Syntactic Structures, vol. 2*. 115-165. New York/Oxford, Oxford University Press.

Cecchetto, Carlo. 2000. "Doubling Structures and Reconstruction". *Probus* 12, 1: 93-126.

Cinque, Guglielmo (ed.). 2002. *Functional Structure in DP and IP. The Cartography of Syntactic Structures, Vol.1*, New York/Oxford, Oxford University Press.

Cruschina, Silvio. 2004. *Il Focus di nuova informazione e la periferia sinistra: la struttura informativa della frase in Siciliano*. Tesi di Laurea Specialistica. University of Siena.

Frascarelli, Mara and Anna Rita Puglielli. To appear. "Focus Markers and Universal Grammar". In M. Marten (ed.) *Proceedings of the 4th International Conference of Cushitic and Omotic Languages*. Rüdiger Köppe Verlag.

Kato, Mary Aizawa. 2003. "The interpretation and derivation of *wh*-in-situ constructions in Brazilian Portuguese". Ms., UNICAMP.

Lonzano, Cristóbal. 2004. "Focus and Split-intransitivity: The acquisition of word order alternations in non-native Spanish". Ms., University of Essex.

Miyagawa, Shigeru. 2001. "The EPP, Scrambling, and wh-in-situ". In M.Kenstowicz (ed.) *Ken Hale: A Life in Language.* 293-338. Cambridge MA, MIT Press.

Moro, Andrea. 1997. *The Raising of Predicates*. Cambridge, Cambridge University Press.

Nicolis, Marco. 2005. *On pro-drop*. Ph.D.dissertation, University of Siena.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In L. Haegeman (ed.) *Elements of Grammar*. 281-337. Dordrecht, Kluwer.

Rizzi, Luigi (ed.). 2004. *The structure of CP and IP. The Cartography of Syntactic Structures, Vol. 2*, New York/Oxford, Oxford University Press.

Ronat, Mitsou. 1979. "Pronoms topiques et pronoms distinctifs". *Langue Française*, 44:106-128.

Sorace, Antonella. 2000. "Differential effects of attrition in the L1 syntax of L2 near native speakers". *BUCLD Proceedings 24*, Cascadilla Press.

Tsimpli, Ianthi-Maria, Antonella Sorace, Caroline Heycock and Francesca Filiaci. 2003. "First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English". Ms., University of Edinburgh.

Classifier incorporation and the locative alternation

Federico Damonte
University of Padua

1. Introduction[*]

In this paper I propose a first approximation towards a *derivational* analysis of the *locative alternation*[1], illustrated for English in (1) with the verb *load*:

(1) a. I loaded the sand on the truck
b. I loaded the truck with sand

More precisely, I will try to show that both variants in (1) share the same initial syntactic structure, and that (1a) more closely corresponds to that structure, while the derivation of (1b) involves additional syntactic processes. I will also argue that a derivational approach can account for a well-known semantic difference between the two variants, the so-called *holistic effect*, namely the fact that (1b) entails that the truck is completely filled with sand, while there is no such entailment in (1a).

What makes a derivational approach to argument alternations attractive is the fact that it makes it possible to restrict initial syntactic structures to those projected by rigid linking principles like the UTAH (Baker 1988).

[*] I wish to thank Alberto Mioni, Marcel den Dikken, Andrea Padovan, Cecilia Poletto, Gemma Rigau, Walter Schweikert and Karen Zagona for discussing with me many issues involved in the analysis proposed here. Thanks also to Andrea Padovan, Mark de Vos and Walter Schweikert for providing me with their judgements about the Italian, English and German examples, respectively. As usual, all mistakes are my own.

[1] The literature on the locative alternation is huge. For a more comprehensive bibliography, the reader is referred to the references in the works quoted in this article, and especially Rappaport and Levin (1988) and Levin (1993). Derivational analyses of the locative alternation were proposed by Hall (1965) and Larson (1990).

Furthermore, if the syntactic processes by which (1b) is derived are the same as those involved in the derivation of other argument alternations, such as the dative alternation (*I gave the book to John*/*I gave John the book*) then we do not need any special mechanism, such as ad hoc linking rules, to account for the fact that some predicates can realize their arguments in more than one way. Finally, if we understand which syntactic mechanism allows a given predicate to realize its arguments in more than one syntactic configuration, we can perhaps also account for why the same predicate in other languages does *not* have this property.

Most modern analyses of the locative alternation, though, follow a lexicalist approach, that is to say, one in which the difference between the two variants is located in the lexicon, not in the syntax. The main motivation for this approach lies in the afore-mentioned semantic difference between the two variants. As argued by Baker (1997), a difference in interpretation between the variants of an alternating predicate means that the alternation cannot be syntactically derived, on the assumption that syntax cannot extend the basic meaning of a predicate. A lexicalist account would then argue that the semantic difference between the two variants is due to the fact that they actually correspond to two different predicates. For instance, Rappaport and Levin (1988) argue that *load* in (1a) is a motion verb, while in (1b) is a change of state predicate, which affects the Locative argument (the truck, in this case) by means of loading it, thus accounting for the holistic effect[2]. The two predicates are then projected onto syntax independently, through highly specific linking rules.

In the minimalistic revision of the theory (Chomsky, 1993), though, intermediate levels of representation such as D-structure are abolished, which in turn means that LF representations of meaning cannot be compared with any other structure than lexical representations (Karen Zagona, pc). Thus, nothing in principle prevents syntax to add specific components of meaning, as long as these are compatible with the core lexical meaning of the predicate. This is the approach started by the seminal work of Hoekstra (1992), and the one that will be adopted here, with some differences.[3] More precisely, I will argue that syntax can produce the interpretation associated with (1b) through the *incorporation* onto the verb of an abstract noun, the

[2] On the further assumption that in order to "affect" a locative container complement like a truck you have to *completely* fill it (Jackendoff 1990, 172).

[3] The main difference being that in the in the line of research started by Hoekstra (1992), Borer (1994), among others, there is no projection from the lexicon to syntax, the meaning of the verb, including its categorial label, being entirely constructed in the syntax. See Levin and Rappaport Hovav (1996, 61 - 65) for some discussion of this "constructivist" approach. I will instead assume that theta roles are projected onto syntax, and that this projection is restricted by UTAH.

classifier of the Theme. In this paper I will focus on the syntactic processes involved in the derivation of (1b), and I will therefore stick to a single predicate, *load*, but I will also consider the systematic difference between Romance and Germanic languages, namely the fact that (western) Romance languages allow yet another variant, in which the Theme is introduced by the preposition *of*, as shown in the following Italian examples:

(2) a. Ho caricato la sabbia sul camion
have-1sg loaded the sand on-the truck
'I have loaded the sand on the truck'
b. Ho caricato il camion con la sabbia
have-1sg loaded the truck with the sand
'I have loaded the truck with the sand'
c. Ho caricato il camion di sabbia
have-1sg loaded the truck of sand
'I have loaded the truck with sand'[4]

This paper is organized as follows: in section 1 I will briefly survey the distribution of the locative alternation cross-linguistically and show that the verb meaning *load* does not alternate in all languages. In section 2 I will argue that the incorporation onto the verb of an abstract preposition[5] allows the Locative argument to become the direct object of the verb. In section 3 I provide an analysis of the *with*-phrase that tries to account for the fact that it seems to have both adjunct and argument properties. In section 4 I will deal with the possibility shown by Romance languages to realize the Theme in the derived variant with an *of*-phrase. In that section I will try to show that the holistic effect is restricted to unbounded Themes and it is triggered by the incorporation onto the verb of the classifier of the Theme. Finally, section 5 provides the conclusions.

2. The locative alternation: a brief cross-linguistic overview

As briefly mentioned in the introduction, lexicalist approaches account for the existence of the locative alternation by assuming that the two variants correspond to two different lexical representations, named by the same verb. While these approaches correctly account for the semantic relationship

[4] From now on, for ease of exposition, I will refer to (2a) as the *basic* variant, and to (2b-c) as the *derived* variants. I will further call (2b) the *with*-variant and (2c) the *of*-variant. The arguments of the verb will be labelled "Theme" and "Locative", and the PPs in (3b-c) the *with*-phrase and the *of*-phrase, respectively.

[5] As proposed by Wunderlich (1987) and Brinkmann (1997), among others, in a lexicalist framework.

between the two variants, they do not say much about the actual syntactic realization of the relevant lexical representation. In particular, they seem rather indifferent to the complex distribution of the locative alternation across languages. Even if we stick to the verb *load*, we see that languages vary as to whether this predicate alternates or not. As shown by the cross-linguistic survey carried out by Kim (1999), in some languages (English, Italian, Hebrew, Malay) this verb alternates; in other languages, only the basic variant is possible (Chinese, Thai, Japanese and Korean), but there is no language in Kim's sample in which the derived variant is the only possible syntactic realization of the verb meaning *load*. Furthermore, the set of languages in which *load* alternates splits into two groups: those languages in which the Theme, in the derived variant, is introduced by the preposition *with*, and those in which it can be also introduced by the preposition meaning *of*. Finally, if we look at other verbs of the load/spray class, not all verbs in this class behave in the same way in all languages: in some languages, like Hebrew and Turkish, *load* alternates, while *spray* only allows the basic variant; viceversa, in Chinese and Thai, *spray* alternates while *load* only allows the basic variant (Kim 1999, 145). The reader is referred to Kim's work for data about other classes of locative verbs.

This uneven cross-linguistic distribution of the locative alternation seems to require some additional explanation besides a purely lexical-semantic one: clearly, a theory in which meaning alone determines whether a predicate alternates or not, implicitly assumes the same verb to have the same behaviour in all languages (Pinker 1989, 67). Thus an important gap in the distribution of the variants of the locative alternation is left unexplained: no language has been as yet reported in which the predicate meaning *load* appears only in the derived variant.

3. Prefixation and the locative alternation

I will propose that the distribution of the locative alternation discussed in the previous section can be accounted for if we assume that the basic variant directly corresponds to the initial syntactic structure generated by UTAH[6] and that the derived variant is possible only if the language has the (morpho) syntactic devices necessary to generate it. This way, the fact that in no language the derived variant is the only means to realize the arguments of the verb *load* is accounted for: since this variant is syntactically derived from the basic one, whenever the derived variant is possible, the basic variant will

[6] Modulo, of course, further movements of the verb to check its features and of the Theme to get case.

be possible as well.[7] In the rest of this paper I will try to define more precisely these (morpho)syntactic mechanisms. In this section I will discuss the promotion of the Locative argument to direct object of the verb, in the next two sections I will examine the syntactic status of the PP in the derived variants.

Following several studies about the locative alternation (Wunderlich 1987, Brinkmann 1997, among others), I will assume that the promotion to object of the Locative argument involves the incorporation of a locative preposition onto the verb. The connection between prefixation and the derived variant of the locative alternation is clearly visible in all Germanic languages, excluding English, where prefixation is still productive and semantically transparent.[8] In German, for instance, the verb in the basic variant is unprefixed while it bears a prefix in the derived variant[9]:

(3) a. Ich lud Heu auf den Lastwagen
I loaded hay on the truck
'I loaded hay on the truck'
b. Ich belud den Lastwagen mit Heu
I loaded the truck with hay
'I loaded the truck with hay'

Intuitively, the preposition in the basic variant alternates with the prefix in the derived variant, and the incorporation analysis captures this intuition straightforwardly: the prefix *be-* is generated as a preposition but it incorporates onto the verb. This analysis has been explicitly proposed by, among others, Gronemeyer (1995) for Swedish, and by Brinkmann (1997) for German.[10] I will bring this idea a step forward, and argue that the incorporation of the preposition is a necessary condition to build the derived variant in the syntax and that it therefore takes places even in languages with less productive prefixation than Germanic languages. More precisely, in case the language does not allow overt prefixation, an *abstract* preposition incorporates. Let us see now how the derivation works in detail.

[7] As pointed out to me by Marcel den Dikken, this only holds if two further assumptions are made: 1) all languages with prepositions have an overt preposition meaning *on*, 2) this overt preposition is selected by the verb meaning *load*.

[8] See, among others, Brinkmann (1997) on German, Gronemeyer (1995) on Swedish, Mulder (1992) on Dutch.

[9] Actually, this is a semplification of the facts: while no prefix is possible in the basic variant (3a), the prefix is optional in the derived variant (3b). See Brinkmann (1997, 68ff) for extensive discussion.

[10] But note that in Brinkmann's framework this incorporation takes place in the lexicon.

Consider first the derivation of the basic variant. Following Hoekstra and Mulder (1990), and Mulder (1992), I will assume that both arguments of the verb *load* are generated in a complement *small clause*. The underlying structure of the basic variant (1a=4a) would therefore be (4b):

(4) a. I loaded the sand on the truck
b. [VP load [SC the sand on the truck]]

Opinions differ as to the internal structure of a small clause, and while these differences do not bear directly on the hypothesis under discussion, I will assume, for ease of exposition, that small clauses have a head, that the locative preposition is the head of the small clause in (4b), and that the Locative argument is the complement of this head while the Theme occupies its specifier position, as shown in (5)[11]:

(5) [VP load [SC/PP the sand [on the truck]]

I will further assume that this is the only possible projection of the arguments of the verb *load* allowed by UTAH. In the case of the basic variant, no special mechanism is necessary to derive it from the structure in (5), either in Italian or English.

On the contrary, in order to obtain from (5) the surface structure of the derived variant, we have to account for the fact that the Locative argument "jumps over" the Theme without causing a shortest move violation. First, following standard assumptions, I will assume that the Theme moves to [Spec, AgrO] to get structural case. Then, following den Dikken's (1995) analysis of the dative alternation in English, I will hypothesize that the verb *load* can licence an empty preposition:

(6) [VP load [SC sand P the truck]]

The empty preposition, in order to be licensed, incorporates by head movement onto the verb:

(7) [VP P-load [SC sand t the truck]]

This makes [Spec, VP] *equidistant* (Chomsky (1993)) to [Spec, SC] and therefore a possible landing site for the Locative argument, which can now

[11] Note that I am assuming that the label "Theme" can be defined on the basis of the meaning of the verb alone, independently of the aspectual make-up of the event. More precisely, in section 4.1 below I will argue that the Theme does not delimit the event in the *of*-variant.

move over the Theme in [Spec, SC] without causing a shortest movement violation.[12] From there the Locative argument can move to [Spec, AgrO] to get structural case, subsequent movement of the verb yielding the surface word order verb - Locative:

(8) [AgrO the truck$_i$ [VP t_i Pj-load [sand t_j t_i]]]

The analysis proposed here leads us to expect that in languages with productive prefixation locative prepositions can surface as prefixes on the verb, which seems to be true for German, at least for the "path" prefixes *um*-, *hinter*-, *durch*- etc.:

(9) a. Das Schiff segelte *um* Kap Horn (Brinkmann 1997, 82)
b. Das Schiff *um*segelte Kap Horn
the ship (around)sailed around Cape Horn
'The ship sailed around Cape Horn'[13]

More generally, the hypothesis under discussion makes two very specific predictions about the correlation between *locative* prefixes and the variants of the locative alternation:[14]

(10) i. Spray/load verbs are prefixed, either overtly or covertly, in the variants where the location argument is the direct object.
ii. All overtly prefixed "verbs of putting" (Levin 1993, 111) do not alternate and only have the option of realizing the location argument as direct object of the verb.

The correlation between prefixation and the ability of a predicate to alternate seems to hold in a number of languages: for an extensive discussion of German, see Brinkmann (1997, 76 - 89) and the works quoted in fn.8 for the other Germanic languages. For English, Brinkmann (1997, 65) notes that almost all of the prefixed locative verbs discussed by Pinker (1989) do *not*

[12] As pointed out to me by Gemma Rigau, this means that the external argument is not merged in [Spec, VP] but in a different position, quite probably [Spec, vP].
[13] On the other hand, not all prefixes might correspond to overt prepositions. This seems to be the case of German *be*-, which does not correspond to *bei*, as discussed extensively by Brinkmann (1997, 76 - 82), see also den Dikken (1995, chapter 5).
[14] Aspectual prefixes like *over*- (as in *overload*) do not follow under (10), as they cannot be licensed as the head of the prepositional small clause. As for the fact that *overload* (and *sovraccaricare* in Italian) only appears in the derived variant, I will tentatively assume that aspectual prefixes require the "perfectivising" aspect introduced by the incorporation of the preposition.

alternate: only two out of 28 prefixed verbs alternate (*inject* and *bestrew*). For a similar conclusion about Italian, see Munaro (1994).

Before leaving this section it is important to point out that the analysis proposed here for the promotion of the Locative argument to direct object is largely compatible with that put forward by den Dikken (1995) for the "Goal" argument in the dative alternation, which in turn shows that the syntactic mechanisms proposed have indeed a general validity.[15]

4. The Theme in the *with*-variant

The main problem for a UTAH-based derivational analysis of the locative alternation is the fact that the Theme in the derived variant surfaces as a prepositional phrase. Furthermore, as noted in the introduction, some languages can use two different prepositions, *with* and *of*, to introduce the Theme. In this section I will only consider the *with*-variant, in the next I will try to account for the possibility, in Romance languages, to realize the Theme with an *of*-phrase as well.

Before proposing a hypothesis about the syntactic process that "demotes" the direct object to a PP, let us define more precisely the semantic and syntactic properties of the *with*-phrase. Semantically, the PP in the derived variant is still the Theme, and therefore an argument, of the predicate. Yet, there are several facts that seem to point to an adjunct status of the *with*-phrase. First, it can be omitted (11), while this is not possible in the basic variant (12):

(11) a. Ho caricato il camion
have-1sg loaded the truck
'I have loaded the truck'
b. I loaded the truck

(12) a.*Ho caricato sul camion
have-1sg loaded on-the truck
'I Have loaded on the truck'
b.*I loaded on the truck

Second, the *with*-phrase seems incompatible with an instrumental adjunct, as pointed out by Mateu (2000, 33), while such an adjunct is possible in the

[15] The main difference with den Dikken's account being that the Goal argument does not move to [Spec, AgrOP], but to the specifier of a head below the verb. Note also that den Dikken argues that verbal prefixes in Germanic languages are particles, not transitive prepositions.

basic variant, which seems to indicate that the *with*-phrase is an instrumental adjunct itself:

(13) a.*Ho caricato il camion con la sabbia con la gru
have-1sg loaded the truck with the sand with the crane
'I have loaded the truck with the sand withthe crane'
b.*I have loaded the truck with sand with the crane

(14) a. Ho caricato la sabbia sul camion con la gru
have-1sg loaded the sand on-the truck with the crane
'I have loaded the sand on the truck with the crane'
b. I loaded the sand on the truck with a crane

Third, it can be clefted (15), as a true adjunct (16), and unlike the PP in the basic variant (17) (Mateu 2000, 33):

(15) a. Ho caricato il camion e l'ho fatto con la sabbia
have-1sg loaded the truck and it-have-1sg done with the sand
'I loaded the truck and I did it with sand'
b. I loaded the truck and I did it with sand

(16) a. Ho caricato il camion e l' ho fatto sabato mattina
have-1sg loaded the truck and it-have-1sg done saturday morning
'I loaded the truck and I did it Saturday morning'
b. I loaded the truck and did it Saturday morning

(17) a.*Ho caricato la sabbia e l'ho fatto sul camion
have-1sg loaded the sand and it-have-1sg done on-the truck
b.*I loaded the sand and I did it on the truck

Yet, the *with*-phrase does not have all the properties of adjuncts. First, it does not seem to behave as an adjunct for weak islands[16]:

(18) a.?Con cosa ti dispiace aver caricato il camion?
with what you(dat) sorry have loaded the truck
'With what are you sorry to have loaded the truck?'

[16] Judgements about this kind of constructions are sometimes not very clear. I checked these sentences with several Italian speakers, and the majority agreed with the judgments provided in (19). English and German speakers seem to agree that wh-movement of the *with*-phrase (19a) is not so bad as that of a true adjunct (19b).

b.*Come ti dispiace aver caricato il camion?
How you(dat) sorry have loaded the truck
'How are you sorry to have loaded the truck?'
c. Cosa ti dispiace aver caricato con la sabbia?
what you(dat) sorry have loaded with the sand
'What are you sorry to have loaded with sand?'

There is a clear difference in grammaticality between (18a) and a sentence involving a true adjunct like *come* "how" in (18b). Example (18a) is slightly worse than one in which a true argument has been moved (18c), but it is slightly better than (19), in which the *of*-phrase has been moved:

(19) ?? Di cosa ti dispiace aver caricato il camion?
Of what you(dat) sorry have loaded the truck
'With what are you sorry to have loaded the truck?'

As far as Italian speakers feel any difference between (18a) and (19), they tend to find (18a) better. The *of*-phrase, in turn, cannot be clefted:

(20) * Ho caricato il camion e l'ho fatto di sabbia
have-1sg loaded the truck and it-have-1sg done of sand
'I have loaded the truck and did it with sand'

This seems to indicate that the cleft-construction is not a good test for argumenthood: the *of*-phrase cannot be clefted, yet it fares rather badly in the weak-island test. Note also that sentence (13) improves considerably, in Italian, if we move the instrumental adjunct to the beginning of the sentence:

(21) ?* Con la gru, ho caricato il camion con la sabbia
with the crane, have-1sg loaded the truck with the sand
'With the crane, I have loaded the truck with sand'[17]

This indicates that (13) is bad for independent reasons. Even a comitative adjunct does not sound perfectly grammatical after the *with*-phrase (22):

(22) a. ?* Ho caricato il camion con la sabbia con Gianni
have-1sg loaded the truck with the sand with John
'I have loaded the truck with sand with John'
b. ?* I have loaded the truck with sand with John

[17] This example is not easily replicated in English, probably because of a difference in the syntax of topicalization.

And finally, an instrumental adjunct cannot be coordinated with the *with*-phrase:

(23) a.*Ho caricato il camion con la sabbia e con la gru
have-1sg loaded the truck with the sand and with the crane
'I have loaded the truck with sand and with the crane'
b.*I have loaded the truck with sand and with the crane[18]

In conclusion, there is no clear evidence that proves that the *with*-phrase is a simple instrumental adjunct. It seems more correct to say that it has both adjunct and argument properties: it behaves as an argument for the weak island test (18a), but it can be dropped, as an adjunct (11). To account for this ambiguous status, I will assume that the *with*-phrase is indeed an argument of the verb, namely the Theme, and is therefore generated as a DP in [Spec, SC], as required by UTAH.[19] Following Kayne (1999), I will assume that the preposition *with* is generated higher up in the functional structure of the clause, and that the surface order P - DP is derived through phrasal movement, driven by case reasons. More precisely, I will follow Kayne's (2002) hypothesis that prepositions select a case projection (labelled KP) and that the Theme DP moves to the specifier of that projection (within a larger phrase) in order to be case-licensed.

To recapitulate, the main steps of the derivation are the following:

(24) i. After incorporation of the abstract P onto the verb, the Locative argument moves out of the VP to [Spec, AgrO]
ii. The verb moves out of the VP to check its features
iii. The Theme argument moves to [Spec, KP], the case projection selected by *with*

The resulting syntactic configuration (abstracting away from the movement of the verb) is the following[20]:

(25) [PP with [KP sand ... [AgrOP truck ...

[18] Of course, both sentences are grammatical in the irrelevant reading in which the crane is also loaded on the truck.

[19] For the possibility of dropping the Theme in the *with*-phrase, see below, sections 4.1 and 4.2.

[20] Notice that, as (25) shows, on the likely assumption that prepositions are merged higher than AgrO, simple spec-to-spec movement of the Theme would result in the word order *I loaded with sand the truck*. This in turn means that a more complex type of phrasal movement is involved in the derivation. For concreteness' sake, I will assume that after (24iii) the preposition *with* attracts in its specifier the maximal projection immediately below KP.

In conclusion, Kayne's hypothesis that prepositions can be merged into the structure of the clause independently from their complements provides a way to account for the "demotion" of the Theme DP in the derived variant of the locative alternation: the Theme cannot get structural case from the verb so it moves to a higher case position, selected by a preposition.

5. The Theme in the *of*-variant

In the previous section I assumed that the *with*-phrase has the same properties in both Italian and English, but there is actually a difference which has not been previously reported in the literature, to the best of my knowledge: in Italian the *with*-variant does not sound completely felicitous with indefinite nouns:

(26) ?* Ho caricato il camion con sabbia
have-1sg loaded the truck with sand
'I have loaded the truck with sand'

Plural bare nouns seem to degrade the sentence even more:

(27) * Ho caricato il camion con tubi
have-1sg loaded the truck with tubes
'I have loaded the truck with tubes'

With an indefinite noun or a plural bare noun the *of*-variant must be used instead (28a). This variant in turn is impossible with definite nouns (28b):

(28) a Ho caricato il camion di sabbia/tubi
have-1sg loaded the truck of sand/tubes
'I have loaded the truck with sand/tubes'
b.*Ho caricato il camion della sabbia
have-1sg loaded the truck of-the sand
'I have loaded the truck of the sand'

English and German, of course, do not have the possibility of realizing the Theme with an *of*-phrase, and therefore do not show this restriction[21]

[21] It might be the case that this restriction does not extend in the same way to all Romance languages. Mateu (2000) judges the Spanish translation of (26) grammatical, and that seems to be true also in my (Latin-American) Spanish. The Spanish equivalent of (27), though, does not sound completely grammatical to me.

Geenhoven (1998) has proposed that (some) indefinites[22] denote a property, and are therefore, from a semantic point of view, predicates. She also argues that these predicative nominals (semantically) incorporate onto the verb, a process which is visible in languages with overt incorporation like West Greenlandic, where incorporated nouns are (predicative) indefinites. I will argue that this incorporation occurs in the *of*-variant in Italian, and that this accounts for several semantic properties of the Theme in this variant. Before examining these properties in detail, let us define precisely the syntactic derivation of the *of*-phrase.

5.1. *The syntactic derivation of the Theme in the of-variant*

As mentioned in section 3, the UTAH-based approach adopted here forces us to adopt the same initial structure for all variants of the locative alternation, which in turn means that the *of*-variant is derived from structure (4b). This in turn means that the Theme is generated in [Spec, SC/PP] as a DP. As the first step in the derivation, I will propose, *contra* Geenhoven's lexicalist approach, that the incorporation of the indefinite noun takes place in the syntax. More precisely, I will assume that the noun denoting the property is the *classifier* of the Theme and that it is this abstract noun that incorporates onto the verb, leaving behind the referring noun:

(29) [VP NOUN_i-P_j-caricare [SC [DP t_i sabbia] t_j camion][23]

In (29) the (empty) classifier defines the *kind* the Theme belongs to, and it moves out of the Theme DP, "stranding" the head noun.[24] Another consequence of this hypothesis is that it automatically makes the Theme DP case-licensed: Baker (1988) argues that a DP from which a noun has incorporated does not need structural case. Thus, while the Locative DP in (29) moves out of the VP to get structural case, the Theme DP stays in situ

[22] More precisely, bare plurals and split noun phrases in German.

[23] Note that incorporation in (29) takes place from a specifier position, contra Baker's (1988) assumption that incorporation takes place only from the complement position of the verb. For empirical arguments against this restriction see den Dikken (1995), Damonte (2004).

[24] This kind of incorporation is overtly visible in polysynthetic languages like Mohawk, in cases like *my father fish-bought eight bullheads* (Mithun 1984, 870), in which the incorporated noun *fish* acts as the classifier of the "doubled", external object *bullheads*; see Mithun (1984) for extensive discussion of this type of incorporation. Note also that two elements are now incorporated onto the verb now, as the abstract P in (29) also incorporates, as claimed in section 2. Again, this kind of prefixation, followed by incorporation of a noun, is visible in languages with more agglutinative morphology, see cases like *hand-Pref-make thing*, "handmade things" in Ainu (Shibatani (1990, 66)).

inside the VP. Finally, I will argue that this incorporation is also the reason why the Theme can be omitted in the derived variant (cf. ex (11) in the preceding section): the incorporated Classifier restricts the set of possible Themes, so that the verb can license an empty category as the Theme.

As for the reason why the two interpretations require two different prepositions in Romance, one would like to link it to the different syntax of the two prepositions in the two languages: more precisely, I will propose that in Italian *di* is generated inside the DP and *con* is merged in the functional structure of the clause. In English, on the other hand, *with* can be generated both inside DP and CP.[25] At the very least, this hypothesis seems to entail that there should be other uses of *di* in Italian that correspond to *with* in English. The prediction seems to be borne out: this correspondence extends to most other classes of locative verbs, such as *spray*, *fill*, *swarm* etc.[26] Let us summarize the proposed derivation of the *of*-variant:

(30) i. An empty P incorporates into the verb, allowing the Locative argument to move out of the VP and get structural case.
ii. The classifier of the Theme incorporates onto the verb, case-licensing the Theme DP in situ.[27]

5.2. *The semantic properties of the Theme in the of-variant*

On the basis of data from West Greenlandic, Geenhoven (1998) shows that incorporated nouns in this language can be modified but not quantified. According to Geenhoven this is due to the fact that the incorporated nominal is, semantically, a pure predicate and it does not introduce a variable that can

[25] This goes against Kayne's (2002) hypothesis that *di* and *of* are not DP-internal. As pointed out to me by Cecilia Poletto, it is probably possible to rephrase the analysis presented here according to the hypothesis that all prepositions are merged in the functional structure of the clause, but I will not investigate this alternative any further here.

[26] But interestingly not to a small set of "privative" verbs like *empty*, *rid, rob*, which take *of* in English.

[27] Note that the classifier cannot incorporate in the basic variant, otherwise the verb would not assign structural case. This in turn means that predicative indefinites in the sense of Geenhoven cannot appear in the basic variant either:

(i) Ho caricato libri sul camion
have-1sg loaded books on-the truck
"I loaded books on the truck"

The prediction seems to be borne out: the Theme in (i) seems to have either a partitive or specific reading, but not the unbounded interpretation associated with the Theme in the *of*-variant.

be interpreted as a member of a presupposed set. Quantifiers presuppose their domain and are therefore incompatible with predicative indefinites. The Theme in the *of*-variant seems to have these properties:

(31) a.* Ho caricato il camion di molta/poca/tanta sabbia
have-1sg loaded the truck of much/little/a lot sand
'I have loaded the truck with much/little/a lot of sand'[28]
b. Ho caricato il camion di sabbia fine/rossa/proveniente dal fiume
have-1sg loaded the truck of sand fine/red/coming from the river
'I have loaded the truck with fine/red/coming from the river/ sand'

As expected, the definite Theme in the *with*-variant can be quantified:

(32) Ho caricato il camion con molta/poca/tanta sabbia
have-1sg loaded the truck with much/little/a_lot_of sand
'I have loaded the truck with much/little/a lot of sand'

Notice that quantification in the *of*-variant is possible if the quantifier takes scope over *kinds*:

(33) a. Context: a truck driver has refused to load dangerous radioactive material on his truck. He says:
b. In vita mia ho caricato il mio camion di molti materiali (ma questa roba aveva paura perfino di toccarla)

[28] According to Geenhoven (1998) numerals can modify a predicative indefinite if they have a cardinal reading. Thus, *ho caricato il camion di tre chili di sabbia*, "I have loaded the truck of three kilos of sand" should be fine if the numeral does not have a partitive reading, i.e the three kilos of sand are not part of a bigger, presupposed set. The prediction does not seem to be borne out:

(i) ?* Ho caricato il camion dei tre chili di sabbia che sono arrivati stamattina
have-1sg loaded the truck of-the three kilos of sand that are arrived this morning
"I have loaded the truck with the three kilos of sand that arrived this morning"

The relative clause in (i) should force a cardinal reading of the numeral (i.e. "those specific three kilos"), yet the sentence does not improve. On the other hand, this might be due to the difficulty of interpreting the numeral as non-partitive. With the same modification, example (31a), as expected, becomes perfectly grammatical:

(ii) ? Ho caricato il camion della poca sabbia che è arrivata stamattina
have-1sg loaded the truck of-the little sand that is arrived this morning
"I have loaded the truck with the little sand that arrived this morning

'In my life I have loaded my truck of many materials (but this stuff I was afraid even to touch)'

Given a context like (33a), which makes it clear that we are talking about different *types* of Themes, quantification is possible, as expected, cf. Geenhoven (1997, 45) for similar examples in German and West Greenlandic. The hypothesis that a classifier incorporates when the *of*-phrase contains a predicative indefinite accounts for this peculiar restriction on quantification, which would otherwise remain unexplained.[29]

I will also propose that this incorporation is the trigger for the *holistic effect*. More precisely, following Brinkmann (1997), I will assume that the fact that the Theme cannot be quantified changes the aspectual structure of the event: the Theme is unbounded, so that the only natural ending point for the process is the filling up of the truck. Consequently, in perfective tenses the *of*-variant will entail that the *Locative* argument (the truck) has been completely loaded, but will *not* entail that the *Theme* has been completely loaded, as the Theme cannot be quantified. Notice that the *Theme* is unbounded, but the overall *event* is not: it is delimited by the Locative argument. This answers Michaelis and Ruppenhofer's (2001) objection that predicates like (31b) are accomplishments, i.e. bounded causative events.

Notice also that I argue that the Theme is *unbounded*, not necessarily *non-incremental*. As Michaelis and Ruppenhofer correctly observe, it is not enough for a Theme to be non-incremental for it to be optional, cf. the ungrammaticality of the sentence *the box contains*. But in this example the Theme is clearly bounded, even if it is non-incremental.

This hypothesis predicts that in Italian, the holistic effect is present in the *of*-variant, but not in the *with*-variant. This seems to be true:

(34) a. Context: One ton of sand has to be loaded onto a three-ton lorry
b.#Ho caricato il camion di sabbia
have-1sg loaded the truck of sand
c. Ho caricato il camion con la sabbia
have-1sg loaded the truck with the sand

Italian speakers agree that (34b) clearly entails that the truck has been completely loaded. While this reading is also possible in (34c), it is possible to build a context, such as (34a), which makes clear that the truck has *not*

[29] Note that indefinite (but not predicative) nouns introduced by *di* can be quantified in Italian: *un autore di molti libri* "an author of many books".

been completely loaded. The *of*-variant cannot be used in this context (34b), while the *with*-variant is still possible (34c).[30]

If we extend this analysis to Germanic languages, it follows that the derived variant in English and German is systematically ambiguous between two different structures, as the *with*-phrase can introduce both definite and indefinite, quantified and mass nouns.

(35) a. I loaded the truck with hay/the books/three tons of hay
b. Ich belud den Lastwagen mit Heu/den Büchern/drei kilo Heu
I loaded the truck with hay/the books/tree kilo hay
'I loaded the truck with hay/the books/tree kilo hay'

It also also follows that the holistic effect should be present in English and German when the *with*-phrase contains an indefinite noun (36), but not in other cases (37):

(36) a. I loaded the truck with sand
b. Ich belud den Lastwagen mit Sand
I loaded the truck with sand
'I loaded the truck with sand'

(37) a. I loaded the truck with the sand
b. Ich belud den Lastwagen mit dem Sand
I loaded the truck with the sand
'I loaded the truck with the sand'

The speakers I consulted tend to agree that a completive reading is possible in the examples in (37), but it is not necessary, so that these sentences can be felicitously used to describe the context in (34a). Germanic languages thus give the impression that there is no restriction on the quantification on the Theme in the derived variant, since the same preposition is used to introduce both bounded and unbounded Themes. But the restriction becomes visible if indefinite nouns are used as Themes: Jackendoff (1990, 172-3) reports that in English bare plurals cannot be quantified by *some* in the derived variant:

(38) a. Felix loaded some books on the truck
b. ?* Felix loaded the truck with some books

[30] Again, I would say the same facts hold for Spanish as well.

In our view, this is due to the incorporation of the abstract classifier noun in (38b). If this analysis is correct, it predicts that *some* (and other quantifiers) should be impossible also with singular mass nouns:

(39) a. ?? Felix loaded the truck with some sand
b. * Felix loaded the truck with little sand
c. * Felix loaded the truck with much sand[31]

As expected, modification with *some* is possible if the quantifier refers to a specific amount, not to part of a larger, presupposed set (see also the Italian example (ii) in fn. 28, above):

(40) Felix loaded the truck with some stones that arrived this morning

Another prediction the theory makes is that when the Theme is dropped in the derived variant, in examples like (11), repeated below, only the holistic interpretation should be possible, in that only the structure in which the classifier incorporates allows for the non-realization of the Theme:

(41) a. Ho caricato il camion
have-1sg loaded the truck
'I have loaded the truck'
b. I loaded the truck

Again, the prediction is borne out: only the completive reading is possible in examples (39).

If the hypothesis proposed here is on the right track, it will lead to a reformulation of Brinkmann's *Nonindividuation Hypothesis*:

(42) *Nonindividuation Hypothesis* (Brinkmann 1997)
For a verb to take its goal as direct object, the quantificational properties of the Theme must be irrelevant. The Theme may then be construed as nonindividuated when it is not specified, i.e., as an unbounded amount of stuff or objects
(Brinkmann 1997: 248).

As we have shown in this section, the non-individuation constraint is only valid for the *of*-variant in Romance languages, but not for the *with*-variant in Germanic languages, as confirmed by the fact that while there is a strong

[31] But as Mark de Vos, who provided the judgements, points out, the ungrammaticality of this example could be due to the archaic use of *much* as modifier of a mass noun.

tendency in linguists's example to use bare plurals as the Theme in the derived variant (Brinkmann 1997, 105), definite and quantified Themes are possible in both German and English (cf. (35)). This in turn means that non-individuation is a necessary condition to trigger the *holistic effect*, but that prefixation alone is sufficient for a verb to take its "goal" as direct object.

To summarize the conclusions reached in this section, I proposed that the holistic effect is linked to a property of a particular class of indefinites, namely predicative indefinites, and that the *of*-variant in Italian is restricted to these indefinites.

6. Conclusions

The derivational analysis proposed here is still sketchy, and many important points still need to be sorted out, yet it already accounts for some significant generalizations about the cross-linguistic distribution of the variants of the verb l*oad*, the most important being that the basic variant is indeed more basic: in no language the verb *load* appears only in the derived variant[32]. If we extend the analysis to other classes of locative verbs, Kim (1999, 102) shows that many languages do not allow *any* verb to appear in the derived configuration only.

Similarly, in a derivational approach the absence of the derived variant in a given language can be accounted for by non-semantic reasons. Returning to the cross-linguistic survey carried out by Kim, and quoted in section 1, it might be that what blocks the derived variant of the verb *load* in isolating languages like Chinese is the fact that these languages do not use prepositions to introduce Locative arguments, but a "serializing" strategy, i.e serial verbs constructions or verb compounding, as pointed out by Kim. In our terms this means that no preposition incorporates onto the verb and therefore the Locative argument cannot get structural case.

The next question, of course, is whether this analysis can be extended to other classes of locative predicates, and account for why some verbs only appear either in the basic or derived variant. Here I will just point out the ways in which the analysis proposed here could account for these facts. A verb may fail to license the abstract preposition and thus appear only in the basic variant. This might be the case of purely spatial predicates like *put* or *place*, which are not inherently defined for a certain kind of motion. On the contrary, some verbs, such as *fill*, might be inherently specified for unbounded Themes, and they will therefore appear in the derived variant

[32] Furthermore, Kim et al. (1999) show that cross-linguistically children generalize the basic variant to verbs like *fill*, but they never assume that verbs like *pour* only appear in the derived variant.

only, as the incorporation of the classifier is not possible in the basic variant. More research is needed to see whether these hypotheses are sound.

References

Baker, Mark. 1988. *Incorporation.* Chicago, University of Chicago Press.

Baker, Mark. 1997. "Thematic roles and syntactic structure". In Liliane Haegeman (ed.) *Elements of Grammar*. 73-137. Dordrecht, Kluwer Academic Publishers.

Borer, Hagit. 1994. "The Projection of Arguments". *University of Massachussetts Occasional Papers* 17:19-47.

Brinkmann, Ursula. 1997. *The locative alternation in German. Its structure and acquisition*. Amsterdam/Philadelphia, John Benjamins.

Chomsky, Noam. 1993. "A minimalist program for linguistic theory". In Kenneth Hale and Samuel Jay Keyser (eds.) *The view from building 20.* 1-52. Cambridge MA, The MIT Press.

Damonte, Federico. 2004 *The Thematic Field: The syntax of argument structure enhancing morphology*. Ph.D dissertation, University of Padua.

den Dikken, Marcel. 1995. *Particles: on the syntax of verb-particle, triadic, and causative constructions*. Oxford, Oxford University Press.

Geenhoven, Veerle. 1998. *Semantic incorporation and indefinite descriptions*. Stanford, CSLI.

Gronemeyer, Claire. 1995. "Swedish applied verbs derived by the prefix *be*". *Lund Working Papers* 44:21-40.

Hall, Barbara. 1965. *Subject and Object in English*, Ph.D. dissertation, MIT.

Hoekstra, Teun. 1992. "Aspect and theta theory". In Iggy M. Roca (ed.) *Thematic Structure: Its Role in Grammar*. 145-174.Berlin, Foris.

Hoekstra, Teun and René Mulder. 1990. "Unergatives as copular verbs; locational and existential predication". *The Linguistic Review* 7:1-79.

Jackendoff, Ray. 1990. *Semantic Structures.* Cambridge MA, MIT Press.

Kayne, Richard. 1999. "Prepositional complementizers as attractors". *Probus* 11:39-73.

Kayne, Richard. 2002. "On some Prepositions that look DP-internal: English *of* and French *de*", *Catalan Journal of Linguistics* 1:71-115.

Kim, Meesook. 1999. *A Cross-Linguistic Perspective On The Acquisition Of Locative Verbs*. Ph.D. dissertation, University of Delaware.

Kim, Meesook, Barbara Landau and Colin Phillips. 1999. "Cross-linguistic differences in children's syntax for locative verbs". In *Proceedings of the 23rd Boston University Conference on Language Development*. 337-348. Brookline MA, Cascadilla Press.

Larson, Richard. 1990. "Double objects revisited: reply to Jackendoff". *Linguistic Inquiry* 21:589-632.

Levin, Beth. 1993. *English Verb Classes and Alternations*, Chicago/London, The University of Chicago Press.

Levin, Beth and Malka Rappaport Hovav. 1996. "From Lexical Semantics to Argument Realization", Ms. Northwestern University/Bar Ilan University.

Mateu, Jaume. 2000. "Lexicalization patterns and the locative alternation". Ms. Universitat Autònoma de Barcelona.

Michaelis, Laura A. and Josef Ruppenhofer. 2001. "Valence Creation and the German Applicative: The Inherent Semantics of Linking Patterns" *Journal of Semantics* 17:335-395.

Mithun, Marianne. 1984. "The evolution of noun incorporation". *Language* 60, 4:847-94.

Mulder, René. 1992. *The aspectual nature of syntactic complementation*. The Hague, HIL.

Munaro, Nicola. 1994. "Alcuni casi di alternanza di struttura argomentale in Inglese". In Gianluigi Borgato (ed.) *Teoria del Linguaggio e Analisi Linguistica: XX Incontro di Grammatica Generativa.* 341-370. Padova, Unipress.

Pinker, Steven. 1989. *Learnability and Cognition: the Acquisition of Argument Structure.* Cambridge MA, MIT Press.

Rappaport, Malka and Beth Levin. 1988. "What to do with θ-roles". In Wendy Wilkins (ed.) *Thematic Relations*. 7-36. San Diego, Academic Press.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge, Cambridge University Press.

Wunderlich, Dieter. 1987. "An Investigation of Lexical Composition: The Case of German *be*-verbs". *Linguistics* 25:283-331.

On merging positions for arguments and adverbs in the Romance Mittelfeld*

Christopher Laenzlinger and Gabriela Soare
University of Geneva

1. Introduction

The major aim of this paper is to propose a computational system based on multiple phases capable of accounting for the positioning of the arguments and of the verb in the domain of the clause referred to as the Mittelfeld. This paper will be mainly concerned with the Romance Mittelfeld though numerous references will be made to English.

The assumption will be made that the domain situated between the complementizer system and the VP-shell consists of Cinque's (1999) functional hierarchy of adverbs and of recursive chunks of A-positions interspersed among each adverb class. The hypothesis to explore is that all arguments must leave the VP-shell in order to have their A-features (Case, phi-features) as well as their IS-features (topic, focus etc) checked/matched. It ensues that scrambling is applicable not only to OV languages like German or Japanese but also to VO languages like English and the Romance languages. Variations in the Information Structure in the Romance languages are responsible for the different configurations found in these languages, esp. with respect to adverb intervention. The configurations to focus on are: (i) SVO, (ii) VSO, and (iii) VOS and the languages to consider are French, Italian, Romanian and Spanish.

* We owe thanks to Ur Shlonsky for having read through the paper. We also thank Andrea Cattaneo, Paola Merlo, Sandra Schwab and Violeta Seretan for judgments.

1.1. *The Framework*

The paper adopts Kayne's (1994) phrase structure theory where the operation Merge forms [$_{XP}$ Spec X Compl] configurations, no further merge of either multiple specifiers or of adjoined elements being possible. Adverbs are thus merged as unique specifiers of semantic-functional projections (Cinque 1999, 2002, Laenzlinger 2000, 2004). Adverbs, and adjuncts in general, are not thematically selected. They must however be semantically licensed in their syntactic position, they must be part of the narrow syntax since they are LF-dependent. Their analysis as specifiers is linked to Cinque's (1999) proposal that they are associated with different semantic projections which are provided by UG.

As hinted at above, the cartographic approach to the clause structure is involved here, the clause structure extending from the thematic VP domain, which contains only the lexical verb and its arguments up to the discourse-related CP domain. Within the framework of the cartographic approach (Belletti 2004, Cinque 2002, Rizzi 1997, 2004a,b), the intermediate inflectional space, equated to the Mittelfeld, constitutes a domain rich in functional projections.

1.2. *A-positions within the Mittelfeld*

The sentences in (1) show that there must be floating positions for the verb and the object among the fixed positions of the three adverb classes.

(1) a. Jean a probablement lu souvent la Bible attentivement
Jean has probably read often the Bible carefully
b. Jean a probablement souvent lu attentivement la Bible
Jean has probably often read carefully the Bible
c. Jean a lu ?probablement la Bible souvent attentivement
Jean has read probably the Bible often carefully
d. Jean a probablement lu souvent attentivement la Bible
'John probably often read the Bible carefully'

The Mittelfeld in the cases above represents the domain of adverb licensing, i.e. it consists of the projections hosting the adverbs *probablement* 'probably', *souvent* 'often', *attentivement* 'carefully'. This domain also contains the projections hosting the moved arguments, i.e. the DP object *la Bible* 'the Bible' and the verb *lu* 'read'. In order to identify these positions, the following hypothesis couched in terms of a principle is made (2). This

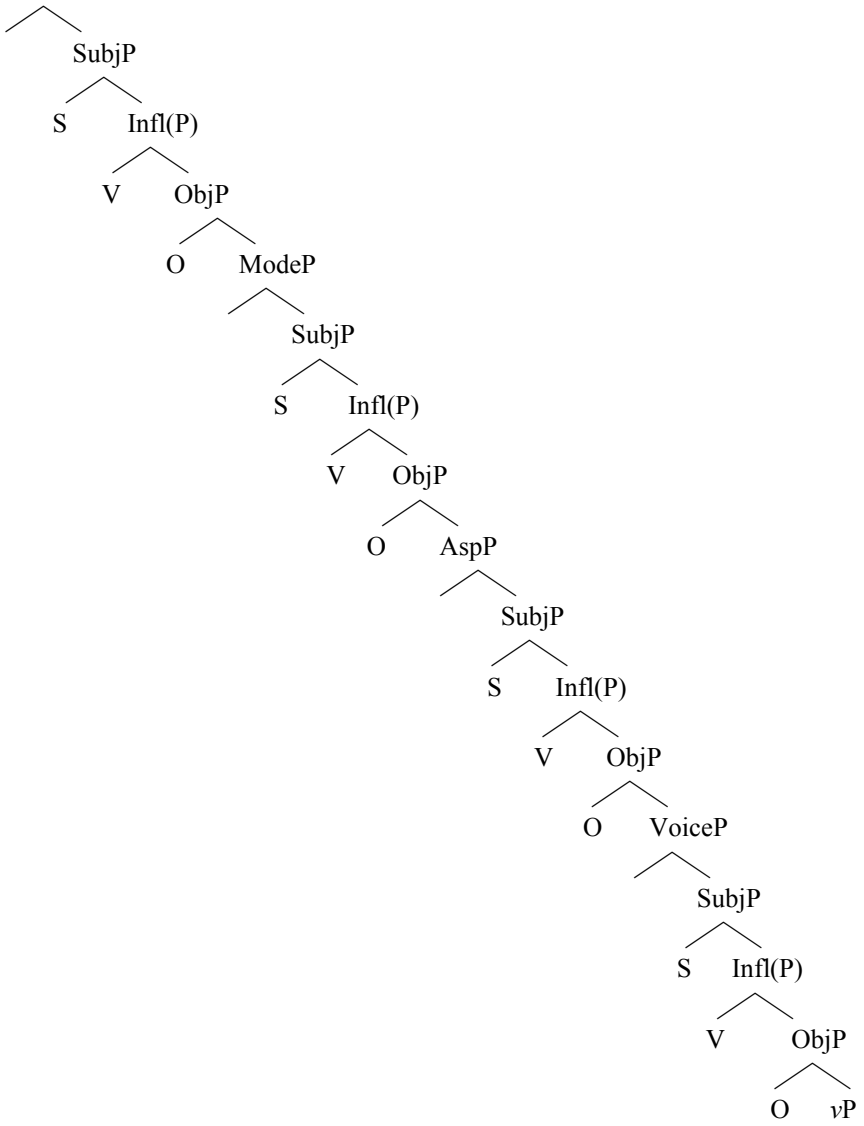
proposal goes against Alexiadou and Anagnostopoulou (2001) who assume that one argument must remain in the VP-shell.[1]

(2) Full VP Evacuation Principle:
"All arguments must leave the vP domain in order to have their A-features (i.e. Case and phi-features) and I-features (i.e. informational features such as top, foc) checked/matched/assigned a value in the overt syntax."

To put it in a nutshell, the Mittelfeld contains not only the adverb-related functional projections but also the verb- and DP-related positions. In order to provide a uniform account of the cross-linguistic facts evinced by the Romance languages under investigation, a system of SVO recursive chunks interspersed among the adverb-related projections is proposed.[2] Such SVO chunks can potentially merge between every semantic-functional projection, as illustrated in (3).[3]

[1] Recall that no adverb merges in the VP-domain, contra Alexiadou and Anagnostopoulou (2001:206).
[2] A chunk can be even more complex, i.e. in the case of multiple complements.
[3] The advantage of the system proposed here can be extended to account for data from Hungarian, a discourse-configurational language, or Greek.

(3)…MoodP



Derivationally, after having merged within the VP-shell, the verb and its arguments are attracted to the relevant positions among the adverb-related projections. The subject targets the specifier position of SubjP, the equivalent of what was called AgrsP or TP, that is, the position where the subject phi- and Case(NOM)-features coupled with the EPP-feature can be matched. As for the object, it is attracted to Spec-ObjP, corresponding to the former AgroP (Belletti 1990) or to Chomsky's (1995) AspP, responsible for

Case(ACC)-feature checking. Since a specifier can be projected, the Obj is also associated with abstract phi-features and an EPP-feature.

As regards the lexical verb, it targets an inflectional projection, InflP (corresponding to AuxP in compound tenses), for checking morpho-selectional features (Aux-$V_{\text{past participle}}$ selectional checking in compound tenses). Verb movement is analysed as head-movement in the Principles & Parameters theory (Pollock 1989). More recently, following Koopman and Szabolcsi's (2000) framework, verb movement can be implemented as remnant VP-movement (Majahan (2000); see also Laenzlinger (2004) for a detailed analysis). The latter approach will be tentatively adopted in this paper. In addition to the selectional feature checking requirement, verbal elements have uninterpretable phi-features which are checked on the head Subj for subject-verb agreement. More precisely, Subj attracts the conjugated auxiliary/lexical verb, thus realizing a Spec-head configuration with the DP subject, if present in the structure. Alternatively, V-phi-features are checked via downward Agree with a lower subject.

In the A-feature system put forth here, the phi-features can be checked/erased in the local relation [Spec-SubjP]-Subj° in the word order SVO. In the word order VS(O) or V(O)S, the phi-features are checked via downward Agree (Chomsky 2001), thus an equally local relation obtaining between the head Subject and its specifier, which is lexically realized lower in the structure. The [Spec-SubjP] hosts an expletive *pro* which is required for EPP-feature checking, this being in line with Rizzi's (2004c) Subject Criterion. The Nominative Case feature is checked/erased in a Spec-head relation of a SubjP projection in a high (SV) or low (VS) "chunk". Insofar as the Accusative case is concerned, little is said in the Minimalist Program (2000) and, in a manner similar to Nominative, we assume that besides the Spec-head relation obtained between [Spec-ObjP] and the head Obj°, the checking/erasure of the uninterpretable Accusative Case can be realized by the mechanism Agree in the word order VSO. The head Obj also contains phi-features such as number and gender which are activated in a configuration where the object DP raises above the participial verb (Kayne 1989). As known, only French and Italian display overt agreement, while Spanish and Romanian do not. This issue will be taken over in section 2.3.

According to the principle stated in 2, besides the set of A-features, arguments move to check features related to the Information Structure (henceforth IS) as well. The Information Structure as envisaged here is organised in function of the chunks, i.e. in the Lexical Array every major constituent of the sentence is assigned a certain value of informational prominence, which amounts to saying that every chunk can potentially be marked for different values of the topic and focus features or simply remain

unmarked for either feature (for a different view see Vallduvi's 1992 Information Packaging, Lambrecht 1994, Choi 1999, Büring 2003). In other words, it potentially contains different types of Foci, such as a Mittelfeld-internal low new information Focus in the sense of Belletti (2001, 2004) or a contrastive Focus in the CP domain in the sense of Rizzi (1997), a Topic-aboutness in the upper layer of the Mittelfeld or a Topic-comment in the CP. IS is thus expressible in terms of such value-assignment to various constituents in various chunks. Put differently, IS is mapped onto syntactic constituents and this mapping is a parametrisable function of the language under investigation and context as well. More research is needed to develop a sound theory of information structure meant to capture the subtle meaning differences resulting of the various distribution of arguments among the adverbs in a sentence.[4]

Argument realization in the chunks thus becomes dependent not only on the positions available for phi- and Case feature checking, but also on different value assignment of informational prominence features. A word of caution is needed here: no IS is associated with the vP.

It may emerge evident that the EPP feature has a special status in the system proposed here in the sense that, along minimalist lines (Chomsky 2000), it is responsible for movement and, at the same time, it seems to be intimately connected to IS. For instance, its presence on the head Subj attracts movement of the DP in the specifier position and, in the light of the above discussion, a certain value of informational prominence may be assigned to the moved argument.

1.3. *More on SVO Chunks and the Computational System*

According to (3) every SVO chunk marks a phase from one adverb-related projection to another. We depart from the view (Chomsky 2001) that only vP and CP are phases which mark a cycle/barrier, that they are movable and thus being sent separately to Spell-Out and propose instead that there are as many phases as there are SVO chunks. Thus, the intermediate chunks between the semantic-functional projections serve as escape-hatches to movement,[5] as represented in (4).

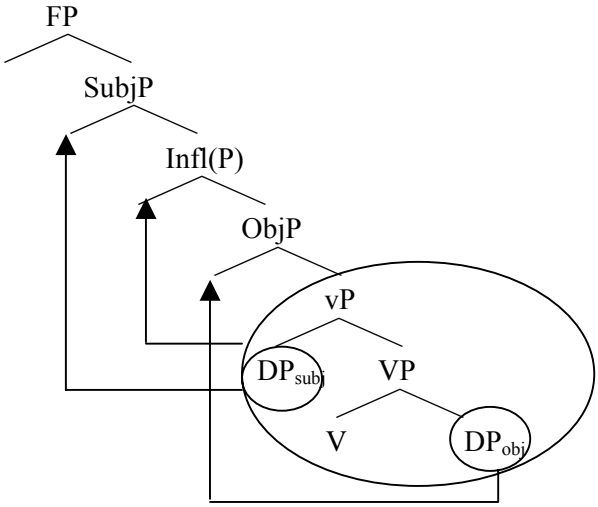
(4) [$_{\text{MoodP}}$ SVO [$_{\text{ModeP}}$ SVO [$_{\text{TP}}$ SVO [$_{\text{AspP}}$ SVO [$_{\text{VoiceP}}$ SVO [$_{\text{VP}}$ ….]]]]]

[4] One interesting approach to Information Structure is put forth by López and Villalba (2000) who take the phase as the IS unit thus developing an Information Structure at the v phase and another at the C phase.

[5] This is in accordance with Chomsky's (2001) Phase Impenetrability Condition.

Within such a system the 'cycle' property of phases derives the order S-(V)-O, the one originally merged within the VP-shell. Successive movement of the object, the verb (projection) and the subject is represented in (5).

(5)

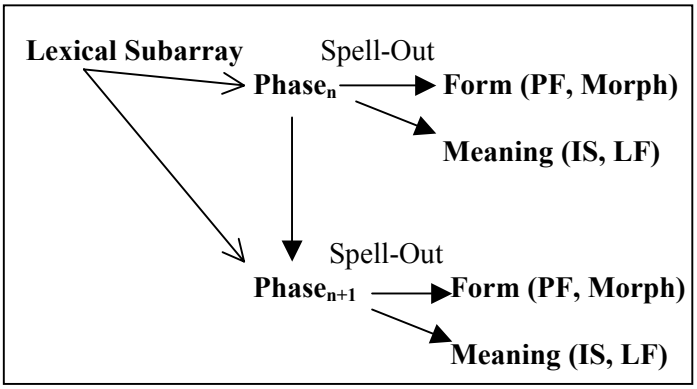


Following a multiple Spell-Out approach[6] (Uriagereka 1999, Platzack 2001, Grohmann 2003), at the end of each phase, the derived substructure is sent to the interface (PF and LF) for interpretation. The grammar system has the representation in (6). Information Structure is also considered an interface interacting with PF (i.e. stress in focalisation) and LF (i.e. covert movement of Focus).

In the light of these theoretical considerations, the following sections will centre on the distribution of adverbs in the word orders SVO, VSO and VOS in English and the Romance languages under investigation: French, Italian, Romanian and Spanish.

[6] The multiple Spell-Out approach raises the question of multiple access to LF, which leads to the compositional semantics of the clause, which is not adequate for a full interpretation of the propositional content of the clause.

(6)



2. Verb-Argument-Adverb Orders in the Mittelfeld

2.1. *Unmarked SVO: English and French*

The word order SV(O) represents the underlying word order in the two non-null subject languages. Consider first the case of high adverbs.

(7) a. [$_{MoodP}$ Fortunately [$_{SubjP}$ John read the book]]
b. [$_{MoodP}$ Heureusement [$_{SubjP}$ Jean a lu le livre]]
'Fortunately Jean read the book.'

Mood adverbs are perfectly acceptable in sentence-initial position which, barred the existence of a topic, represents their merge position, i.e. [Spec-MoodP] in (7). The subject DP can float very high in the upper Mittelfeld, i.e. above such adverbs to check the EPP, phi and Case features (8a). Conversely, French does not allow any adverb in between the subject position and the auxiliary/verb in the head SubjP (8b). One micro-parametric property of French is that feature-checking of the subject and the auxiliary/conjugated verb must be realized in a Spec-head configuration.

(8) a. John fortunately read this book.
b.*Jean heureusement/probablement/souvent/récemment a lu ce livre.
Jean fortunately/probably/often/recently has read the book.'

Turning to the epistemic adverb *probably*, the examples below show that it may fill its root-merge position, at the boundary between the Mittelfeld and the Vorfeld in (9a). The corresponding French sentence in (9b) allows only a parenthetical reading of the adverb. As with the case above, the DP subject in (9c) can target the appropriate position in a chunk above ModP in English,

such a possibility is ruled out in French due to the violation of the adjacency subject-auxiliary/verb, as indicated by the ungrammatical (9d). In (9e,f), a chunk made up of the subject and the auxiliary is positioned above ModP. In English and French, its sentence-final position results in a parenthetical reading, as illustrated in (9g,h).[7] However, French displays one more alternative in which this high adverb may intervene between the participial verb and the object DP (9j), English displaying the well-known Case-resistance constraint (9i).[8] Translated into the theory of chunks, Stowell's (1981) constraint forces the verb and the nominal DO to occur within the same (S)VO chunk.[9]

(9) a. Probably John has read the book.
b. Probablement, Jean a lu le livre.
c. John probably has read the book.
d.*Jean probablement a lu le livre.
e. John has probably read the book.
f. Jean a probablement lu le livre.
g.*John has read the book probably/(^{ok}, probably).
h.*Jean a lu le livre probablement/(^{ok}, probablement).
i. *John has read probably the book.
j. ?Jean a lu probablement le livre.

Let us further consider the examples containing a $TP_{anterior}$ adverb like *recently* or an $AspP_{frequencyI}$ adverb like *often* .

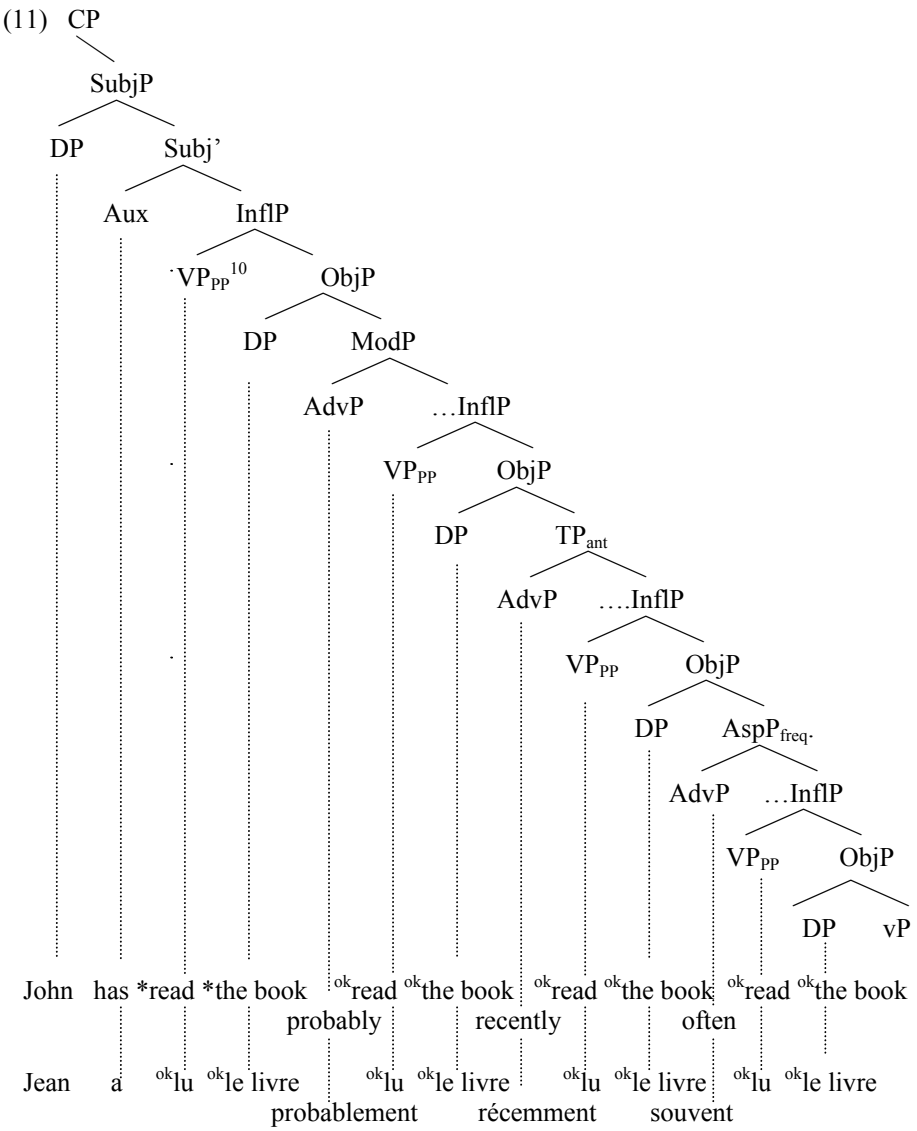
(10) a. Mary has (often/recently) read (*often/*recently) the book (often/recently).
b. Marie a (souvent/récemment) lu (souvent/récemment) ce livre (souvent/récemment).

In (10a) the adjacent constituents, the verb and the direct object check their features in a position higher than both the $AspP_{frequency}$ licensing the adverb *often* and the TP licensing the adverb *recently*. In French (10b), the participial verb can float to a position preceding or following the time and aspect adverbs. In order to account for the verb always preceding the direct

[7] It is not the aim of this paper to tackle the issue of parenthetical phrases.
[8] The function of *probably* as a phrasal modifier is not considered here.
[9] As suggested by Ur Shlonsky (personal communication), the direct object DP in English is a light element being morphologically/prosodically poor. Informationally, they are weak elements and, as such, must belong to the same chunk as the verb. This view finds support in the process of heavy NP-shift. If the object is a heavy DP as well as a PP or CP, it acts as a free element with respect to the verb.

object in the Mittelfeld in English as well as in Romance (only as far as the SVO word order is concerned), the condition must be observed that the verb c-command the direct object DP.

(11)



To sum up, the moved arguments and verbs' positions among the functional projections of the mode, time and aspect adverbs considered thus

[10] $VP_{PP} = VP_{Past\ Participle}$

far are provided in (11) above.[11] As the structure above shows, in French the auxiliary must move very high, i.e. to a head Subj in the upper Mittelfeld in order for the Spec-head configuration to obtain.

Finally, let us consider argument floating with respect to unambiguous manner adverbs such as *beautifully* and *merveilleusement* 'beautifully'.

(12) a. John has sung this song beautifully.
b. John has beautifully sung this song.
c.?? John beautifully has sung this song.
d.* John has sung beautifully this song.
e. Marie a chanté la chanson merveilleusement.
f. Marie a chanté merveilleusement la chanson.
g. Marie a merveilleusement chanté la chanson.
'Marie has sung the song beautifully.'

Examples (12a,e) show that a whole chunk made up of the subject, auxiliary, participial verb and object may float to their appropriate positions above the adverb-related projection. Alternatively, in both languages only the subject and the auxiliary can target higher positions (12b,g). Contrary to other classes of adverbs, a manner adverb can only marginally occur between the subject and the auxiliary in English, as illustrated in (12c). As already pointed out, the ungrammaticality of (12d) can be accounted for in terms of a violation of the requirement that the verb and its nominal direct object occur within the same chunk. In French, a chunk consisting of a subject, auxiliary and participial verb may alternatively occur higher than the adverb, with the object remaining in a lower position in the space between VoiceP and vP, as (12f) shows.

Mention must be made at this point that the variable floating nature of the chunks conducive to word order variations with adverbs does not solely

[11] Further evidence for the existence of several derived positions of the arguments in the Mittelfeld comes from floating quantifier placement. Whatever the approach adopted, a floating quantifier must be in a local relation with respect to the associate subject/object, as illustrated in (i).

(i) Ces filles, les professeurs les ont (récemment) tous (récemment) toutes
these girls the professors them-have recently all_{masc} recently all_{fem}
(récemment) félicitées.
recently congratulated

Though the sentence contains a topicalised direct object, the co-occurrence of a subject and object floating quantifiers must observe the strict order Subj>Obj. Their variable position with respect to the time adverb further shows that there are several intermediate positions for the subject and object to have landed in.

depend on morpho-syntactic properties (the A-system). The role of adverbs in information structure is not very well known. To our knowledge, very few analyses exist on the topic and particular attention should be paid to this issue especially given the view of the apparently 'free' placement of (some) adverbs. Consider, for instance, the placement possibilities of the manner adverb around the participial verb and its object in French in (12e-g). The sentence final position of the adverb in (12e) is appropriate if the adverb conveys new information, i.e. as an answer to the question "*Comment a-t-il chanté la chanson*?*" "How did he sing the song*?*"* , in which case the verb and its object are informationnally unmarked and target a chunk above VoiceP, at the centre of the Mittelfeld. According to the analysis, this space is the locus of neutral information. When the adverb is preverbal, as in (12g), it overtly marks its scope to the verb plus its complement, which remain in a chunk below VoiceP. Finally, the order in (12f) results in an informational balance among of the verb, the adverb and the object. The verb occurs in a chunk above VoiceP, while the object is located in a chunk below VoiceP. The conclusion we can draw from such observations is that the distribution of constituents among the different chunks is necessarily triggered by information structure.

2.2. *Marked SVO: Italian, Spanish and Romanian*

One common property of these languages is that the preverbal subject is assumed to bear a topic-like feature in the sense that, as Rizzi (2004c) puts it, if events are conceived of in the subject-predicate format, the description of an event may start by selecting an argument, with the event presented as being about that argument.[12] Therefore, preverbal subjects in these null-subject languages share the features [+aboutness, -D-linking] that they can check in [Spec-SubjP]. The Topic feature, strongly connected to the IS interface is thus checked together with Case, phi- and the EPP features present on the featural make-up of the head Subj. An answer taking on the form SV(O) can be given to a question like "*What about X*?" Though more marked, the subject is still part of the Mittelfeld.

Argument floating among the same adverb classes will be analysed in this section. Let us first consider MoodP adverbs.

[12] As far as Spanish is concerned, there seems not to be a consensus as to its basic word order. Some linguists (Suñer 1994, Ordóñez 2000) consider that the basic order is SVO, yet allowing its subject to appear postverbally, while others (Zubizarreta 1998, Costa 2001) claim that the basic word order is VSO. However, in chapter 2, fn. 3, Ordóñez himself considers examples in which the answer to the question "*Qué pasó*?" "What happened?" is given in the order XP/que V S O. Spanish will be taken in this paper to have VSO as the unmarked order.

(13) a. Francamente Gianni si era formato una pessima opinione di voi.
Frankly Gianni se_{refl} was made a very bad opinion of you
'Frankly Gianni had a very bad opinion of you.'
b. Gianni francamente si era formato una pessima opinione di voi.
c. Evidentemente (,) Maria dirá la verdad.
'Evidently, Maria will tell the truth.'
d. Maria felizmente ha leido este libro.
'Fortunately Maria read this book.'
e. Din fericire, Ion a citit cartea.
'Happily Ion read the book.'
f. Ion, din fericire, a citit cartea.

As the Italian sentences (13a,b) show, a subject position may be available in the upper Mittelfeld, the same holding true for Spanish though, according to some speakers, such adverbs may be parentheticals (13c,d). In Romanian (13e,f), such high adverbs, i.e. $MoodP_{evaluative}$, $MoodP_{evidential}$ and $MoodP_{speech-act}$ have a parenthetical reading, whatever their position in the sentence.[13]

Insofar as the epistemic adverb *probabil* 'probably' is concerned, the three languages pattern with English in that a subject position may be projected above it, as illustrated in (14a,b,c). One parametric variation of Spanish and Romanian is that they both exhibit the auxiliary-verb adjacency which can be accounted for in terms of one-step movement of Aux° to Subj°.[14] As the example in (14d) shows, Italian exhibits only auxiliary raising above the functional projection of the adverb, participial verb movement above the epistemic adverb is not allowed (14e), while in Romanian the chunk DP_{Subj}-Aux-VP_{PP} may occur above ModP as in (14f). Spanish opts for none of the three possibilities.

(14) a. Gianni probabilmente ha letto il libro.[15]
Gianni probably has read the book
b. Ion probabil a citit cartea.
Ion probably has read the-book
c. Juan probablemente ha leído este libro.
Juan probably has read this book

[13] Romanian exhibits classes of adverbs in which an adverb has an adverbial PP counterpart, the latter having a higher frequency of language use. A case in point is the class of $MoodP_{evaluative}$ adverbs which contains only PPs, such as: *în mod regretabil* 'regrettably', *în mod neaşteptat* 'unexpectedly, *din (ne)fericire* '(un)fortunately'.

[14] In Romanian the auxiliary-verb adjacency can be interrupted only by a restricted class of clitic adverbs.

[15] To some speakers the adverb in this configuration does not display a parenthetical use.

d. Gianni ha probabilmente letto il libro.
e.*Gianni ha letto probabilmente il libro.
f. Ion a citit probabil cartea.

It has been seen above that the subject can raise very high in Italian, Spanish and Romanian, which means that different adverb classes may appear between this position and that of the verb. Such is the case of time and aspect adverbs, as illustrated in (15a-c) below for Spanish and Romanian.

(15) a. Juan recientemente/a menudo ha leído este libro. (Spanish)
Juan recently/often has read this book.
b. Ion recent/RECENT a citit această carte. (Romanian)
Ion recently has read this book.
c. Ion adesea/ADESEA a citit această carte.
Ion often has read this book.
d. Gianni *recentemente/*spesso ha letto questo libro. (Italian)
Gianni recently/often has read this book
e. Gianni ha recentemente/spesso letto questo libro.
f.. Gianni ha letto recentemente/spesso questo libro.
g. Gianni ha letto questo libro recentemente/spesso.

Interestingly, Romanian seems to accommodate a Mittelfeld-internal Focus position hosting the two adverbs (as will be seen, a manner adverb can as well be focalised). This process of clause-internal focalisation remains mysterious on syntactic grounds. Either the focalised adverb occurs in its root-merge position where it is stressed (in-situ Focus), or it is displaced in a Mittelfeld-internal FocusP.

Though exemplified only for Italian (15g), the three languages are alike in that a chunk containing the participial verb and the object may precede the adverbs under analysis. Alternatively, the object DP may remain lower in the structure, as illustrated in (15f). As already mentioned, Italian does not display the auxiliary-verb adjacency and thus the two adverbs may occur in between, as seen in (15e). However, Italian departs from Spanish and Romanian in not allowing a subject position right above the two adverb classes (15d).[16]

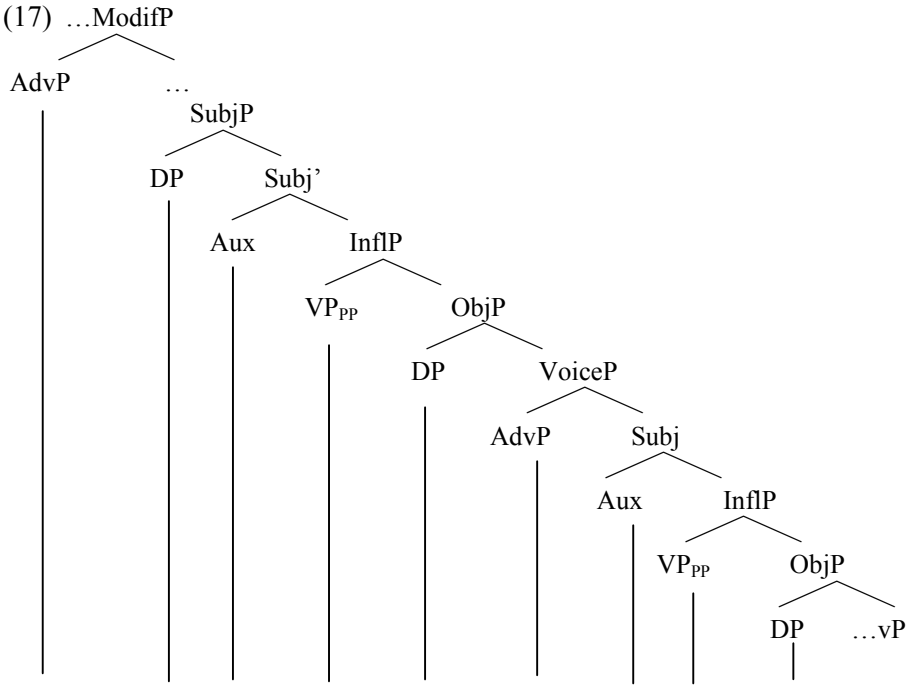
Argument floating is further analysed with respect to truly manner adverbs. As illustrated in the examples in (16a,b), Italian and Spanish pattern with French in that a pre-auxiliary position of the adverb leads to a

[16] Surprisingly, as seen above, the DP subject can move above ModeP and MoodP.

parenthetical reading of the adverb.[17] Romanian does not project a subject position above the manner adverb (16c). The three languages exhibit the same behaviour with respect to object and participial verb movement to [Spec-ObjP] and [Spec-InflP] in a chunk above VoiceP. The structure in (17) provides the movement possibilities of the sentences below.

(16) a. Juan (‚glotonamente,) ha comido (glotonamente) la manzana (glotonamente). (Spanish)
'Juan ate the apple greedily'.
b. I bambini (‚dolcemente,) hanno (dolcemente) accarezzato (dolcemente) il gattino (dolcemente). (Italian)
'The children caressed the cat gently. '
c. Copiii (*frumos) au împachetat (frumos) cadourile (frumos).
'The children wrapped up the presents beautifully.' (Romanian)

(17)



[17] Spanish also exhibits a Mittelfeld-internal Focus position hosting only manner adverbs in the pre-auxiliary position.

The structure above also contains the [Spec-ModifP] at the left periphery of the clause. Such a position is activated only in Spanish, while the clause-initial position of (unambiguous) manner adverbs in Romanian results in their parenthetical use.

To sum up this section, the three Romance languages make available a SubjP in the upper Mittelfeld where the moved subject checks the Topic-like feature, intimately connected to IS, this position can be above MoodP adverbs Italian and Spanish but not higher than $ModP_{epistemic}$ adverbs in Romanian.

2.3. *The Word Order VSO*

As mentioned in the previous section the word order VSO is more natural in Romanian and Spanish and may constitute an appropriate answer to a question like "*What happened*?" As noted by Belletti (1999) and as will be discussed below, VSO is impossible in Italian. Another major difference between Italian and the two languages is that in cases of free inversion an adverb may follow the postverbal subject only in latter. Compare (18a) to (18b,c,d).

(18) a. Ha mangiato Gianni *golosamente/*a volte/*recentemente.
Has eaten Gianni greedily/sometimes/recently
'Gianni has eaten greedily/sometimes/recently'. (Italian)
b. Ha mangiato golosamente/a volte/recentemente Gianni.
c. Ha llorado Juan recientemente/a menudo/discretamente.
Has cried Juan recently/often/gently
'Juan has recently/often cried gently.' (Spanish)
d. A plâns Ion recent/adesea/zgomotos.
Has cried Ion recently/often/noisily
'Ion has recently/often cried /noisily.' (Romanian)

The subject, which expresses unmarked information, can float above VoiceP, AspP and TP in Romanian and Spanish. Alternatively, the floating subject can remain in a chunk below and among the adverb-related projections. This is represented in (19).

(19) $[_{SubjP}$ Aux $[_{InflP}$ V_{pp} $[_{SubjP}$ Subj $[_{TP}$ Adv $[_{SubjP}$ Subj $[_{AspP}$ Adv $[_{SubjP}$ Subj $[_{VoiceP}$ Adv $[_{SubjP}$ Subj $[_{vP}$ …]]]]]]]

However, in Italian, an adverb can only precede the subject, which means that in the system proposed here the postverbal subject in (18b) fills the

appropriate position in the lowest chunk, this lowest subject position being marked as 'new information' focus.

Among the adverb classes considered thus far, all except mood and mode adverbs (unless parenthetically used) may occur in the postverbal domain between the participial verb and the subject, as illustrated in (20a) for Romanian and (20b) for Spanish.

(20) a. A citit probabil/recent/adesea/atent Ion această carte.
Has read probably/recently/often/carefully Ion this book.
'Ion probably/recently/often/carefully read this book'. (Romanian)
b. A leído probablemente/recientemente/a menudo/atentamente Juan
Has read probably/recently/often/carefully Juan
este libro.
this book.
'Juan probably/recently/often/carefully read this book.' (Spanish)

Such data can be interpreted in terms of participial movement to a chunk above VoiceP, AspP and ModeP. These adverbs may occur after the subject, which means that the subject DP can alternatively check its features in a position above these adverb classes, the object remaining in a lower chunk, as illustrated in (21) below.

(21) a. A citit Ion probabil/recent/adesea/atent această carte.(Romanian)
Has read Ion probably/recently/often/carefully this book.
'Ion probably/recently/often/carefully read this book .'
b. A leído Juan probablemente/recientemente/a menudo/
Has read Juan probably/ recently/often/
atentamente este libro. (Spanish)
carefully this book.
'Juan probably/recently/often/carefully read this book.'

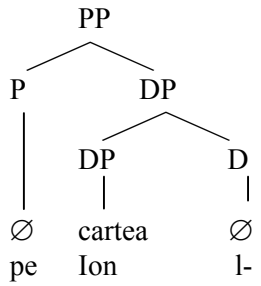
It is to remark that movement of the subject past the object, both being nominal chains, does not induce any minimality violation effect in either Romanian or Spanish but does so in Italian, hence the impossibility of the word order VSO (**Ha letto Gianni questo libro*). We suggest that the explanation for the absence of intervention effects is to be sought in the process of clitic doubling. Following proposals on clitic doubling constructions (Uriagereka 1995, Belletti 1999), we assume that nominal direct objects in clitic doubling languages are more than DP categories. More precisely, they always project a PP, with a dummy preposition and an

overt clitic if the object is doubled as in (22a,b), or with an empty preposition and a null clitic if it is not doubled, as in (22c,d).

(22) a. L-a văzut pe el/Ion. (Romanian)
Him$_{Cl}$-has seen pe$_{Accprep}$ him/Ion.
'S/he saw Ion.'
b. Lo vimos a el/Juan. (Spanish)
Him see a him/Juan.
'We see Juan.'
c. A citit cartea. (Romanian)
Has read book-the.
'S/he read the book.'
d. A leído el libro. (Romanian)
Has read book-the.
'S/he read the book.'

Thus, the nominal direct object in Spanish and Romanian always projects a PP with the following root-merge structure.

(23)



Being a PP, the nominal object does not act as an intervener with respect to the displaced subject. The word order VSO in which both DPs are nominal is ruled out in Italian because it does not display clitic doubling with direct objects, a minimality violation being induced on the subject's chain *Gianni*).[18,19]

[18] In cases where the internal complement of the verb is a PP or a CP, the word orders V S PP and V S CP are grammatical in Italian, the former being, however, slightly marginal.

(i) ?Ha telefonato Gianni a Maria.
Has phoned Gianni to Maria.
'Gianni phoned Maria.'

To conclude the discussion on floating arguments, time, aspect and manner adverbs can occur sentence-finally, i.e. they follow the object. In this position mood and mode adverbs must be parentheticals. In other words, the object can reach a chunk above VoiceP, AspP and TP, but not above ModP.

2.4. *The Word Order VOS*

Italian, Spanish and Romanian also exhibit the word order VOS though things are more complicated in Italian. As Belletti (2001) argues, this word order is highly restricted in Italian since the verb and the direct object must constitute directly accessible 'given' information, or reiterates the information contained in the question, as seen in (24) below.

(24) A: Chi ha letto questo libro ?
Who has read this book
'Who read this book?'
B: Ha [letto questo libro] Gianni.
Has [read this book] Gianni.
'Gianni read this book.'

Romanian and Spanish allow VOS without such a contextual restriction, the postverbal subject filling the new information focus.

(25) a. A citit cartea Ion. (Romanian)
Has read book-the Ion.
'Ion read the book.'
b. A leído este libro Juan. (Spanish)

(ii) Ha detto Silvia che ha telefonato Gianni.
Has said Silvia that has phoned Gianni.
'Silvia said that Gianni phoned.'

In (i) the nominal object which is a PP is not an intervener on the subject's chain, PP's being autonomously licensed. Within the framework proposed here, the clausal object in (ii) is analysed as filling the lowest objetc position. As a matter of fact, the only position it can fill is sentence-final, its raising past the subject would lead to ungrammaticality. The clausal object's occupying the lowest object position may be related to its informational weight, i.e. it must occur at the right edge of the sentence because it is a heavy element in need of checking some feature associated to information structure. As Belletti notes, a clausal object can however precede the subject with some verb classes.

[19] An interesting consequence of the analysis of the Romanian/Spanish nominal direct object as PPs is that once displaced, they do not show past participial agreement, as opposed to French and Italian.

Ordóñez (2000) argues for Spanish that VOS structures are the result of scrambling of the object past the subject DP.[20] Similar analyses were proposed for Romanian by Cornilescu (1997) and Alboiu (1999). For Italian, Belletti extensively argues for an analysis of such structures in terms of clause internal remnant topicalisation. The subject raises to the low new information focus position and the [VO] constituent containing the subject trace raises above, in the Spec of the TopP. Since the direct object is embedded in a larger constituent, the ObjP-SubjP chains do not show any intervention effects. Such an analysis in terms of remnant VP movement is not without problems as it presupposes that the object remains in the VP, which runs counter to the Full VP Evacuation Principle used throughout the analysis. Thus in (25) the verb and the object move separately to [Spec-InflP] and [Spec-ObjP], respectively, for feature-checking purposes.

Let us concentrate on Romanian and Spanish data taking into account the distribution of adverbs.

(26) a. A citit (probabil/recent/adesea/atent) această carte (?probabil/
Has read probably/recently/often/carefully this book
recent/adesea/atent) Ion (‚probabil/recent/adesea/atent).[21]
'Ion recently/often/carefully read the book.' (Romanian)

b. A leído (probablemente/recientemente/a menudo/atentamente) este libro (?probablemente/recientemente/a menudo/atentamente) Juan (‚probablemente/recientemente/a menudo/atentamente).
'Juan recently/often/carefully read the book.' (Spanish)

These sentences show that the two languages pattern alike in that a subject DP can raise to a position as high as TP and AspP, but not ModP (given the parenthetical reading of *probably* in sentence-final position). In such a position the subject is associated with a 'floating' new information Focus feature.[22] In other words, the subject is attracted by the head Subj containing not only Case and phi-features, but also an EPP-feature associated with

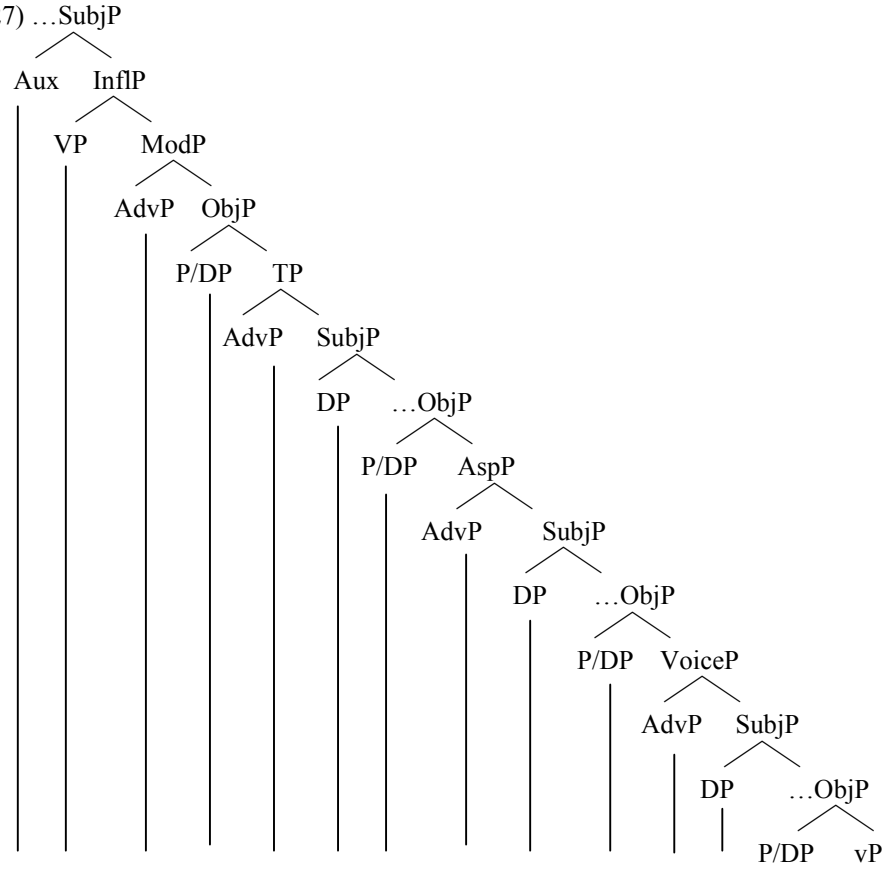
[20] Within a somewhat similar vein, Ordóñez (2000) proposes a uniform analysis of the VOS word order for Spanish, Catalan, Italian and even French in terms of subject movement to the Focus position and object movement to a higher c-commanding position. His analysis includes LPR (light predicate raising) of the remnant TP with any XP that follows the subject scrambling out of the TP before LPR applies.

[21] The time, aspectual and manner adverbs can marginally be focalised in Romanian when filling a position between the participial verb and the object in the order VOS.

[22] Mention must be made that in Romanian the subject in VOS can also be associated with contrastive Focus feature. It may be suggested that there is raising of the focused DP followed by remnant IP movement to CP.

Focus and determined by IS. The object DP is attracted to an Accusative Case position, i.e. [Spec-ObjP], which may range from VoiceP to AspP and TP, as indicated by the interference of such adverbs between the subject and the object, and between the participial verb and its object. The floating positions for the object and the subject are represented in (27).

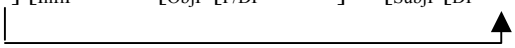
(27)



A citit probabil cartea recent Ion cartea adesea Ion cartea atent Ion cartea
A leído prob..e.libro recien. Juan e.libro a men. Juan e.libro atent. Juan e.libro

Mention must be made that in keeping with the analysis of clitic doubling advanced here, the nominal direct object in Romanian and Spanish is a PP and thus, unlike in Italian, the object does not act as a DP intervener in the subject's chain in Romanian and Spanish. The representation in (28) shows that the nominal object as a PP is invisible to the probe and the Minimal Link Condition required for the operation downward Agree on the subject is observed. Such a configuration clearly holds for prepositional objects as well.

(28) $[_{SubjP}\,[_{Subj}\,A]\,[_{InflP}\,citit\,...[_{ObjP}\,[_{P/DP}\,cartea]\,...\,[_{SubjP}\,[_{DP}\,Ion]\,...]]]]$



As a summation, among the micro-parametric properties of the three languages discussed in the configuration VOS, it is to mention that Subj-Aux agreement can be done by the operation downward Agree, that the subject is associated with a 'floating' new information focus feature, and, finally, that there is no intervention effect between Subj and Obj as the latter is a kind of PP category.

3. Conclusions

This paper has proposed a computational system based on multiple Spell-Out which is capable of accounting for word order variations in the Romance Mittelfeld taking as an essential criterion Cinque's (1999) rigid hierarchy of adverbs. In so doing the Full VP Evacuation Principle has been forwarded according to which all arguments must leave the vP domain for A- and IS-feature checking. The role of IS has been pointed out, in the sense that it is intimately related to the EPP-feature. Arguments move in order to express some informational value, marked or unmarked, the information value they convey being organised in function of the chunks. The system described here relies on the existence of SVO chunks, interspersed among adverb-related projections, each chunk marking a phase. Based on the rigid position of adverb classes (mode, mood, time, aspect and voice adverbs have been considered here), it has been extensively showed that there are floating positions for the subject, object and verb in the configurations investigated, SVO, VSO and VOS. It seems, therefore, that the richer the A-system is, the more IS can avail itself of the system of chunks.

References

Alboiu, Gabriela. 1999. "Object Raising and Two Types of Focus in Romanian". *Bucharest Working Papers in Linguistics* 1:79-106

Alexiadou, Artemis and Elena Anagnostopoulou. 2001. "The Subject-in-Situ Generalization, and the Role of Case in Driving Computations". *Linguistic Inquiry* 32:193-231.

Belletti, Adriana. 1990. *Generalized Verb Movement: Aspects of Verb Syntax*. Turin, Rosenberg and Sellier.

Belletti, Adriana. 1999. "Italian/Romance Clitics: Structure and Derivation". In Henk van Riemsdijk (ed.) *Clitics in the Languages of Europe*. 543-579. Berlin, Mouton de Gruyter.

Belletti, Adriana. 2001. "Inversion as focalization". In A. Hulk and J.-Y. Pollock (eds.) *Subject inversion in Romance and the theory of universal grammar*. 60-90. New York/Oxford, Oxford University Press.

Belletti, Adriana (ed.). 2004. *Structures and Beyond: The Cartography of Syntactic Structures Vol 3*. New York/Oxford, Oxford University Press.

Büring, Daniel. 2003. "Intonation, Semantics and Information Structure". Ms. (handbook article), UCLA.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford, CSLI Publications.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA, MIT Press.

Chomsky, Noam. 2000. "Minimalist Inquiries: the Framework". In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik.* 89-155. Cambridge MA, MIT Press.

Chomsky, Noam. 2001. "Derivation by Phase". In M. Kenstowicz (ed.) *Ken Hale: A Life in Language.* 1-52. Cambridge MA., MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York/Oxford, Oxford University Press.

Cinque, Guglielmo (ed.). 2002. *Functional Structure in IP and DP*: *The Cartography of Syntactic Structures*, vol.1. New York/Oxford, Oxford University Press.

Cornilescu, Alexandra. 1997. "The Double Subject Construction in Romanian". *Revue Roumaine de Linguistique,*101-104.

Costa, João. 2001. "Word Order Typology in Optimality Theory". In Aafke Hulk and Jean-Yve Pollock (eds.) *Subject inversion in Romance and the theory of universal grammar*. 91-107. New York/Oxford, Oxford University Press.

Grohmann, Kleanthes. 2003. *Prolific Domains*. Amsterdam/Philadephia, John Benjamins.

Kayne, Richard. 1989. "Facets of Romance Past Participle Agreement". In Paola Benincà (ed.) *Dialect Variation and the Theory of Grammar*. 85-104. Dordrecht, Foris.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge MA, MIT Press.

Laenzlinger, Christopher. 2000. "More on Adverb Syntax". In Artemis Alexiadou and Peter Svenonius (eds.) *Linguistics in Potsdam*, vol.6, 103-132. University of Potsdam.

Laenzlinger, Christopher. 2004. "A Feature-based Theory of Adverb Syntax". In J. Austin, S. Engelberg and G. Rauch (eds.) *Adverbials. The*

Interplay between Meaning, Context and Syntactic Structure. 205-252. Amsterdam/Philadelphia, John Benjamins.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge, CUP.

López, Luis and Xavier Villalba. 2000. "Phases and Information Structure". *GLOW Newsletter* 44:80-81.

Mahajan, Anoop. 2000. "Eliminating Head Movement". *GLOW Newsletter* 44:44-45.

Ordóñez, Francisco. 2000. *The Clausal Structure of Spanish: A Comparative Study*. New York, Garland.

Platzack, Christer. 2001. "Multiple interfaces". In U. Nikanne and E. van der Zee (eds.) *Cognitive Interfaces*. 21-53. New York/Oxford, Oxford University Press.

Pollock, Jean-Yves. 1989. "Verb Movement, UG and the Structure of IP". *Linguistic Inquiry* 20:365-424.

Rizzi, Luigi. (1997). "The Fine Structure of the Left Periphery". In Liliane Haegeman (ed.) *Elements of Grammar*. 281-337. Dordrecht, Kluwer Academic Publishers.

Rizzi, Luigi. 2004a. "Locality and the Left Periphery". In Adriana Belletti (ed.) *Structures and Beyond. The Cartography of Syntactic Structures*. 223-252. New York/Oxford, Oxford University Press.

Rizzi, Luigi (ed.). 2004b. *The Structure of CP and IP. The Cartography of Syntactic Structures*. New York/Oxford, Oxford University Press.

Rizzi, Luigi. 2004c. On the form of chains: criterial positions and ECP effects. Talk delivered at IGGXXX, Venice/Treviso, February 2004.

Stowell, Tim. 1981. *Origins of Phrase Structure*, Ph.D. dissertation, MIT.

Sũner, Margarita. 1994. "Verb Movement and the Licensing of Argumental Wh-Phrases in Spanish". *Natural Language and Linguistic Theory* 12:335-372.

Uriagereka, Juan. 1995. "Aspects of the Syntax of Clitic Placement in Western Romance". *Linguistic Inquiry* 26:79-124.

Uriagereka, Juan. 1999. "Multiple Spell-Out". In Samuel D. Epstein and Norbert Hornstein (eds.) *Working Minimalism*. 251-282. Cambridge MA, MIT Press.

Vallduví, Enric. 1992. *The Informational Component*. New York, Garland.

Zubizarreta, María-Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge MA, MIT Press.

The order of prepositional phrases

Walter Schweikert
University of Venice

1. Introduction

Modifiers in general and prepositional modifiers in particular were often considered to be adjuncts to some functional projection above the VP (be it vP or TP). This would prohibit any syntactic base order among themselves. If some rigid ordering was found, this was usually attributed to some semantic property.

This view changed radically with the publishing of Cinque (1999). In this book Cinque showed that certain types of adverbial modifiers namely adverbs, auxiliaries and modifying affixes of agglutinating languages obey strict ordering restrictions among themselves. A large sample of data from very different languages revealed that this order is universal:

Sentence modifying adverbs can be subdivided in a finite group of classes which obey a strict order relation among themselves. I give here the labels of these classes together with a typical representative:

$\text{Mood}_{\text{speech act}}$	frankly
$\text{Mood}_{\text{evaluative}}$	fortunately
$\text{Mood}_{\text{evidential}}$	allegedly
$\text{Mod}_{\text{epistemic}}$	probably
T (Past)	once
T (Future)	then
$\text{Mood}_{\text{irrealis}}$	perhaps
$\text{Mod}_{\text{necessity}}$	necessarily
$\text{Mod}_{\text{possibility}}$	possibly
$\text{Mod}_{\text{volition}}$	willingly
$\text{Mod}_{\text{obligation}}$	inevitably

$Mod_{ability,\ permission}$	cleverly
$Aspect_{habitual}$	usually
$Aspect_{repetitive\ I}$	again
$Aspect_{frequentative\ I}$	often
$Aspect_{celarative\ I}$	quickly
T (Anterior)	already
$Aspect_{terminative}$	no longer
$Aspect_{continuative}$	still
$Aspect_{perfect}$	always
$Aspect_{retrospective}$	just
$Aspect_{proximative}$	soon
$Aspect_{durative}$	briefly
$Aspect_{generic/progressive}$	characteristically
$Aspect_{prospective}$	almost
$Aspect_{SgCompletetive\ I}$	completely
$Aspect_{PlCompletive}$	tutto
Voice	well
$Aspect_{celarative\ II}$	fast, early
$Aspect_{SgCompletetive\ II}$	completely
$Aspect_{repetitive\ II}$	again
$Aspect_{frequentative\ II}$	often

(Cinque 1999:106)

Affixes in agglutinating languages, if realised as suffixes, obeyed the exact reversed order. If found as prefixes they are either in the original (direct) order, or in very rare cases such as Navajo, in reversed order. Auxiliaries which serve the purpose of these affixes in fusional languages such as English, show up in direct order.

In order to give an explanation to these facts Cinque proposed a syntactic hierarchy of functional projections between CP and VP. This was in fact an extension of the Split-Infl theory of Pollock (1989). Auxiliaries and Affixes representing modifiers are sitting in the heads of the respective projections, while their specifiers host the adverbs (AdvP).

In subsequent work (Cinque 2000) he showed that the above proposal could be extended to certain modals, the so called restructuring verbs.

In my dissertation I wanted to see whether it was possible to apply the idea of a rigid hierarchy to prepositional phrases which modify the VP. In order to verify this I had to find a suitable subdivision of PPs in classes, and then test whether an ordering relation among them could be found.

This work presents syntactic tests from my dissertation and their results together with certain statistical control methods, which might find application in other fields of linguistic research.

2. Thematic Roles as PP classes

Given their semantic content, thematic roles seemed to be the natural candidate for a subdivision of these modifying prepositional phrases into suitable classes. Two members of the same class cannot be added without a syntactic coordinator if referring to different entities:

(1) a. I decorated the box with a spray can and (with) a paint brush.
b.*I decorated the box with a spray can with a paint brush.

“with a spray can” and “with a paint brush” are both bearers of the same thematic role (instrumental). There is no semantic reason which prohibits having two instruments in the same sentence as example (1a) shows. Nevertheless, without a coordinator the sentence becomes ungrammatical. Since coordination is a syntactic device, I conclude that the thematic roles constitute syntactic classes.
Some sentences with two locative or temporal PPs which seem at first sight to be counterexamples, instead turn out to reinforce the analysis:

(2) a. I met John in Italy in Venice.
b. I met John in Venice in Italy.
c. I met John on Thursday at 8 o'clock.

In none of the cases the two PPs refer to different referents. In (2a) *in Venice* is a specification of the location and can be considered a modifier of the PP *in Italy* while in (2b) *in Italy* is a specification of *Venice* (the Venice in Italy, not the one in California). In (2c) *at 8 o'clock* is a modification of *on Thursday*.

If I want to express having met John in two different places or at two different times I again have to use coordination.

(3) a. I met John in Paris *(and) in Venice.
b. I met John on Thursday *(and) on Friday.

Using the above considerations as guidelines I stipulated in a first approach the following thematic roles as classes :

2.1. *Benefactive*

The Benefactive introduces a participant who benefits from the action done by the actor. In German the preposition is always “für”.

für seine Frau	for his wife
für seinen Chef	for his boss

2.2. *Comitative*

Comitatives add a person, which share the role of the subject. If the subject is an agent, they are semantically also agents. These additional agents are not introduced via coordination, but by means of a prepositional modifier. The accompanying preposition is in many languages the same as the one introducing instruments. In German this is “mit”, in Russian “s”, in English “with” and in Italian “con”. I do not think this is sheer coincidence, but for the moment I have no explanation for it. The syntactic tests, which will be illustrated in section 3, show clearly that its position is much higher than the one of the instrument.

mit einem Kollegen	with a colleague

2.3. *Evidential*

This group of prepositional modifiers adds the source of the proposition. This can be a person, but legends, stories and rumours can also be stated. German has two adpositions, which introduce them, “nach” and “gemäß”. Both can be used as prepositions or postpositions. “Nach” is more common with non human DPs. “Gemäß” as a preposition can have either a genitive or a dative complement; as postposition it always follows a dative DP.

einem Zeugen gemäß	according to a witness
gemäß eines Zeugen	according to a witness
nach einer alten Legende	according to an old legend
einer alten Legende nach	according to an old legend

2.4. *Goal*

This is a special kind of directional modifier which adds the goal of a movement. Since in many languages Goals are introduced by the same prepositions as Locatives, Directionals and Locatives are often grouped

together. In English you have to add the particle "-to" to some of the locative prepositions: "into", "onto", others like "under" are ambiguous. The preposition "to" by itself is only directional. In German, nearly all locative prepositions can be used in directional goal modifiers ("in", "auf", "unter", but not "bei"). Additionally, there exists "nach"

nach Hamburg	to Hamburg

2.5. *Instrumental*

This thematic role determines the instrument, the tool, which was used in order to commit the action. In German this role is exclusively realised by the preposition "mit".

mit einem Schraubenzieher	with a screwdriver

Since the same preposition is used with Comitatives and Means, they are often confused with each other though it remains to be determined whether Means and Instrumentals take different positions, but Comitatives and Instrumentals have rather different semantics and occupy distinct positions.

2.6. *Locative*

This is maybe the most common, in any case the most described thematic role. It determines the place where the action occurs. This is usually done by relating the event to an object, described by a DP. A great variety of prepositions make this relation explicit.

in Venedig	in Venice
hinter der Schule	behind the school
vor der Schule	in front of the school
neben der Schule	beside the school
auf dem Tisch	on the table
unter dem Tisch	under the table
über dem Tisch	above the table

2.7. *Malefactive*

This modifier adds an opponent, an obstacle to the proposition, a person or a (weather) condition which wants to block the action. Malefactives can also introduce a rival. The principal preposition in German is "gegen".

gegen das schlechte Wetter	against the bad weather
gegen seinen Erzkonkurrenten	against his arch-rival

2.8. *Manner*

This might be the most problematic group. Prepositional modifiers determine the manner in which a certain action was done. Frequently used prepositions introducing this thematic role are “mit” and “auf”. Speed modifiers are very often subsumed under this category. Since Cinque establishes frequentative and celerative adverbs as separate classes in his hierarchy, I was careful to use only certain expressions. In order to be always in the same class, I constructed examples with PPs of the following type.

auf besondere Art und Weise	in a special way

If taken in a broader sense, you would find examples such as:

mit Vorsicht	carefully
mit hoher Geschwindigkeit	with high speed

2.9. *Matter*

With this somehow artificial term, I named a group of modifiers that give the topic of a talk, the subject of research or a book. In German it is used with the preposition “über”.

über Mathematik	about mathematics

2.10. *Means of Transportation*

Cars, public busses, bicycles, airplanes are all examples of instruments, which can be used for movement. It is not clear whether this thematic role has to be distinguished from Instrumentals. But since verbs of movement display a particular behaviour, I decided to make this distinction. The results showed that Instrumentals and Means PPs are close neighbours, if separate at all. In German as in many other languages, they share the same preposition “mit”. In English, Means modifiers are often introduced by “by”.

mit dem Bus	by bus
mit einem Ferrari	with a Ferrari

2.11. *Path*

In addition to source and goal of a journey we can name a place, which has been passed by. In German, the preposition "über" introduces the place, sometimes you find "durch"

über Mainz	through Mainz
durch Mainz	through Mainz

2.12. *Reason*

This role determines the reason or motive a certain action was done. Typical prepositions are "wegen" and "aus":

wegen einer Krankheit	because of illness
aus Angst	out of fear

Reason modifiers are more sensitive to scope effects than most of the other types. There is a big difference between "Vincent painted because of the splendid light in Provence" and "Vincent painted in Provence because of the splendid light". The second sentence is indicating that Vincent went to Provence to paint, because of the splendid light there; this shows, that the reason modifier takes into its scope the Locative. In the first sentence the reason modifier only takes the nuclear event – Vincent painted – into its scope. The fact that the act of Vincent painting because of the splendid light takes place in Provins is just an additional information.

2.13. *Source*

Source modifiers specify the origin of a movement. They belong to the group of Directionals and are also related to Locatives. In many languages, combinations of a preposition like "from" and a locative preposition are used together to form something like "from under". Standard German does not allow for this construction, but several dialects have it ("von unter der Brücke"). Source modifiers are usually introduced by "von".

von München	from Munich

2.14. *Temporal*

These expressions determine the time interval in which the actual event takes place. It could be a year, a month, a certain day, an hour etc. The preposition in German is either “an”/“am” (with day), “um” (with time) or “in”/“im” (with month, year, season):

am Sonntag	(on) sunday
am gestrigen Tag	yesterday
um 14 Uhr	at 2 pm
im Dezember	in December
im Jahre 1492	in 1492
im Herbst	in autumn

3. Syntactic Tests

As a next step I had to check for ordering restrictions. But unfortunately there is no strict rigid surface order, as the following examples show.

(4) a. Canova sculpted with marble in Venice.
b. Canova sculpted in Venice with marble.
c. Leonardo worked for Sforza in Milan.
d. Leonardo worked in Milan for Sforza.

In the sentences (4a) and (4b) the thematic roles of Instrumental and Locative are reversed, but both sentences are grammatical. The same is valid for the Benefactive and Locative in the sentences (4c) and (4d). If there is a base ordering among thematic roles then movement must have produced (at least) one of the orders of each pair. Therefore, the next step was to look for syntactic tests which are sensitive for movement. Since German is my mother language I concentrated on this language, especially on the German Mittelfeld.

3.1. *Quantifier Scope*

The first test exploits the fact that sentences with two operators, where the lower has moved across the higher, exhibit scope ambiguity. I used sentences with two PPs in which one contains a universal quantifier and the other an existential. If the lower operator never crosses the higher we expect sentences with only one interpretation, the one with the higher operator taking scope over the lower:

$\forall x \quad (\exists y)$

or

$\exists x \quad (\forall y)$

In case of movement, however, we find scope ambiguity. Two interpretations are available, one with the moved element taking scope over the other and another interpretation with the originally higher one over the trace:

$\exists x_i \quad (\forall y \quad t_i)$
$\exists x_i \quad \forall y \quad (t_i)$

or

$\forall x_i \quad (\exists y \quad t_i)$
$\forall x_i \quad \exists y \quad (t_i)$

The ambiguity is often explained in terms of optional reconstruction. If two different thematic roles were base inserted in different but fixed positions, this test should give us in one order only one interpretation while in the other an obvious ambiguity.

Applying this test to the pair of matter PP and temporal PP results in a clear contrast. I evaluated two couplets of sentences. Each couplet retains the order of the operators but reverses the thematic roles. In the first couplet the existential operator comes first, in the second couplet the universal operator is in front.

3.1.1. *Matter – Temporal*

(5) Tony hat an mindesten einem Tag über jede
Tony has on at_least one day about every
Massenvernichtungswaffe gesprochen.
mass_destruction_weapon spoken
'Tony spoke about every mass destruction weapon on at least one day.'

(6) a. Tony hat an mindesten einem Tag über jede Massenvernichtungswaffe gesprochen.
∃ (time) ∀ (matter)
?? ∀ (matter) ∃ (time)

b. Tony hat über mindestens eine Massenvernichtungswaffe an jedem Tag gesprochen.
   ∃ (matter) ∀ (time)
   ∀ (time) ∃ (matter)
c. Tony hat über jede Massenvernichtungswaffe an mindesten einem Tag gesprochen.
   ∀(matter) ∃ (time)
   ∃ (time) ∀ (matter)
d. Tony hat an jedem Tag über mindestens eine Massenvernichtungswaffe gesprochen.
   ∀ (time) ∃ (matter)
   *∃ (matter) ∀ (time)

The prevalent interpretation of (6a) is that there is at least one special day on which Tony spoke about every mass destruction weapon. The reversed scope interpretation, that for every weapon there is at least one day on which he spoke about it – but not necessarily the same day for every weapon is nearly excluded.

In (6b) however we get both interpretations: 1) that there is a special weapon about which Tony spoke every day and 2) that he spoke every day about at least one weapon, but not necessarily the same one each day. From this we can conclude that (6a) represents the base order: Temporal is higher generated than Matter, while in (6b) the lower Matter PP is moved across the (original) higher Temporal.

The contrast in the second couplet with the universal operator both times coming first is even sharper. In (6d) the reversed scope interpretation is totally excluded, while in (6c) both interpretations are available.

Note also that, for me, the reverse interpretation in both (6b) and (6d) is prevalent, which I indicated with bold face.

So far this seems to be a convincing result, but before continuing let's have a look on another pair.

3.1.2. *Temporal - Locative*

(7) Georg hat an mindestens einem Tag in jedem Sandkasten
George has on at_least one day in every sand_box
Krieg gespielt.
war played
'Georg played war in every sandbox on at least one day.'

(8) a. Georg hat an mindestens einem Tag in jedem Sandkasten Krieg gespielt.
∃ (time) ∀ (place)
∀ (place) ∃ (time)
b. Georg hat in mindestens einem Sandkasten an jedem Tag Krieg gespielt.
∃ (place) ∀ (time)
∀ (time) ∃ (place)

Here in both cases I get scope ambiguity. (8a) could mean that there was a special day on which George played war in every sand box. But it could also mean that for each sand box there was at least one day in which he played in it.

(8b) reveals the analogous ambiguity. I get the interpretation that there is at least one sand box in which George played war every day and that there is for each day at least one (maybe different) sand box in which he played.
Does this mean that Locatives and Temporals belong to the same class? Let's look at the couplet with the universal quantifier coming first.

(9) a. Georg hat in jedem Sandkasten an mindestens einem Tag Krieg gespielt.
∀(place) ∃ (time)
∃ (time) ∀ (place)
b. Georg hat an jedem Tag in mindestens einem Sandkasten Krieg gespielt.
∀ (time) ∃ (place)
?? ∃ (place) ∀ (time)

This time I get a clearer contrast. Only (9a) is clearly ambiguous. In (9b) the reverse scope interpretation is much less available than the direct scope interpretation, though not totally excluded.

The fact that in some couplets only one order gives rise to scope ambiguity and in others there is only some asymmetry, raises the question of the validity of the test. In order to get a significant result I had to take some precautions:

1. I tested all possible combinations (91) of the thematic roles. For each pair of thematic roles I compared two couplets, one with the existential quantifier always to the left and thematic roles exchanged and the other with the universal quantifier to the left. This should show whether the resulting hierarchy is transitive.

2. I had to give a precise definition of "asymmetry". Each judgement was furnished with an evaluation. I concentrated on the comparison of the pairs. An interpretation got a "*" if it was not available at all (e.g. reverse scope interpretation in (6d)). If the reverse scope interpretation was only marginally available it was furnished with "??" (e.g. (6a) and (9b)). If I got only an asymmetry in availability (meaning in both sentences of a couplet the reverse scope is available but in one of them less available) I gave the less available interpretation a "?". I assigned a number to each of the symbols: "?" evaluates to "1", "??" to "2" and "*" to "3". In some of the cases the reverse scope interpretation was prevalent ((6b) and (6c)). These interpretations were indicated with bold face. The equivalent number in this case is "1", otherwise "0".

Table 1: Symbols of judgement evaluation

Symbol	Explanation	Numeric Evaluation
?	Reverse scope interpretation available, but more marked than the reverse scope in the partner sentence of the same couplet	1
??	Reverse scope interpretation marginally available	2
*	Reverse scope interpretation not available	3
bold	Reverse scope prevalent	1

In order to quantify the judgement I assigned a number to each pair of thematic roles. It is the sum of the elements of a quadrupel of numbers which consists of:

1. The number of question marks in the first couplet (the one with the existential operator in front), counting the "*" as "3".
2. The number of question marks in the second couplet (the one with the universal operator in front).
3. The number "1", if in the first couple in one of the sentences the reverse scope interpretation was salient, otherwise "0".
4. The number "1", if in the second couple in one of the sentences the reverse scope interpretation was salient, otherwise "0".

The resulting number is the sum of these four numbers. In the previous examples we get:

For the pair Matter – Temporal:
Result(QS) (2,3,1,1) Σ = 7 Temporal > Matter

For the pair Temporal – Locative:
Result(QS): (0,2,0,0) Σ = 2 Temporal > Locative

3.2. *Informational Focus*

A well known property of the German Mittelfeld is the fact that among two constituents the one behind can always bear informational focus, i.e. be understood as answer to a constituent question, while the one in front can bear it only when base generated higher. (Lenerz 1977). This works especially well for indirect and direct objects. Take the following base sentence:

(10) Ich habe dem Kassierer das Geld gegeben.
I have (the cashier)+DAT the money given
'I gave the money to the cashier.'

If we question the indirect object, sentences with two possible word orders are acceptable answers:

(11) Wem hast du das Geld gegeben?
'To whom did you give the money?'

(12) a. Ich habe *dem Kassierer* das Geld gegeben.
b. Ich habe das Geld *dem Kassierer* gegeben.

If, however, the direct object is questioned, only the word order with the direct object following the indirect is acceptable as an answer.

(13) Was hast du dem Kassierer gegeben?
'What did you give to the cashier?'

(14) a. Ich habe dem Kassierer *das Geld* gegeben.
b. ?? Ich habe *das Geld* dem Kassierer gegeben.

The indirect object with informational focus can be positioned before or after the direct object; therefore we take it to be generated higher.

If PPs realising different thematic roles were base generated in different position this test should give rise to an asymmetry amongst the two possible orders.

I start with a base sentence having a Benefactive and a Temporal:

(15) Donald hat am Dienstag für Georg gelogen.
Donald has on Tuesday for George lied
'Donald lied for George on Tuesday.'

If I question the Temporal I get two possible answers:

(16) Wann hat Donald für Georg gelogen?
'When did Donald lie for George?'

(17) a. Donald hat für Georg *am Dienstag* gelogen.
b. **Donald hat *am Dienstag* für Georg gelogen.**

But if I question the Benefactive, putting it in front of the Temporal, the sentence becomes odd:

(18) Für wen hat Donald am Dienstag gelogen?
'Who did Donald lie for on Tuesday?'

(19) a. Donald hat am Dienstag *für Georg* gelogen.
b.?? Donald hat *für Georg* am Dienstag gelogen.

In analogy to the above example we can deduce that Temporals are base generated higher than Benefactives.

As in the case of the Quantifier Scope Test the results were not always clear cut yes/no distinctions, although an asymmetry was always detectable. Again I quantified the judgements. If it was not possible to have the questioned constituent in front, a sentence was marked by a "*". If it was only marginally possible it got a "??". If there was just an asymmetry, i.e. the positioning of the questioned element in front of the other was possible but less acceptable than in the partner pair, I gave it a "?".

The evaluation of a pair of thematic roles consists of a pair of numbers and their sum. The first element of the pair equals to the number of question marks, again counting the "*" as "3". The second element of the pair equals to "1" if the focussed element is preferred in first position (marked in the sentences in bold face).

Summing up the two numbers gives the strength of the judgement. In the above example we get:

Result(IF) (2,1) $\Sigma = 3$ Temporal > Benefactive

3.3. *Pair List Reading*

This test is another application of scope ambiguity, this time between an interrogative operator and a universal quantifier (proposed by May 1988; see also Bruening 2001).

If the interrogative is generated above the quantifier and moves up to the left periphery, it always has the quantifier in its C-command. It allows only one possible answer containing the universal quantifier:

(20) a. Who reads all the books?
b. John reads all the books.

But if the interrogative is base generated below the quantifier and moves across it to its surface position, we get scope ambiguity. In the first case the *wh*-element is interpreted as taking scope over the quantifier, as in the above case. We expect only one simple answer:

(21) a. Which book did all the boys read?$_1$
b. All the boys read "The Minimalist Program"

In the second case, the quantifier is interpreted as taking scope over the trace of the interrogative. Now the answer is a list of pairs:

(22) a. Which book did all the boys read?$_2$
b. Bob read "Aspects",
c. Bill read "Barriers" and
d. John read "The Minimalist Program".

Applying this test to modifying PPs gave even clearer results than the other two tests.

If we take the combination of Comitative and Temporal and question the Comitative we get two types of answers. A simple one with the universal quantifier and a list of pairs:

(23) Mit welchem Freund hat Georg in jedem Jahr Krieg gespielt?
'With which friend did George play war every year?'

(24) a. Georg hat in jedem Jahr mit Tony Krieg gespielt.
'George played war with Tony every year.'

b. Georg hat 2002 mit Tony und Gerhard gespielt,
2003 mit Tony und José
2004 mit Tony und Silvio.
'George played with Tony and Gerhard in 2002, with Tony and José on 2003 and with Tony and Silvio in 2004.'

If the Temporal becomes the wh-element and the Comitative has the universal quantifier, the pair list reading becomes unavailable:

(25) Wann hat Georg mit jedem Freund Krieg gespielt?
'When did Georg play war with every friend?'

(26) Georg hat 2002 mit jedem Freund Krieg gespielt.
'George played war with every friend in 2002.'

4. The Results

During the research four questions became relevant:

1. Do all three tests result in a linear order?
2. Do all three tests give the same result?
3. What exactly is the resulting order?
4. Does the weighting give some clue?

4.1. *Do all three tests result in a linear order?*

A relation ">" is resulting in a linear order if it is

a) transitive

If A > B and B > C then A > C

b) antisymmetric

If A > B then not (B > A)

c) total
For all possible pairs (A,B) there is a relation between them so that either A > B or B > A.

All three tests resulted in nearly perfect linear order. The only deviations were the following:

(i) Deviations from Transitivity

Only the Pair List Reading Test gave a slight deviation from transitivity. It gave

Means of Transport > Malefactive
Malefactive > Instrumental
Means of Transport = Instrumental

(ii) Deviations from Antisymmetry

There were few cases where there could not be detected an asymmetry between two thematic roles.

In the Quantifier Scope Test:

Path = Means of Transport

In the Informational Focus Test:

Instrumental = Path
Instrumental = Means of Transport

In the Pair List Reading Test:

Instrumental = Means of Transport (see also above)

All these deviations concern the same low part of (Path / Means of Transport / Instrumental). This could indicate that they do not really constitute different thematic roles but occupy the same position. Semantically, Means of Transport and Instrumental are quite similar.

(iii) Deviations from Totality

Matter and Means of Transport are not compatible. Means of Transport needs a motion verb which seems to be incompatible with a Matter modifier. Furthermore, there is a problem of having a Goal and a Means of Transport modifier together. In this case, motion verbs tend to take the Goal as a (selected) complement. I wanted to avoid mixing complements with modifiers. So this is not a real incompatibility of thematic roles.
All together, the above deviations can be reduced to few cases which can be explained by model of a hierarchy of functional projections above VP.

4.2. *Do all tests give the same results?*

Table2: Resulting orders

QS	PLR	IF
Evidential	Evidential	Evidential
Temporal	Temporal	Temporal
Locative	Locative	Locative
Comitative	Comitative	Comitative
Benefactive	Benefactive	Benefactive
Reason	Reason	Reason
Source	Source	Source
Goal	Goal	Goal
Malefactive	Malefactive	Malefactive
Path /Means	Instrumental / Means	Instrumental/Means
Instrumental	Path	Path/Instrumental
Matter	Matter	Matter
Manner	Manner	Manner

As can be seen all three tests give the same order, again with the exception of the region of Path / Means of Transport / Instrumental. Therefore, I would answer the question with a clear yes. This leads directly to the answer of the third question.

4.3. *What exactly is the resulting order?*

Evidential > Temporal > Locative > Comitative > Benefactive > Reason > Source > Goal > Malefactive > Instrumental / Means /Path > Matter > Manner

4.4. *Does the weighting give some clue?*

The most surprising result was the observation that the judgement about the asymmetry was stronger the further away two elements in the hierarchy were. The evaluation number can be interpreted as a measurement of distance. The thematic roles cannot be grouped into classes where members of the same class behave less asymmetrically with respect to the test and members of different classes have sharper distinction. The sharpness of the judgement increases gradually with the distance.

This becomes clear when taking the average over all distances from the lowest element Manner. It is defined by:

$$n_{Ave}(TR1) = \Sigma_{nTR2}(Distance(TR2, TR_{Ref}) - Distance(TR2,TR1)) / n_{hits}$$

The average distance of a certain thematic role TR1 is evaluated by taking for each other thematic role TR2 the distance of this role to Manner (TR_{REF}) minus the distance between TR1 and TR2 and summing all up. The interesting finding is that this results in exactly the same hierarchy as revealed by the individual tests.

An interpretation of this effect can be achieved if we assume that in order to scramble a lower PP across a higher there are (at least) two different derivations, one in order to reverse scope and another for focus effects.

The above tests detect always for only one effect, either scope or focus. If movements existed only for scope reasons, the Quantifier Scope and Pair List Reading Tests would give sharp yes or no results. But there can be additional movements for focus reasons as can be seen for the Informational Focus Test. This explains the remaining interpretations of the reverse kind.

The fact that the sharpness of the judgement increases with the distance indicates that the "scrambled" PP has to do more work, it has to move around all intervening functional projections. This in turn shows that these intervening projections always exist, even if not represented overtly in the pronounced string.

A few remarks to the validity of this hierarchy. When doing the test I tried carefully to avoid seeing in the data what I expected, especially when having evaluated a pair of thematic roles with another test. Of course the judgements especially between neighbours can be subtle. But the order between one thematic role and the one following its direct neighbour in the above hierarchy seems to me very clear.

References

Bruening, Benjamin. 2001. "QR obeys superiority: Frozen scope and ACD". *Linguistic Inquiry* 32.2: 233-273.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective* New York/Oxford, Oxford University Press.

Cinque, Guglielmo. 2000. "'Restructuring' and functional structure". *University of Venice Working Papers in Linguistics,* 11: 45-127.

Frey, Werner and Karin Pittner. 1998. "Zur Positionierung der Adverbiale im deutschen Mittelfeld". *Linguistische Berichte* 176: 489-534.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen.* Tübingen, Gunter Narr Verlag.

May, Robert. 1988. “Ambiguities of quantification and wh: A reply to Williams”. *Linguistic Inquiry* 19: 118-135.

Pollock, Jean-Yves. 1989. “Verb movement, universal grammar and the structure of IP”. *Linguistic Inquiry* 20: 365-424.

The acquisition of Italian questions by Tamil speakers[1]

Jacopo Garzonio and Sandra Gracci
University of Padua - University of Pisa

1. Introduction

Formal approaches to language like Generative Grammar assume that linguistic variation is highly constrained and that all languages share the same fundamental structure, or *Universal Grammar* (UG). The principles that make up UG in turn are supposed to be active not only in the stable language of adults, but also during the acquisition of language itself. But while First Language Acquisition (see Friedemann-Rizzi, 2000, among others) has been extensively studied to see how much it conforms to the constraints imposed by UG, less is known about the role played by universal principles in Second Language Acquisition and their interaction with interference from the first language of the learner (see Flynn-O'Neil, 1988, on this topic).

In this paper we will study the syntax of interrogative clauses and *wh* items in the interlanguage of Tamil learners of Italian, living in Italy.[2] In particular, we will focus on the acquisition of *wh* movement, since Tamil does not have overt movement of *wh*-items. Several hypotheses were proposed to account for the parametric difference between *wh*-movement and *wh in-situ* languages; we will base our analysis on the *Clausal Typing hypothesis* proposed by Cheng (1997), according to which languages can check the interrogative feature either through *wh*-movement or base-

[1] The authors would like to thank Cecilia Poletto and Federico Damonte for their precious help and advice. We also thank an anonymous reviewer, who gave us some useful suggestions. However, we are responsible for any errors.

[2] Tamil is a Dravidian language, spoken in Southern India (Tamil Nadu) and in the island of Ceylon.

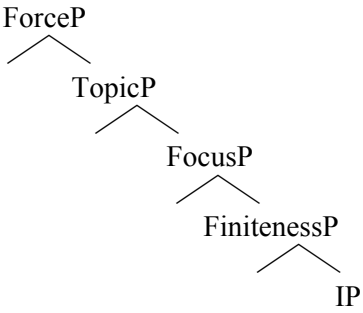
generation of an interrogative particle. We will also adopt Rizzi's (1997) split-CP hypothesis and assume that the interrogative feature is spelled out in the Focus Phrase in Italian. We will argue that our findings show that both UG principles and interference from the native language play a role during second language acquisition (as also claimed by Flynn-O'Neil, 1988). Moreover, it will be shown that interference from the first language is restricted to formal properties of the *wh* items and that the universal structure of the left periphery is accessible during the acquisition process.

The paper is organised as follows: section 2 presents the theoretical background we assume and introduces briefly Italian and Tamil interrogative structures, pointing out some major differences between the two languages; section 3 describes the data collection techniques we used and contains some essential information about our informants; section 4 presents the results of our survey, and some generalizations that can be made about them. Some peculiar phenomena will be pointed out and investigated in more detail; section 5 provides our analysis of these phenomena. We argue that UG principles are visible in the interlanguages of Tamil learners and that interference from the first language is restricted to the formal properties of the *wh* items; section 6, lastly, provides some conclusive remarks.

2. Theoretical Background

It is well known that sentences can have different functions. Some of these functions are overtly marked by different morpho-syntactic mechanisms. Together, these mechanisms are labelled “Clause Typing” (see, for example, Cheng, 1997). Informally, the “type” a clause belongs to indicates its illocutionary Force, such as declarative or interrogative (see, among others, Chomsky, 1995: §4). In his seminal work on the “left periphery” of the sentence, Rizzi (1997) argues that the Force of a sentence is encoded in the left periphery (and more precisely in the upper part of it): he identifies a functional projection (which he calls Force Phrase), where several features have to be checked. These features, in turn, determine the Clause Type of the sentence. In (1) we report the map of the left periphery as proposed by Rizzi (1997) and slightly modified by Benincà (2001):

(1)



Following Chomsky's (1995) Feature Checking Requirement, we assume that the interrogative clausal type is related to an interrogative feature [+Q], which must be checked in order to license an interrogative interpretation. In some languages, for example Italian, this feature must be checked in a low projection in the left periphery, namely Focus Phrase (Rizzi, 1997: §5), as it can be seen in (2), where a left dislocated phrase in [Spec, TopicP] precedes an item in [Spec, FocusP]:

(2) Il biglietto quando lo compri?
the ticket when it buy-2sg
'The ticket, when are you buying it?'

As it has been shown by Cheng (1997) for *wh* questions, languages vary as to the way they check the interrogative feature [+Q]. Some languages use *wh* movement to the left periphery (*wh* movement languages), while others allow the *wh* item to stay *in situ* (*in situ* languages). According to Cheng, *in situ* languages mark questions by means of specific interrogative particles (which can be phonologically null). This hypothesis is based on the observation that all *in situ* languages have phonetically realized interrogative particles in *yes/no* questions.[3]

These considerations are summarized in Cheng's (1997: 22) *Clausal Typing* Hypothesis:

[3] Baker (1970: 207), who first introduced the concept of the [Q] feature, considers English *if* and *whether* as interrogative particles and notes that there is a close connection between the position of *yes/no* particles and syntactic movement of *wh* elements. However, his conclusion that only languages which place *yes/no* particles in clause-initial position permit *wh*-movement turned out to be wrong: it seems that all *in situ* languages have some morphologically overt process to mark *yes/no* questions.

> Every clause needs to be typed. In the case of typing a -question, either a -particle in C° is used or else fronting of a -word to the Spec of C° is used, thereby typing a clause through C° by Spec-head agreement.

In *wh-in situ* languages overt *wh* movement is not necessary, because the interrogative feature does not have to be checked before Spell-Out. This variation is explained by the parametrisation of features' strength: in some languages [+Q] is strong and must be checked before Spell-Out; on the contrary, if [+Q] is weak, *wh* movement takes place at LF.

2.1. *Comparing Italian and Tamil Interrogatives*

Italian is a *wh* movement language, while Tamil is an *in situ* (verb final) language. Italian is very similar to English, as it can be seen in (3), where the *wh* word is fronted:

(3) Cosa fai per il tuo compleanno?
what do-2sg for the your birthday
'What are you doing for your birthday?'

In Tamil, a declarative sentence and its correspondent question have the same surface word order (Savio, 1991):

(4) a. avan sātam sāppiṭṭān.
he rice ate-3sg
'He ate some rice'
b. avan enna sāppiṭṭān?
he what ate-3sg
'What did he eat?'

In the Italian sentence (3) the *wh* direct object is raised to [Spec, FocusP], and the inflected verb *fai* is raised to Focus° (as argued by Rizzi, 1997); in the Tamil question (4b), the *wh* word *enna* stays *in situ* and the [+Q] feature is checked by a null particle. We assume the existence of such null interrogative particle, since in Tamil questions are marked by special morphemes on the verb or the interrogated argument. In (5) an example of a *yes/no* interrogative sentence is given; the final verb shows the interrogative suffix -ā:

(5) Rāman vandār-ā?
Raman came-*yes/no*Prt
'Did Raman come?'

Similarly, *wh* interrogative words are marked by the prefix *e-*, which distinguishes them from correspondent adverbs, adjectives and pronouns, as it can be observed in table 1, where some interrogative and demonstrative pronouns are reported:

Table 1: Interrogative and demonstrative pronouns

PRONOUNS		
DEMONSTRATIVE		*INTERROGATIVE* ***e-***
Proximate **i-**	Distant **a-**	
Idu "this thing"	*Adu* "that thing"	*Edu* "which thing?"
Ivaru "this person"	*Avaru* "that person"	*Evaru* "which person? who?"
Ivan "this male"	*Avan* "that male"	*Evan* "which male?"

(Schiffman 1999: 119)

While Tamil uses morphological means in the formation of *wh*-questions, Italian resorts to both morphological and syntactic means. It should be pointed out that Tamil is different from *in situ* languages taken into consideration by Cheng (1997), as Chinese: in Tamil there are not independent interrogative particles. This fact shows that a third type of language should be added to Cheng's typology: *in situ* languages with interrogative bound morphemes and null particles in CP.

3. Methodological Remarks

In this section we describe the data collection techniques we used and provide some information about the Tamil speakers interviewed for this study. The data collection took place in a series of encounters with the informants during the period from May to September 2003.

3.1. *Typology of the Elicitation Tasks*

The data were obtained by asking informants to perform three different types of linguistic tasks. These in turn were written down or recorded on tape.

TASK 1: the first task consisted in the translation in Italian of Tamil sentences;

TASK 2: the second task asked for the question corresponding to a given answer (in Italian), such as the one in (6), the expected output being "Perché sei tornato?" ('Why did you return?'):

(6) Sono tornato perché è tardi.
'I returned because it is late.'

TASK 3: in the third task informants had to complete a short dialogue in Italian where some questions were omitted. This type of linguistic task is different from the previous one because in this case the omitted questions are inserted in a context, which helps the informants to reconstruct the missing question.

3.2. *The informants*

The data were obtained from five Tamil speakers who have been living in Italy for different periods of time. The informants reside in three different cities of Central Italy (Bologna, Florence and Pisa). The main information about the informants are shown in table 2. In order to respect the privacy of our informants, nicknames are used in place of real names.

Table2: Informants

NICKNAME	SRI	RAJA	SUJE	SUJATHA	KABILAN
SEX	male	male	female	female	male
AGE	35	34	28	27	23
PLACE OF ORIGIN	Sri Lanka (Jaffna)	Sri Lanka (Jaffna)	Sri Lanka (Jaffna)	Tamil Nadu (Nagercoil)	Tamil Nadu (Madras)
JOB	employee	workman	housewife	shop-assistant	engineer
TIME IN ITALY	13 years	10 years	5 years	5 years	1,5 years
COURSE OF ITALIAN	No	No	No	Yes	Yes
OTHER LANGUAGES	English (a little)	English	English (a little)	English	English, Hindi, Kannada

4. Results of the investigation

Before presenting the results of our study, it should be noticed that the aim of our research is that of identifying the specific properties of the syntax of Italian interrogative sentences produced by Tamil speakers. It is beyond the scope of this paper to give a complete picture of the different stages of acquisition of Italian interrogative syntax by Tamil speakers. For this reason we did not carry out a longitudinal study, and we will not try to qualify the phenomena we observed as belonging to an early (or late) acquisitional stage.

4.1. Main Divergences from the Italian Interrogative Structure

About 10% of produced sentences can be considered totally incorrect. These sentences were produced mainly by speakers who have been living in Italy for a short period of time or who have few contacts with Italian people.

There are two main types of mistakes:

1. The wrong *wh* item is used;
2. The *wh* item is left *in situ*.

First we analyse some sentences with a wrong *wh* item. In (7a) and (8a) examples of this kind of error are given: (7a) was produced as translation (task 1) of the Tamil sentence (7a), while (8a) was produced as translation of (8b):

(7) a. Lei dove suo paese? (SUJE)
you (polite) where your (polite) country
b. Nīṅga eppō ūrukkup pōrīṅga?
you when journey-dat go-2pl
'When do you make a trip (to your country)?'

(8) a. Dove andato Tiruchi? (SUJE)
where gone Tiruchi'
b. Tiruccikku eppadi pōrīṅga?
Tiruchi-dat how go-2pl
'How do you go to Tiruchi?'

In these examples the *wh* item *dove* is used instead of *quando* 'when' and *come* 'how'. An other similar case is shown in (9), which is produced as a question (task 3) to the given answer (10):

(9) Quando volta vai in India? (SUJE)
when time go-2sg in India

(10) Ci torno due volte l' anno.
there return-1sg two times the year
'I return there two times in a year.'

Here *quando volta* is used instead of *quante volte* 'how many times' (but this output could be motivated by a phonetic interference). Another kind of error, where L1 structures play a more evident role, is the *wh in situ* construction, which is presented below (11a), which was produced as translation (task 1) of (11b):

(11) a. Mio pena è dove? (KABILAN)
my pen is where
b. En pēnā eṅgē irukkudu?
my pen where is
'Where is my pen?'

It should be noticed that in (11) the verb-final structure of Tamil is not transferred to the Italian translation. This means that at this stage of acquisition the position of the inflected verb in Italian has already been learned, and the *wh* item occupies the position of a locative PP in an unmarked declarative order. Similar errors occur only in sentences produced as translations (task 1).

4.2. *Specific Phenomena*

We observed two peculiar phenomena in our corpus. The first can be observed in sentence (12):

(12) Comsichiama tu? (RAJA)
how-oneself-call-3sg you
'What's your name?'

In this sentence the *wh* item and the verb (with the reflexive clitic) are written by the informant as one word. This is a written example, but this phenomenon seems quite frequent also in the spoken data. The informant seems to consider the *wh-item*-verb complex as a single unity.

The second phenomenon is present in 20% of the sentences and thus can be considered very frequent: a topicalised DP (mostly the subject or the

direct object of the sentence) appears before the *wh* item and the inflected verb:[4]

(13) a. Mia pena tove sono? (RAJA)
my pen where are
'Where is my pen?'
b. Tua fratello qando parte? (RAJA)
your brother when leave-3sg
'When does your brother leave?'
c. Lei che lavoro? (SUJE)
you (polite) what job
'What's your job?'

In these sentences the position of the *wh* word can be considered correct, since, as we have shown above in example (2), a topic can precede a *wh* item, in Italian as in many other languages.

5. Analysis

Hawkins (2001) assumes that the clause structure of interlanguages is deprived of the CP layer, at least in the first stages of acquisition. We will show that this is not true, at least in the Italian clause structure of our informants. Instead, the data presented in the previous section lead us to think that UG, and the hierarchy of the left periphery in particular, are active during the acquisitional process.

Consider the examples in (13): the topicalised DP always precedes the *wh* item. Without a precise configuration of the CP layer, this is suprising, but, if we assume the left periphery mapping of Rizzi (1997) and Benincà (2001), the reason for such order is clear: a topic (either a Hanging Topic or a Clitic Left Dislocated phrase) must precede a *wh* word in [Spec, Focus]. This is the reason why in Italian (14a) is acceptable but (14b) is ungrammatical:

(14) a. Il libro quando lo compri?
the book when it buy-2sg
'The book, when are you buying it?'
b.*Quando il libro lo compri?

Our informants make several types of morphological and syntactic mistakes,

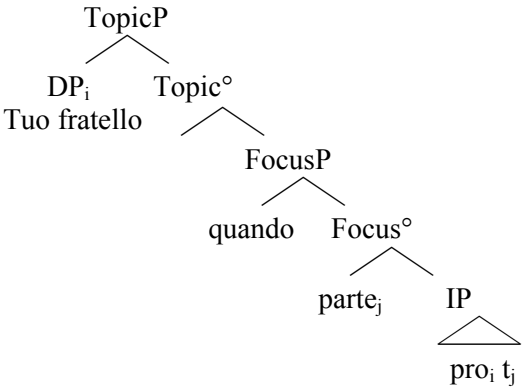
[4] An anonymous reviewer points out that the order Subj-*wh*-Verb in (14) reflects the Tamil word order. So, even if we analyzed the initial DP as a topic, it is not possible to exclude the influence of the Tamil word order in such examples.

but they never put a topic after a *wh* item. Even if our informants are not able to produce some complex structures involving the left periphery, such as cleft sentences, they seem to have a correct "map" of the order of elements in the CP domain. We leave open the question whether the topicalised phrases in (14) are Hanging Topics or Clitic Left-Dislocated phrases, but it should be noticed that they correspond to either the subject or the object of the verb: probably this is due to the fact that DPs are more naturally topics than adjunct PPs.

The structure of sentences like (13b), repeated below as (15) without orthographical errors, is shown in (16):

(15) Tuo fratello, quando parte?

(16)



The DP *tuo fratello* is located in [Spec, TopicP], while the *wh* item *quando* and the inflected verb *parte* are in [Spec, FocusP] and Focus°, respectively. They are in a Spec-Head configuration, and therefore the [+Q] feature in Focus° is checked. Here we will not address the question whether the topic DP is raised from IP or it is base-generated in [Spec, TopicP]; in the former case it leaves a trace, in the latter it is coreferential with a *pro* in subject position (as shown in (16)).

The other peculiar phenomenon, exemplified by sentences like (12), repeated as (17), is also worth discussing:

(17) Comsichiama tu?
how-oneself-call-3sg you
'What's your name?'

The fact that the informant writes the *wh* word and the inflected (reflexive) verb as one word means that such sentences are produced through a

reanalysis of the *wh* item as a *wh* agreement morpheme on the inflected verb in a way parallel to the strategy used in Tamil for *yes/no* questions (see (5)). As we said above, in Tamil there are no free phonetically realized interrogative morphemes, be they *wh* items or interrogative particles. Such examples, triggered by L1 interference, may be considered as evidence of an intermediate stage of acquisition. In other words, since in their mother language interrogative sentences are always marked by bound morphemes, our informants interpret the *wh* word *come* 'how', which appears before the verb, as an interrogative prefix.

These facts show that what is transferred from L1 to L2 in second language acquisition are the formal properties of the elements rather than the clause typing strategy. Tamil learners of Italian learn very quickly that *wh* items and verbs have to be moved to CP in interrogatives (as pointed out above, *wh* items *in situ* are very rare in our corpus), but they treat *wh* words as bound morphemes.

6. Conclusions

At the end of this study, it is possible to say that Tamil speakers acquire Italian interrogative strategies very soon in the acquisition process. *Wh* words, even if not always the correct ones, are fronted in most sentences. Furthermore, even when the *wh* word is not fronted, informants correctly move the inflected verb past the direct object, both in declaratives and interrogative sentences. The verb final configuration of Tamil is not transferred to Italian. It should be added that also Italian interrogative intonation (which is sometimes the only way to distinguish a *yes/no* question from a declarative) is acquired without any difficulty.

Two kinds of interference phenomena can be observed. The first is related to Tamil word order: as pointed out above in section [3], sometimes *wh* words appear *in situ.* The second one, which is more frequent and relevant, is related to Tamil affixal morphosyntax: Italian *wh* items are treated as interrogative morphemes incorporated onto the verb.

The fact that very often a topic DP precedes the *wh* item reveals that informants have access to the internal structure of the CP and know that in Italian [+Q] is checked in FocusP. This leads us to conclude that at least some of the principles of UG are active during second language acquisition. It also seems to indicate that errors made at intermediate stages of the acquisition process are caused by the lack of movement operations rather than gaps in the structure of the clause. Second language learners have a CP layer in their syntactic representations, but they lack the morphological and lexical means to produce complex structures involving this representation. This is the reason why cleft sentences are not produced and why topics lack

resumptive clitics, as it should be the case in clitic left dislocation constructions.

References

Baker, C. Lee. 1970. "Notes on the Description of English Questions". *Foundations of Language* 6:197-219.

Benincà, Paola. 2001. "The Position of Topic and Focus in the Left Periphery". In G. Cinque and G. Salvi (eds.) *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*. 39-64. Amsterdam, Elsevier.

Cheng, Lisa Lai-Shen. 1997. *On the Typology of Wh-Questions*. New York/London, Garland.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA, MIT Press.

Flynn, Suzanne and Wayne O'Neil. 1988. "Introduction". In S. Flynn and W. O'Neil (eds.) *Linguistic Theory in Second Language Acquisition*. 1-24. Dordrecht, Kluwer.

Friedemann, Mark Ariel and Luigi Rizzi (eds.). 2000. *The Acquisition of Syntax*. Harlow, Longman.

Hawkins, Roger. 2001. *Second Language Syntax*. Oxford/Malden, Blackwell.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In L. Haegeman (ed.) *Elements of Grammar: Handbook in Generative Syntax*. 281-337. Dordrecht, Kluwer.

Savio, S. Dominic. 1991. "*Wh*-Questions in Tamil". *CIEFL Occasional Papers in Linguistics* 3:55-67.

Schiffman, Harold F. 1999. *A Reference Grammar of Spoken Tamil*. Cambridge, Cambridge University Press.

The left periphery of Hungarian exclamatives

Anikó Lipták
ULCL/Leiden University

1. Introduction[1]

Wh-exclamatives are quite similar to interrogatives crosslinguistically, as far as their syntax is concerned. This similarity does not come as a surprise though, given that both *wh*-exclamatives and interrogatives involve an A-bar *wh*-dependency. Viewed from this perspective, systematic differences between exclamatives and questions are of more descriptive and theoretical interest for the study of sentence mood types as well as *wh*-constructions in general.

In this paper I undertake the description and analysis of Hungarian exclamative clauses with respect to word order variation found in exclamatives and questions. I will show that the distribution of *wh*-phrases in the two constructions differs in ways that to my knowledge have not yet been found in other languages. The prime variation we will be interested in concerns the free availability of verb–preverb inversion in some exclamatives, illustrated in (1):[2]

[1] I thank the audience of the 30th Incontro di Grammatica Generativa as well as Hans Broekhuis and István Kenesei for helpful comments on the material presented in this paper.

[2] Hungarian ortography does not require a dash between the uninverted preverb and the verb, but a dash will be used throughout this paper for transparency reasons. The glosses in this article are as follows: ' = emphatic stress; ACC = accusative; Adj = adjective; Adv = adverb; E = exclamative; N = noun; PL = plural; PV = preverb(al element); Q = question; REL = relative morpheme; V = verb. Nominative case is not glossed. Subject person and number morphemes as well as object definiteness morphemes on verbs (1-3SG/PL, DEF/INDEF) are glossed only when relevant. SMALL CAPS on lexical words indicate identificational focus.

(1) a. Mennyi könyvet **olvastál el**!
how.many book-ACC read-2SG PV
'How many books you read!'
b. Mennyi könyvet **el-olvastál**!
how.many book-ACC PV-read-2SG
'How many books you read!'

(1a) illustrates an exclamative sentence with verb–preverb inversion, and (1b) without. While the two sentences clearly differ in word order, the basic exclamative meaning is the same in both, as reflected in the translations. The same freedom of inversion is not available in questions. (1a) is string-identical to a well-formed question, but (1b) is not. This shows that more positions are available to exclamative *wh*-phrases in exclamatives than to question words in questions. The reason behind this has to do with the way exclamative meaning is expressed in the sentence. Unlike question-words, which are necessarily interpreted as contrastive focus and occupy the contrastive focus positions, exclamatives are not interpreted this way and as a result have a greater freedom of placement.

This paper consists of two parts. The first part is devoted to description and the second part to analysis. The descriptive part will start out with a brief introductory section on exclamatives (section 2). The next section will present the Hungarian data that form the empirical basis for the rest of the discussion (section 3). In the analytical part, I first identify the positions that are substantiated in Hungarian exclamatives, using the evidence of distributional facts and meaning differences that can be observed in the different constructions (section 4). Section 5 will compare the syntax of questions and exclamatives, and section 6 will summarize the results.

2. Exclamatives

2.1. *The exclamative sentence type*

Exclamative utterances are a minor sentence type (Sadock and Zwicky 1985). They differ from declaratives in that exclamatives are not assertive, and they are distinct from interrogatives and imperatives in that they do not elicit an answer or an action from the listener. Exclamatives are utterances with which the speaker expresses astonishment or surprise about something that he takes to be a fact. They are expressive utterances conveying an emotional reaction about something that is unexpected or extraordinary.

What the speaker finds surprising can range from a property of an entity to surprise about a whole state of affairs. Accordingly, there are

exclamatives which contain a sentential constituent (mostly a nominal or adjectival phrase) as the exclamative item, and exclamatives which are about a proposition. This paper will only be concerned with the first type of exclamatives, thus the term "exclamative" below will refer to this type only. For more on exclamatives in general, see among others Bennis (2000), Zanuttini-Portner (2003), Villalba (2003).

2.2. *Types of exclamative phrases in Hungarian*

Hungarian exclamatives have two major subtypes depending on what expression is used as the exclamative constituent: a *wh*-phrase (*wh-exclamatives)* or a DE-phrase (*de-exclamatives*).[3] In both types, the exclamative constituent fronts to a preverbal (left peripheral) position. The exact nature of this position will receive close attention in the coming sections:

(2) a. Milyen rohadt hideg van! *wh*-exclamative
how rotten cold is
'How awfully cold it is!'
b. Hányszor elmondta ezt a viccet Áron!
how.often PV-told this-ACC the joke-ACC Áron
'How often Áron told his joke!'
c. Miket tud ez a gyerek!
what-PL knows this the child
'The things this child knows!'

(3) De rohadt hideg van! DE-exclamative
DE rotten cold is
'How awfully cold it is!'

[3] Next to these major types, there are minor types of exclamatives as well, which are not discussed in this paper. One syntactically interesting type contains a preverbal nominal phrase in which predicate–article inversion has taken place, and the whole exclamative fronts to the preverbal focus position (for more on these, see Den Dikken and Lipták 1997):

(i) a. [$_{NP}$Hideg egy napok] voltak azok!
cold a days were those
'Those days were bitterly cold!'
b. *Azok voltak [$_{NP}$ hideg egy napok]!
those were cold a days

The other type is expressed by an isolated free relative:

(ii) Amennyit ez a gyerek megeszik!
REL-how.much-ACC this the child eats
'The amount (of food) this child consumes!'

The word *de* (literally 'but') is a degree expression and has the same distribution as *milyen (N/Adj/Adv)* 'what kind (N)', 'how (Adj/Adv)' in Hungarian.

As will be seen below, the distribution of the exclamative constituent *within* the exclamative clause shows the same regularities in *wh*- and DE-exclamatives. DE-exclamatives differ from *wh*-exclamatives, however, in their *external* distribution, first and foremost, in their compatibility with a complementizer. DE-exclamatives, unlike *wh*-exclamatives, cannot be introduced by a finite complementizer in root contexts (cf. 4) and cannot be embedded under a matrix exclamative predicate (5):

(4) *Hogy de rohadt hideg van! DE-exclamative
COMP DE rotten cold is
'How awfully cold it is!'

(5) *Elképesztő, hogy de rohadt hideg van!
incredible that how rotten cold is
'It's incredible how awfully cold it is!'

Both of these options are freely possible with *wh*-exclamatives, as (6) and (7) show: the finite complementizer can optionally introduce mono-clausal, root *wh*-exclamatives and can subordinate these under a matrix predicate as well.

(6) (Hogy) milyen rohadt hideg van! *wh*-exclamative
COMP how rotten cold is
'How awfully cold it is!'

(7) Elképesztő, hogy milyen rohadt hideg van!
incredible that how rotten cold is
'It's incredible how awfully cold it is!'

The presence of the finite complementizer in root *wh*-exclamatives (as in (6)) gives extra emotional emphasis to the exclamative utterance as a whole.
The differences between *wh*- and DE-exclamatives illustrated in this section will be not further be discussed in this article, as they are orthogonal to the internal build-up of the immediately preverbal domain, which is the locus of variation this paper is concerned with. The remainder of this paper is fully dedicated to the study of word order patterns of *wh*-exclamatives in root contexts.

3. Word order variation in Hungarian exclamatives

As pointed out above, certain *wh*-exclamatives in Hungarian can occur with two possible word orders. If the verb has a preverb, they can occur with inversion or without inversion. This is in stark contrast to *wh*-questions, which always feature inversion. To illustrate this difference, the discussion in this empirical section will start with a short exposition of the syntax of questions and proceed with the description of exclamatives.

3.1. *The point of contrast: constituent questions*

Question formation in Hungarian always involves the placement of the *wh*-phrase into the immediate preverbal position of the clause, resulting in *wh*-V adjacency, exemplified in the following:

(8) a. Hol lakik Elemér?
where lives Elemér
'Where does Elemér live?'
b.*Hol Elemér lakik?
where Elemér lives

As (8b) illustrates, nothing can intervene between the *wh*-phrase and the verb in interrogatives. If the verb has a non-referential lexical complement or *preverb* (PV for short), constituent question formation involves *inversion* between the verb and its preverb, next to the movement of the *wh*-phrase to preverbal position, as illustrated in (9) and (10):

(9) Elemér el-költözött Párizsba. declarative
Elemér PV-moved Paris-to
'Elemér moved to Paris.'

(10) Elemér hova költözött el? question
Elemér where moved PV
'Where did Elemér move?'

Verb–preverb inversion is also called preverb stranding in the literature on Hungarian, and will be given a structural account in section (4) below.

Preverbal placement and verb–preverb inversion characterize all constituent questions without exception, regardless of what constituent is questioned: bare *wh*-phrases or complex *wh*-phrases; D-linked or non-D-linked constituents. There is no difference between nominal and numerical or degree expressions, either, as evidenced by the following two sentences:

(11) Melyik kocsit adta el Adél?
which car-ACC sold PV Adél
'Which car did Adél sell?'

(12) Hányszor költözött el Elemér?
how.often moved PV Elemér
'How often did Elemér move out?'

The uniformity of (11) and (12) will become relevant when in comparison with exclamatives in further sections.

3.2. *Word order properties of exclamative wh-phrases*

In Hungarian, all *wh*-words can occur in exclamatives.[4] Just like *wh*-phrases in questions, *wh*-phrases in exclamatives obligatorily front to a preverbal position, which is a left periphery position in Hungarian:

(13) a. (Hogy) hány helyen lakott már Elemér!
COMP how.many place-ON lived already Elemér
'How many places Elemér already lived in!'
b.*(Hogy) Elemér lakott már hány helyen!
COMP Elemér lived already how.many place-ON

The preverbal position that exclamative *wh*-phrases can occupy can be of two types: in one, the *wh*-expression triggers verb–preverb inversion, in the other, it does not. It is not the case that both positions are available to all *wh*-phrases though. Some can only occur with inversion and some can occur with inversion or without. The obligatoriness of verb–preverb inversion thus splits exclamative *wh*-phrases into two types. What I will call *flexible* wh-*phrases* optionally trigger inversion. *Inflexible wh*-phrases on the other hand force inversion obligatorily. The next two sections will introduce these two classes in more detail. Before turning to these, a note is in order about the data to be presented below.

The data in this article were collected in 2002/2003 from 15 native Hungarian speakers in the form of written questionnaires. The survey revealed uniform judgements in case of some *wh*-phrases and some individual variation in case of others (indicated by % below). The most important finding of the survey was the existence of two types of

[4] According to Kálmán et al. (2001) *miért* 'why' cannot occur in exclamatives in Hungarian. In my survey, this item was judged grammatical by many speakers, although individual variation in this domain is substantial.

exclamative *wh*-constituents (flexible vs. non-flexible), to which we turn directly. This is in concert with what is reported in Kenesei et al. (1998), who also recognize the existence of two classes. My survey and Kenesei et al. (1998) are in contrast to Kálmán et al. (2001), who do not acknowledge the existence of two types of exclamative *wh*-phrases. Rather, it is claimed that all *wh*-phrases can freely occur with or without inversion (i.e. are all flexible in the present terminology).

After these introductory notes, let us turn to the illustration of the two basic types of *wh*-phrases: the flexible and inflexible class.

3.2.1. *Flexible wh-phrases*

Flexible wh-phrases can occur with or without inversion in Hungarian exclamatives:

(14) a. (Hogy) hány filmet meg-néztél!
COMP how.many film-ACC PV-watched-2SG
'How many films you watched!'
b. (Hogy) hány filmet néztél meg!
COMP how.many film-ACC watched-2SG PV
'How many films you watched!'

(15) a. (Hogy) ki mindenki el-jött az ünnepségre!
COMP who everyone PV-came-3SG the celebration-TO
'How many (different) people came to the celebration!'
b.%(Hogy) ki mindenki jött el az ünnepségre!
COMP who everyone came-3SG PV the celebration-TO
'How many (different) people came to the celebration!'

Flexible *wh-phrase*s, when occurring without inversion, are always adjacent to the verbal head. They do not tolerate any intervening material, including focus, between themselves and the verb:

(16) *??(Hogy) hány filmet BÉLA nézett meg!
COMP how.many film-ACC Béla watched-3SG PV
intended: 'How many films Béla watched (and not someone else)!'

There is a primary semantic distinction between the two word order patterns with *wh*-expressions that have an amount reading. Without inversion only

the high amount reading is possible, while with inversion, both low and high amount readings are fine (Kálmán et al. 2001)[5]:

(17) a. (Hogy) hány filmet meg-néztél!
COMP how.many film-ACC PV-watched-2SG
'How many films you watched! / *How few filmes you watched!'
b. (Hogy) hány filmet néztél meg!
COMP how.many film-ACC watched-2SG PV
'How many films you watched! / How few films you watched!'

Flexible *wh*-expressions comprise the phrases formed with the *wh*-words *mennyi (N)* 'how much/how many (N)', *hány (N)* 'how many (N)' and their case-marked forms (e.g. *hányszor* 'how often')[6], as well as phrases with the same lexical content, like *milyen sok (N)* lit. what many/much (N), 'how many/much (N)'. The flexible class also comprises *kik* 'who-PL', and to a lesser degree, any plural nominal *wh*-expression (some speakers reject these without inversion), as well as quantified *ki mindenki* 'who all', *mi minden* 'what all', which are preferred to be used without inversion by some speakers.

3.2.2. *Inflexible wh-phrases*

Inflexible wh-phrases can only occur with inversion and do not allow any intervener (including focus) between themselves and the verb:

(18) a. (Hogy) ki ment el moziba Annával!
COMP who went PV cinema-TO Anna-WITH
'The person who went to the cinema with Anna!'

[5] Further distinctions in meaning between the two orders will be mentioned in section 4 below.

[6] The only exception in this domain is *mennyire* 'how much/to what extent', and in the same meaning, *hogy* 'how (much)' when modifying a VP. These two *wh*-phrases differ from the rest of the flexible class (as well the inflexible one as well) in that they obligatorily occur *without* inversion:

(i) a. (Hogy) mennyire/hogy meg-nőtt Éva!
COMP how.much /how PV-grew Éva
'How much Éva has grown!'
b.*(Hogy) mennyire/hogy nőtt meg Éva!
COMP how.much/how grew PV Éva

In this respect, *mennyire/hogy* follow the distribution of *milyen nagyon* 'how very.much', which, due to *nagyon* being an inclusive adverb (see fn.7 below), cannot occur with inversion.

b.*(Hogy) ki el-ment moziba Annával!
COMP who PV-went cinema-TO Anna-WITH
c.*(Hogy) ki ANNÁVAL ment el moziba!
COMP who Anna-WITH went PV cinema-TO
intended: 'The person who went to the cinema with Anna (and not someone else)!'

(19) a. (Hogy) hova mentél el!
COMP where went-2SG PV
'The place you went to!'
b.*(Hogy) hova el-mentél!
COMP where PV-went-2SG
c.*(Hogy) hova PÉTER ment el!
COMP where Péter went-3SG PV
intended: 'The place Péter went to (and not someone else)!'

(20) a. (Hogy) milyen drága könyvet vettél meg!
COMP how expensive book-ACC bought-2SG PV
'How expensive a book you bought!'
b.*(Hogy) milyen drága könyvet meg-vettél!
COMP how expensive book-ACC PV-bought-2SG
c.*(Hogy) milyen drága könyvet MA vettél meg!
COMP how expensive book-ACC today bought-2SG PV
intended: 'How expensive a book you bought today (and not some other time)!'

The inflexible class of *wh*-phrases is made up of those that do not belong to the flexible class. These involve the following items and their case-marked or derived forms: *ki* 'who', *mi* 'what', *mikor* 'when', *hol* 'where', *hogy(an)* 'how', *miért* 'why', *milyen (Adj)* 'how (Adj)', *milyen (Adv)* 'how (Adv)'[7] and the singular *melyik (N)* 'which N', *milyen (N)* 'what kind of (N)'.

[7] *Milyen (Adv)* 'how (Adv)' behaves as an inflexible *wh*-phrase only if the adverb in it is not a special, so called inclusive adverb. *Inclusive* adverbs (completion and intensity adverbs as well as degree quantifiers, like *könnyedén* 'lightly', *teljesen* 'completely', *alaposan* 'thoroughly', see Kiefer 1967) obligatorily occur in preverbal position without inversion (not precisely, but approximately in *many*P, see section 3), including exclamative *wh*-phrases with these adverbs. (i) and (ii) illustrate this parallel.

(i) a. Ágnes alaposan meg-fázott. incl. Adv, indicative
Ágnes thoroughly PV-cold.caught
'Ágnes caught a thorough cold.'
b. *Ágnes alaposan fázott meg.
Ágnes thoroughly cold.caught PV

3.2.3. *The data at a glance*

This section has provided a characterization of the basic word order properties of exclamative *wh*-expressions (E-*wh*-phrases for short) in Hungarian. The syntactic distribution of these with respect to the possibility of verb–preverb inversion is summarized in Table 1.

Table 1: *Wh*-expressions in Hungarian exclamatives

		with inversion	**without inversion**
flexible E-*wh*	*mennyi* 'how many/much'	✓	✓
	hány 'how many'	✓	✓
	plurals (*kik* 'who-PL', *mik* 'what-PL')	✓	%
	quantified wh (*ki mindenki* 'who all')	%	✓
inflexible E-*wh*	*ki* 'who'	✓	*
	mi 'what'	✓	*
	mikor 'when'	✓	*
	hol 'where'	✓	*
	hogy(an) 'how'	✓	*
	melyik (N) 'which (N)'	✓	*
	miért 'why'	✓	*
	milyen (N/Adj) 'what (N), how (Adj)'	✓	*

The table helps us capture the following generalization: putting aside quantified *wh*-expressions and adverbials mentioned in footnote 7, all *wh*-phrases can occur with inversion, and a subset of them, namely the flexible class, can also occur without.

(ii) a. (Hogy) milyen alaposan meg-fázott Ágnes! incl. Adv, excl.
COMP what thoroughly PV-cold.caught Ágnes
'What a thorough cold Ágnes caught!'
b. *(Hogy) milyen alaposan fázott meg Ágnes!
COMP what thoroughly cold.caught PV Ágnes

In contrast, *exclusive* adverbs (with a negative connotation like *bonyolultan* 'in a complicated manner', *hibásan* 'with mistake', *hasztalan* 'in vain', etc.) obligatorily occur with inversion (i.e. in focus) in all their occurrences, including the case when they are used as a *milyen (Adv)* 'how (Adv)' exclamative.

To complicate matters, there are adverbs (e.g. *gyakran* 'often', *szépen* 'nicely'), which can freely occur with both patterns above (with or without inversion in the left periphery). These can freely occur with both word orders in exclamatives as well.

The next two sections will elaborate on the observed patterns. Section 4 will offer a structural analysis of the positions *wh*-expressions occupy in exclamatives and will explain the above split between the two classes of *wh*-phrases. Section 5 will compare exclamatives to questions.

4. The fine structure of the left periphery in exclamatives

A good starting point for identifying the various syntactic positions in exclamatives is the structure of constituent questions, which was first described in sufficient details by Horvath (1981) and É.Kiss (1987). This, as well as the syntactic map of the left periphery (section 4.1. and 4.2.) will pave the way to the discussion of exclamative sentences in section 4.3 and section 4.4.

4.1. *The fine structure of the left periphery in questions*

Recall from section 3 above that *wh*-questions in Hungarian obligatorily show inversion. What structural position do the *wh*-phrases occupy in these? The distribution of *wh-phrase*s in single constituent questions in Hungarian is the same as that of identificational focus constituents: the immediate preverbal position.

(21) a. Kit hívtál meg? question
who-ACC invited-2SG PV
'Who did you invite?'
b. ÁRONT hívtam meg. contrastive focus[8]
Áron-ACC invited-1SG PV
'It was Áron whom I invited.'

This preverbal position is a distinct position in the left periphery of the Hungarian clause. Since Brody (1990, 1995), it came to be known as the position of (contrastive) *focus*: FocP. When the specifier of FocP hosts a lexical focus or a *wh*-phrase (in complementary distribution), the head of FocP has to be filled by the verb. Verb raising to F^0 has the fine structure illustrated in (22). If the verb has a preverb, the preverb is stranded in a position lower than FocP, in AspP. This phenomenon was referred to as (verb–preverb) inversion in the previous sections.[9]

[8] Contrastive focus is called identificational focus in the terminology of É.Kiss (1998).
The default place of the stranded preverb is right after the verb, but this is a tendency rather than a rule: material can intervene between the verb and the preverb, albeit with varying degree for some speakers.

(22) $[_{FocP}$ focus/*wh* $[_{Foc'}\ V^0_i \ldots [_{AspP}$ PV $[_{Asp'}$ … $[_{VP}\ t_i$]]]]]

Verb–preverb inversion is therefore indicative of focusing/question formation. Whenever Spec,FocP is filled, verb–preverb inversion is obligatory.

4.2. *The higher left periphery: the properties of DistP and manyP*

So far we have seen that Spec,FocP hosts the immediately left-adjacent element to the verb. FocP is a rather low projection in Hungarian. It is dominated by a set of other left peripheral projections, most importantly the functional projections hosting distributive quantifiers (DistPs), topics (TopPs) and the complementizer projection (CP). These projections are ordered in the following way:

(23) $[_{CP}$ $[_{TopP*}$ $[_{DistP*}$ $[_{FocP}$ focus/*wh* V^0 $[_{AspP}$ PV ...]]]]]

Following the complementizer and topics, DistP is the projection of universal quantifiers in the left periphery. This projection was termed QP in É.Kiss (1987), and later came to be known as a distributive projection (DistP) due to Szabolcsi (1997), who argued that this position is unique in only hosting distributive constituents. The reason why we need a separate projection from FocP for universal quantifiers in Hungarian has to do with the fact that these quantifiers (i) cannot occur in Spec,FocP themselves, i.e. do not trigger inversion and (ii) are not in complementary distribution with a focused expression. These properties are illustrated in (24) and (25). (24) shows that a universal quantifier does not trigger inversion and is not compatible with it:

(24) a. Mindenkit meg-hívott János az ünnepségre.
everyone-ACC PV-invited János the celebration-ON
'János invited everyone to the celebration.'
b.*Mindenkit hívott meg János az ünnepségre.
everyone-ACC invited PV János the celebration-ON

(25a,b) illustrate that universal quantifiers allow for a focus to intervene between them and the verb, as the only constituent that is allowed to appear there:

(25) a. Mindenkit JÁNOS hívott meg az ünnepségre.
everyone-ACC János invited PV the celebration-ON
'It was János who invited everyone to the celebration.'

b.*Mindenkit tegnap meghívott az ünnepségre János.
Everyone-ACC yesterday PV-invited the celebration-ON János
'Yesterday János invited everyone to the celebration.'

This is in accordance to the structure in (23) which registers the fact that DistP dominates FocP in Hungarian.

Recent work (Kálmán et al. 2001) has argued that the DistP projection should rather be characterized as a DistP *field* comprising several slightly distinct projections. The split of the DistP projection is most notably required by the empirical properties of emphatic *many/much*-phrases, which also occupy a DistP position when in the left periphery (' stands for emphasis)[10]:

(26) 'Sok lányt meg-hívott János az ünnepségre.
many girl-ACC PV-invited János the celebration-ON
'János invited many girls to the celebration.'

Emphatic *many/much*-phrases express the speaker's judgement about a high amount or numeric degree. (26), for example, indicates that according to the speaker's judgement, there were many invited girls (above average, above expectation or contrasting with just a few girls).

Many/much-phrases are different from universal quantifiers in that they do not allow for a focus to follow them (compare (25a) to (27a)) and they themselves can occur in Spec,FocP as focused constituents (compare (24b) with (27b)):

(27) a.*'Sok lányt JÁNOS hívott meg az ünnepségre.
many girl-ACC János invited PV the celebration-ON
'It was János who invited everyone to the celebration.'
b. SOK LÁNYT hívott meg János az ünnepségre.
many girl-ACC invited PV János the celebration-ON
'It was many girls who János invited to the celebration.'

A further difference between universal quantifiers and *many/much*-phrases concerns co-occurrence restrictions with other types of quantifiers. DistP is recursive and can host more than one universal quantifier, in any order:

[10] *Many/much*-phrases without emphasis have a wider distribution. They can occur as topics or postverbal constituents as well.

(28) Mindenkit mindenhova meg-hívott János.
everyone-ACC everywhere PV-invited János
'János invited everyone to every place.'

Emphatic *many/much*-expressions, on the other hand, are unique and when they occur with universal quantifiers, they have to follow these:

(29) a.*'Sok lányt 'sok helyre meg-hívott János.
many girl-ACC many place-ON PV-invited János
'János invited many girls to many places.'
b. Mindenhova 'sok lányt meg-hívott János.
everywhere many girl-ACC PV-invited János
'János invited many girls to every place.'

To accommodate the observed co-occurrence restrictions we are led to a more detailed map of the left peripheral quantificational positions than the one in (23). The quantificational layer needs to be split into an iterable DistP for universal quantifiers, and a unique *many*P for *many/much*-expressions. The two positions emphatic *many/much*-phrases can occupy in the left periphery are indicated in (30a-b):

(30) a. ([$_{\text{DistP*}}$ ∀) [$_{\textit{many}\text{P}}$ *many/much*-XP [$_{\text{AspP}}$ PV-V [...]]]
b. ([$_{\text{DistP*}}$ ∀) [$_{\text{FocP}}$ *many/much*-XP V^0 [$_{\text{AspP}}$ PV [...]]]

One option for emphatic *many/much*-phrases is to occur in a quantificational position, adjacent to a non-inverted PV-V verb (30a), which is found right below *many*P.[11] The other option for *many/much*-phrases is to occupy the focus position and thereby force verb–preverb inversion, as in (30b).

The structures in (30) will form the basis of my analysis of exclamatives in the next section.

4.3. *The fine structure of the left periphery in exclamatives*

With the structure of the quantificational and focus layer firmly in place, I now turn to the structural properties of exclamatives. Recall from section 3 that *wh*-expressions in exclamatives split into two types according to what kind of left periphery position they occur in. Some obligatorily trigger

[11] This adjacency requirement has not yet received attention in the literature on Hungarian. Putting it down to the selectional restriction of *many*P suffices for the purposes of this paper but it is nothing more than a mere descriptive coding of the observed facts. The real explanation behind this adjacency presumably lies somewhere else.

inversion, others only optionally. In other words, all of them can occur with inversion and a subset of them can also occur without.

Identifying the position of *wh*-phrases with inversion is really easy, given that inversion is an earmark of focusing in Hungarian. When triggering inversion, exclamative *wh*-expressions are in FocP:

(31) ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{FocP}$ E-*wh* V^0 [$_{AspP}$ PV ...]]]]]

The position of those without inversion is not difficult to establish, either, once we recognize the parallels between *many/much*-phrases and flexible *wh*-phrases. These items have the exact same distribution in the left periphery. There are two key properties that characterize them uniquely and single out the same position for them:

(i) both of them are adjacent to the PV-V sequence (without inversion);
(ii) both disallow a focus to intervene between them and the verb.

These properties are illustrated again here, with examples repeated from above. The first example in both pairs is an exclamative, the second is an indicative clause with a *many/much*-expression. Property (i) is illustrated in (32)-(33) and property (ii) in (34)-(35).

(32) (Hogy) hány filmet meg-néztél! (ex. 14a)
COMP how.many film-ACC PV-watched-2SG
'How many films you watched!'

(33) 'Sok lányt meg-hívott János az ünnepségre. (ex. 26)
many girl-ACC PV-invited János the celebration-ON
'János invited many girls to the celebration.'

(34)*??(Hogy) hány filmet JÁNOS nézett meg! (ex. 16)
COMP how.many film-ACC János watched-3SG PV
intended: 'How many films János watched (and not someone else)!'

(35) *'Sok lányt JÁNOS hívott meg az ünnepségre. (ex. 27a)
many girl-ACC János invited PV the celebration-ON
intended: 'It was János who invited everyone to the celebration.'

These properties single out one structural position in Hungarian: *many*P. Other inhabitants of the DistP domain, namely universal quantifiers, do not have property (ii): they allow for a focus following them (see (25a) above).

Neither do higher left peripheral constituents have this property. This is also the case with topics. They do not only differ from exclamatives in property (ii), but also in property (i): they need not be adjacent to a verb and can be followed by focus as well as other quantifiers.

This identifies the position of flexible exclamative *wh*-phrases without inversion as that of emphatic *many*P, which, as argued before is higher than the position of focus.

(36) ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{manyP}$ flexible E-*wh* [$_{AspP}$ PV-V [...]]]]]]

Next to this, flexible items can also occupy the focus position, just like any inflexible *wh*-phrases:

(37) ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{FocP}$ flexible E-*wh* V^0 [$_{AspP}$ PV ...]]]]

Does this distribution of flexible *wh*-phrases come as a surprise? The following section will show that it does not.

Before showing why, let me illustrate some further distinguishing properties that characterize flexible E-*wh*-phrases in the two different syntactic positions, *many*P and FocP. These also provide further evidence for the structures proposed in (36) and (37), and show that syntactic placement has an effect on semantic and syntactic behaviour. I briefly exemplify each effect in the rest of this section, using the evidence from distributive readings and the licensing of postverbal superlatives respectively.

As noted above, *many*P is part of the quantificational field of Hungarian, which although comprises more than one functional projection, forces a uniformly distributive reading onto all constituents that appear there. This is in stark contrast with FocP, which can host constituents with both distributive and collective readings (Szabolcsi 1997). Due to this essential difference, the meaning of *many/much*-phrases differs in distributivity depending on their structural position. When they are in *many*P, i.e. in the quantificational field, they are obligatorily distributive. When in FocP, they are optionally distributive:

(38) a. 'Sok gyerek fel-emelte a zongorát.
many kid PV-lifted the piano-ACC
'Many kids lifted the piano (separately).'
b. SOK GYEREK emelte fel a zongorát.
many kid lifted PV the piano-ACC
'Many kids lifted the piano (separately/together) (not just few).'

The exact same phenomenon can be observed with flexible *wh*-phrases that are allowed to appear in both positions:

(39) a. (Hogy) hány gyerek fel-emelte a zongorát!
COMP how.many kid PV-lifted the piano-ACC
'How many kids lifted the piano (separately)!'
b. (Hogy) hány gyerek emelte fel a zongorát!
COMP how.many kid lifted PV the piano-ACC
'How many kids lifted the piano (separately/together)!'

The positional difference, however, does not only have interpretational consequences, but syntactic ones, too. When flexible E-wh-phrases occur in FocP, they assume the syntactic behaviour of focused constituents. As such, they license postverbal superlative expressions for example. Superlative licensing can only be done by focused elements, and not by quantificational ones, as was argued in É. Kiss and Farkas (2001). Thus, a lexical focus can, a *many/much*-phrase cannot license a superlative, as illustrated in (40):

(40) a. JÁNOS itta meg a legkevesebb bort.
John drank PV the least wine
'John drank the least wine (and not someone else).'
b.*'Sok fiú meg-itta a legkevesebb bort.
many boys PV-drank the least wine
'John drank the least wine (and not someone else).'

The same is true about exclamatives, when they appear in different positions. Exclamative wh-phrases can only license a superlative phrase from the focus position (i.e. with inversion), but not from *many*P (without inversion):

(41) a. (Hogy) hányszor voltál kész a leggyorsabban!
COMP how.often were-2SG ready(=PV) the quickest
'How often were you ready the quickest!'
b.*(Hogy) hányszor kész voltál a leggyorsabban!
COMP how.often ready(=PV) were-2SG the quickest

This section has argued that the placement of flexible exclamative *wh*-expressions follows that of emphatic *many/much*-phrases in Hungarian, and provided supporting evidence for the distinct syntactic positions these constituents can occupy. The next section will argue that the observed parallel between exclamatives and *many/much*-phases is not a coincidence,

but lies in the semantic properties of these items. This finding will allow us to define the flexible class of *wh*-phrases in terms of lexical properties.

4.4. *The positions of flexible E-wh phrases: evaluation, high degree and quantification*

Flexible E-*wh*-phrases pattern with emphatic *many/much*-phrases in their distribution: they can be either in Spec,*many*P or Spec,FocP. Given that emphatic *many* phrases express the speaker's own judgement concerning a high numeric degree or high amount, this does not come as a surprise for various reasons.

First of all, exclamative phrases in general express the speaker's own judgement about something unexpected, so they are evaluative, just like emphatic *many/much*-phrases. Second, the most unambiguously flexible E-*wh*-expressions are numerical or amount expressions that have a high amount reading in this position (see ex. 17 above, repeated here):

(17) a. (Hogy) hány filmet meg-néztél!
COMP how.many film-ACC PV-watched-2SG
'How many films you watched! / *How few filmes you watched!'

b. (Hogy) hány filmet néztél meg!
COMP how.many film-ACC watched-2SG PV
'How many films you watched! / How few filmes you watched!'

Emphatic *many/much*-phrases and flexible E-phrases thus share two crucial meaning components: evaluative value and the high degree reading.

Since flexible *wh*-phrases do not have the latter property when they occupy Spec,FocP, we have to conclude that the high amount reading is due to the functional syntactic position they occupy: *many*P is in fact a specialized projection for high amount evaluatives. It hosts constituents that are either lexically specified for a high amount reading (*many/much*-phrases) or are compatible with that (flexible E-*wh*-expressions).

The explanation provided here covers amount *wh*-expressions, but does not at the same time straightforwardly extend to plural phrases and quantified *wh*-phrases, which are also in the flexible class. What can we say about these?

Let me start with quantified items. The fact that quantified *wh*-phrases (*ki mindenki* 'who all', *mi minden* 'what all') are preferred to be used without inversion by many speakers, indicates that these escape from being focused. I put this down to the fact that they are modified by universal quantifiers, and therefore externally behave like those: they denote a high number of

individuals (recall (15)). This explains why they front to the quantificational field, rather than to FocP in exclamatives.

Variation among speakers can be described by saying that speakers split into two groups when it comes to assigning an externally quantificational value to quantified *wh*-phrases. For those speakers who treat these as externally quantified, these items behave as quantifiers in that they are excluded from FocP (compare (24b)). For those speakers who do not treat these as quantifier phrases, they behave like any other flexible *wh*-phrases, due to the fact that they are amount expressions.

The explanation for plural *wh*-phrases, among which *kik* 'who-PL' is the most frequently occurring one, runs along similar lines, I believe. The reason why these items, at least for some speakers, follow the distribution of flexible *wh*-phrases has to do with the fact that plurality can be conceived of as quantification in semantics (Link 1983). That plural *kik* 'who-PL' has something to do with quantification can be seen from the translation in (42) for example. *Kik* here refers to a high number of different types of individuals:

(42) (Hogy) kik el-jöttek!
COMP who-PL PV-came-3PL
'The (many) kind of people who came!'

We have thus seen that next to amount readings, quantification and plurality play a key role in the placement of exclamative items. Those items that are lexically specified or morphologically marked for amount readings or plurality can occupy a quantificational position without triggering inversion in exclamatives.

This finding gives us a handle on what kind of *wh*-phrases belong to the flexible and inflexible classes of *wh*-expressions, and allow us to define these in terms of semantic properties. As could be seen above, flexible *wh*-phrases can be defined *as* wh-*phrases that are quantificational or amount expressions*. Inflexible *wh*-phrases are those without these properties.

Summing up, in this section I have examined the distributional behaviour of left peripheral focus, quantificational and amount-phrases in Hungarian and presented empirical arguments to the effect that *wh*-phrases in exclamatives can occupy two hierarchically distinct positions: emphatic *many*P and FocP. *Many*P was identified as a possible position of flexible *wh-phrase*s, which in turn were defined as quantified/amount expressions. FocP was identified as the position for both flexible and inflexible *wh*-phrases, the latter being *wh*-phrases without a quantified/amount reading. The fact that Hungarian thus renders available two left peripheral positions for

exclamative items along the lines specified here is a general property of the language independent of the formation of exclamatives. Since this language "wears its LF on its sleeve", it has a wide range of specialized quantificational positions in the left periphery.

This however does not mean that the above identified positions are both available to *wh*-constructions in any context in Hungarian. *Wh*-phrases in questions are restricted to a singular position only. To show why this is not the case, the next section returns to the other explananda: the observed word order differences between questions and exclamatives.

5. Differences between questions and exclamatives

The previous sections have illustrated that Hungarian questions do not allow for the same amount of freedom when it comes to the placement of *wh*-expressions as exclamatives do. While all *wh*-expressions are found in Spec,FocP in questions, not all of them are found there in exclamatives. The placement possibilities for these two sentence types are repeated in (43)-(44):

(43) ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{FocP}$ Q-*wh* V^0 [$_{AspP}$ PV ...]]]]] question

(44) a. ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{manyP}$ E-*wh* [$_{AspP}$ PV-V]]]]] exclamative
b. ([$_{CP}$ [$_{TopP*}$ [$_{DistP*}$) [$_{FocP}$ E-*wh* V^0 [$_{AspP}$ PV ...]]]]] exclamative

The difference in placement originates from the underlying structure of interrogatives and exclamatives, more precisely the role focus plays in these.

The formation of constituent questions in Hungarian makes use of the FocP position, as a result of the fact that *wh*-constituents in questions are contrastive focus items semantically (É. Kiss 1987, Lipták 2001). This forces all *wh*-expressions to appear in FocP and receive contrastive focus interpretation, just like any lexical focus item.[12]

The situation is different in exclamatives: exclamative constituents are not a case of contrastive focus. While focus is a necessary ingredient of exclamatives, the focus here is not contrastive. Focus is inherent to all exclamatives due to the fact that exclamatives are emotive expressions and have a *scalar* implicature (Michealis and Lambrecht 1996). Exclamatives are

[12] A more precise account would have it that Hungarian *wh*-expressions, which are lexically variables, are bound by a word-level question operator morpheme (Q_{wh}) in questions. This operator provides them with question semantics, and carries the contrastive <+f> focus feature that drives overt movement of *wh*-phrases to FocP. For further details see Lipták (2001).

characterized by a scalar operator that establishes a scale and picks an unexpected or unusual high or low value from this scale.

Such a scalar operator is associated with two positions of the Hungarian clause. One is the FocP, and other is the evaluative high degree *many*P, which was identified in section 3 above. This follows from the fact that *wh*-constituents in FocP and *many*P do not differ in exclamative meaning, suggesting that both are associated with an exclamative scalar operator, which provides meaning to the *wh*-expression they host.

The two positions differ in other properties, as was established in the previous section. *Many*P is more selective for the constituents it can host: unlike FocP, it can only host amount/numeric expressions (*flexible wh*-phrases). Interpretation-wise exclamative phrases in *many*P can only have a high degree reading (ex. (17) above). Further differences between *many*P and FocP are properties that stem from the fact that they belong to two different layers of the left periphery: the quantificational layer and the focus layer. *Many*P, just like the quantificational projection DistP, forces a distributive reading on its specifier. FocP, not being part of the quantificational layer of the left periphery,[13] does not force a distributive reading on its specifier.

6. Summary and conclusions

In this article I presented a structural account of *wh*-exclamatives, which aimed at explaining the observed word order variation found within these clauses and in contrast to interrogatives. Such variation and contrast are not unknown from work on other languages (Munaro 2003, Zanuttini-Portner 2003).

In characterizing the Hungarian word order patterns, we found evidence for the following claims:

(i) *wh*-constituents in exclamatives have two distinct left peripheral positions open for them: focus (FocP) and a higher quantificational projection (*many*P);
(ii) both projections are associated with a scalar exclamative operator, but have different selectional properties: *many*P can only host amount/numerical expressions, while FocP has no selectional restrictions;
(iii) as a result, amount/numerical *wh*-expressions thus have a greater freedom of placement in exclamatives;

[13] Which is not to say focus is not a *quantificational,* A-bar element. The claim above only concerns its position: it is not one of the *positions* in the quantifier domain.

(iv) exclamatives and interrogatives differ in the *type of focus* they assign to *wh*-expressions: contrastive focus in interrogatives and evaluative scalar focus in exclamatives;
(v) the two types of focus have a distinct syntax (forcing or allowing inversion); only contrastive focus is unambiguously linked to FocP (inversion).

The properties of Hungarian *wh*-exclamatives that were revealed in this paper contribute to the growing knowledge about typological variation among exclamatives, focus and *wh*-constructions as well.

References

Bennis, Hans. 2000. "Exclamatives!". *Linguistics in Netherlands 2000.* Amsterdam/Philadelphia, John Benjamins.

Brody, Mihály. 1990. "Remarks on the order of elements in the Hungarian focus field." In I. Kenesei (ed.) *Approaches to Hungarian* 3. 95-122. Szeged, JATE Press.

Brody, Mihály. 1995. "Focus and Checking Theory." In I. Kenesei (ed.) *Approaches to Hungarian* 5. 31-41. Szeged, JATE Press.

Den Dikken, Marcel and Anikó Lipták. 1997. "Csoda egy nyelv – Nominal

Internal Predication in Hungarian". In J. Coerts and H. de Hoop (eds.) *Linguistics in the Netherlands 1997*. 177-214. Amsterdam/Philadelphia, John Benjamins.

Kálmán, László et alii. 2001. *Magyar leíró nyelvtan I.* [Hungarian descriptive grammar I]. Budapest, Tinta Könyvkiadó.

Kenesei, István, Anna Fenyvesi and Rovert Vago. 1998. *Hungarian*. Routledge Grammars.

Kiefer, Ferenc. 1967. *On Emphasis and Word Order in Hungarian*. Bloomington, Indiana University Press.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian.* Dordrecht, Kluwer.

É. Kiss, Katalin. 1998. "Identificational focus versus information focus". *Language* 74:245-273.

É. Kiss, Katalin and Donka Farkas. 2001. "On the comparative and absolute readings of superlatives". *Natural Language and Linguistic Theory* 18:417-445.

Link, Gödehard. 1983. "The logical analysis of plurals and mass terms". In R. Bäuerle, C. Schwarze and A. von Stechow (eds.) *Meaning, Use and Interpretation of Language.* 250-269. Berlin, Walter de Gruyter.

Lipták, Anikó. 2001. *The syntax of wh-phrases in Hungarian*. LOT disseration series.

Michealis, Laura and Knud Lambrecht. 1996. "The exclamative sentence type in English". In *Conceptual Structure, Discourse and Language*. 375-389. CSLI Stanford.

Munaro, Nicola. 2003. "On Some Differences between Exclamative and Interrogative Wh-phrases in Bellunese: Further Evidence for a Split-CP Hypothesis". In C. Tortora (ed.) *The syntax of Italian dialects*. 137-151. New York/Oxford, Oxford University Press.

Sadock, Jerrold M. and Arnold M. Zwicky. 1985. "Speech act distinctions in syntax". In T. Shopen (ed.) *Language typology and syntactic description*. 155-196. Cambridge, Cambridge University Press.

Szabolcsi, Anna. 1997. "Strategies for scope taking". In A. Szabolcsi (ed.) *Ways of scope taking.* 109-145. Dordrecht, Kluwer.

Villalba, Xavier. 2003. "An exceptional exclamative sentence type in Romance". *Lingua* 113:713-745.

Zanuttini, Raffaella and Paul Portner. 2003. "Exclamative Clauses at the Syntax-Semantics Interface". *Language* 79:36-81.

COMP: a multi-talented category. Evidence from Romance

Sandra Paoli
University of Cambridge, UK

1. Introduction

Traditionally the CP has been thought of as the projection where complementation obtains. Its head, COMP, has been analysed as the position in which complementisers are generated (Pesetsky 1982, Chomsky 1986a). More recently (cf. Chomsky 1995, Frajzyngier 1995, Rizzi 1997, Benincà 2001, Benincà and Poletto 2004), studies within the generative tradition have argued that COMP does more than simply mark clause boundaries: it encodes information on the semantic status of the whole proposition, such as illocutionary force/clause-typing, as well as acting as an interface between discourse and propositional content.

Assuming Rizzi's (1997) split-CP system, further refined by Benincà (2001) and Benincà and Poletto (2004), the data put forward in this paper presents evidence from a number of diverse Romance varieties (two Northern Italian dialects (Tur(inese) and Mar(ebbano)), various early Romance (eR) texts, Tuscan, French and Castilian, and French acquisitional (FrAcq) data) for the existence of [mood], [topic] and [φ] features expressed overtly through various heads in the left periphery[1]. The conclusion is that, clearly, C is a 'multi-talented' category, which reinforces the idea of a split-CP structure, a field of syntactically and semantically distinct projections.

[1] The sources of the data presented in this article are as follows:
Tur and Mar are data that I have collected myself;
eR: data obtained through searches on the OVI site (http://www.lib.uchicago.edu/efts/ARTFL/projects/OVI/), an article by Wanner (1995);
FrAcq: data obtained from Rohers and Labelle (2003).

The article is organised as follows. Sections 1 and 2 investigate a superficially identical construction, namely the apparent repetition of the finite complementiser *che/que*, in Tur, eR texts and FrAcq, and by highlighting the different functions performed by the second *che/que*, it is argued that while in eR texts and FrAcq the second complementiser lexicalises Top°, in Tur it is the overt expression of mood features in Fin°. Section 3 turns to the alternation of the relative pronouns *che* and *co* in Mar, and claims that [φ] features, too, can be expressed at the C level.

2. The 'Double *che/que* Construction' in Turinese

A conservative variety of Piedmontese spoken in the city of Turin, Italy, to which we refer simply as Tur, allows, in some embedded contexts, the apparent repetition of the finite complementiser *che*. The verb in the main clause belongs to the so-called 'verb-of-belief' type and it selects a '*that*-clause': the first complementiser, *che1*, appears immediately after the main verb, and the second one, *che2*, after the embedded subject[2]:

(1) a. Gioanin a spera **che** Ghitin **ch**' as në vada tòst
John scl hope.pr.1s that Margaret that scl+rf pt go.S.3s soon
'John hopes that Margaret leaves as soon as possible'
b. Majo a chërde **che** Luch **ch**' a sia dësmentiass-ne
Mario scl believe.pr.3s that Luke that scl be.S.3s forget.pple.rf-pt
'Mario believes that Luke has forgotten about it'
c. Majo a pensa **che** Franchin **ch**' as n' ancorza
Mario scl think.pr.3s that Frank that scl+rf pt realise.S.3s
'Mario thinks that Frank will realise it'

The use of *che2* is not compulsory, in fact all the above examples are grammatical without it; an account of this optionality is suggested later in this section. Its presence is, nevertheless, totally excluded from some contexts, which allows us to draw firm conclusions on its role. The licensing condition on *che2* is that the embedded verb be in the subjunctive mood: present and future indicative and conditional do not trigger it, as shown in (2), in which the overt realisation of *che2* makes the sentences ungrammatical.

[2] Throughout the examples presented in the paper the following abbreviations are used: scl = subject clitic; pr = present indicative; 1,2,3 = first, second, third person; s/p = singular/plural; rf = reflexive clitic; pt = partitive clitic; neg = negation; S = present subjunctive; cond = conditional; inf = infinitive; fut = future.

(2) a. A dis **che** Marìa e Gioann **(*ch')** a mangio nen` d rane
scl say.pr.3s that Mary and John scl eat.**pr**.3p neg of frogs
'S/He says that Mary and John do not eat frogs'

b. Giòrs a spera **che** Majo **(*ch')** as n' andarà tòst
George scl hope.pr.3s that Mario scl+rf pt go.**fut**.3s soon
'George hopes that Mario goes away soon'

c. Majo a pensa **che** Franchin **(*ch')** as n' ancorzerìa
Mario scl think.pr.3s that Frank scl+rfpt realise.**cond**.3s
'Mario thinks that Frank would realise it'

Given that the subjunctive mood selected by 'verb of belief' type of verbs seems to be the triggering factors for *che2*, it could be argued that *che2* is some sort of *irrealis* marker. This hypothesis is not empirically sustainable. Examples (2b) and (2c) on one hand, and (1a) and (1c) on the other are minimal pairs that allow us to make a revealing comparison: although in all of these sentences the main verb is of the same type, namely a 'verb of belief', only in those sentences with the embedded verb in the subjunctive *che2* is licensed. The relevance of the role played by the selecting verb in the licensing of *che2* is further questioned by two other pieces of evidence. The first one is that there are instances of *che2* not licensed by a verb but by a conjunction (cf. (3a)), by what we could analyse as a clause-type operator for imperative clauses (cf. (3b)), and by a relative operator (cf. (3c)):

(3) a. I veno volonté, basta mach **che** Gioann **ch'** a staga
scl come.pr.1s willingly as long as that John that scl stay.S.3s
nen solo
neg alone
'I will come willingly as long as John is not left on his own'

b. **Che** ij cit **ch'** a vado a pluché sùbit!
that the children that scl go.S.3p to sleep.inf immediately
'The children should go to sleep immediately!'

c. Giòrs a veul parlé con un dotor **che**, ant la meisin-a
George scl want.pr.3s speak.inf with a doctor who in-the medicine
autërnativa, **ch'** a-j chërda
alternative that scl+loc believe.S.3s
'George wants to speak to a doctor who believes in alternative medicine'

The second one is that there are cases of *che2* being triggered by verbs that do not belong to the 'verb of belief' type:

(4) Marìa a regreta **che** Giòrs **ch'**a sia dësmentiass-ne
Mary scl regret.pr.3s that George that scl be.S.3s forget.pple-pt
'Mary regrets the fact that George forgot about it'

The verb in the main clause is a factive verb which, by implying the truth of its complement clause, is semantically opposed to the 'verbs of belief', that, by definition, do not assert the truth of their complement clause. The conclusion that can be drawn is that the triggering factors and the *raison d'être* of *che2* are not lying with the syntactic and semantic properties of the selecting verb, nor in some sort of *irrealis* feature in the embedded clause[3]. With these premises it seems plausible to discard the link to a specific type of selecting verb, to acknowledge the unequivocal dependence of realisation of *che2* on the presence of the subjunctive, and to claim that *che2* is a subjunctive marker. In order to support this claim we concentrate on the morpho-syntactic properties of the subjunctive.

It has been argued in the literature (cf. Giorgi and Pianesi, 1997; von Stechow, 1995, among others) that the subjunctive is a deficient tense. The term 'deficient' could be interpreted both semantically and morphologically. Semantically, because on its own the subjunctive does not give rise to any real temporal interpretation; morphologically, because of the lack of morphological differentiation between some forms of the present indicative and the subjunctive. This deficiency is also witnessed in Tur: the differentiation between the present tense in indicative and subjunctive is minimal, as shown in the following table (a similar, but not identical, pattern is also found in verbs belonging to the second and third conjugation):

Table 1: Tur: first conjugation *parlé* 'to speak'

	Pres. Indicative	**Pres. Subjunctive**
1s	**mi i parlo**	**che mi i parla**
2s	ti it parle	che ti it parle
3s	chiel a parla	che chiel a parla
1p	**noi i parloma**	**che noi i parlo**
2p	voi i parle	che voi i parle
3p	lor a parlo	che lor a parlo

Recall that it was mentioned that the use of *che2* is a matter of preference in those cases in which it is licensed. There is a striking, almost 100%, correspondence between the poor morphological distinction between

[3] The terms *realis* and *irrealis* have been, and still are, subject to an ongoing discussion concerning their linguistic validity. Here this debate will not be addressed, and the reader is referred to, among others, Elliott (2000), Palmer (2001) and references cited there.

indicative and subjunctive present tenses and the occurrence of *che2*, which is more 'readily' produced precisely in those cases in which the verb is morphologically identical in the subjunctive and the indicative. For example, just like in the paradigm illustrated in table 1, *che2* is 'preferred' with the second and third person singular and plural. Conversely, it is felt to be less 'necessary' with the first person singular and plural.

All these facts strongly suggest that *che2* serves as a morphological subjunctive marker, expressing overtly those morphological features that remain under-specified on the verb. A similar function is performed by the Romanian particle *să*, which appears alongside the indicative to form the subjunctive.

The unequivocal link between *che2* and the subjunctive mood as well as the facts discussed in this section have clearly suggested that *che2* is a modal particle. Cinque (1999) posits several heads with modal content at the edge of his extended IP; Rizzi (1997) assumes the lower of the heads in the CP, Fin°, to encode modality. Whether mood is encoded in a head within the IP or the CP spaces is not a straightforward issue, further complicated by the high degree of decomposition of the two functional domains and by the idea (cf. Rizzi 1997) that some inflectional features are reduplicated as an impoverished version at the CP level. There is nevertheless some evidence that suggests that *che2* belongs to the left periphery of the clause.

In all the examples *che2* forms a cluster with what has been glossed as 'scl', 'subject clitics'. Poletto (2000) proposes a categorisation of these elements based on an investigation of more than a hundred Northern Italian dialects: scls fall into four different types, two placed in the IP and two in the CP. Each type displays particular properties, and in turn, specific properties characterise each type. An investigation of Tur scls based on Poletto (2000), reveals that the scls with which *che2* forms a cluster belong to the Deictic type, one of the two CP types, placing *che2*, in turn, in the left periphery. Given its mood content, it is reasonable to conclude that *che2* lexicalises Fin°. The fact that mood is expressed at the C level is not a new idea, and it goes back to at least den Besten (1983) and Stowell (1982). In Enç (1987) the link between T and C is expressed in terms of semantic anchoring: the specifier of tense, or in other words, the source of its value, is located in COMP. This relation is interpreted syntactically through a binding relation between T and C which holds across categories.

A final observation. The relative order between *che2* and the scl with which it forms a cluster requires assuming that *che2* has moved into that position in a system that does not allow right adjunction. We propose that *che2* originates in the embedded v° that dominates the VP projection, and then raises via head-to-head movement to Fin° to check its [mood] features.

The whole VP, and the lexical verb included, then raises into the Specifier of T, deleting the uninterpretable [D] feature on it: this movement ensures that no XP intervenes between *che2*, the scl and the lexical verb. This choice is not as surprising as at first it may appear. vP is associated with the functional as well as with the semantic content of the verb phrase it dominates, and it seems plausible to assume that v° expresses information relating to non-lexical properties of the verb. Recall that *che2* does just this: it overtly expresses those mood features that are morphologically deficient on the subjunctive verb. From a purely formal point of view this analysis satisfies the requirement that the functional features on v° only be licensed by a functional particle or by an expletive. In those cases in which *che2* is licensed but can be omitted, we assume that v° is filled by a null operator that moves to Fin° as a bundle of silent features (cf. Chomsky 1995)[4].

Concluding, this section has shown that [mood] features can be expressed at the CP level. No clear semantic link between *che2* and a specific type of modality has been established; we therefore suggest that these [mood] features are solely related to the presence of the subjunctive, more precisely to the indicative/subjunctive contrast. Interestingly, this bears in turn on the question of the degree of 'reduplication' of inflectional information. It seems that the inflectional features expressed at the C level are, in this case, a reduced version of those encoded at the I level, only offering a binary [±] choice for a specific mood, rather than ranging over the whole array of mood specifications found in the IP.

We now turn to a similar construction in early Romance.

3. The 'Double *che/que* Construction' in early Romance

A similar construction to the one we have just discussed is found in 13th century Romance texts: two finite complementisers (*che/que*, henceforth *QUE1* and *QUE2*) are simultaneously realised in embedded clauses. This happens when some syntactic material appears between *QUE1* and the verb in the embedded clause. Keniston (1937:675), in his collection of Castilian texts, describes this construction as *'... (a) special use of annunciative que ... repeated, after another word or phrase… it is a common practice in the sixteenth century to repeat annunciative que when some element of the sentence intervenes between que and the verb of the clause. This usage is especially common when an adverbial clause precedes the verb; but it is*

[4] There are two possible alternatives to the proposed analysis. It could be assumed that either *che2* and the scl form a phonological cluster, and hence, no movement is involved, or that the scl raises into Fin° to adjoin to *che2* from a lower C position in which it is merged. Neither seem to affect the conclusions reached. We are grateful to an anonymous reviewer for these suggestions.

also found after relative clauses, or even after other elements, such as the subject or object of the verb'.

In the examples collected, there is no evidence that the two *QUE* were allowed to appear in a sequence: this is taken as a reflection of the fact that *QUE1* and *QUE2* could not be simultaneously realised unless they were separated by phonetically realised syntactic material. There is a variety of elements that can appear between the two *QUE*:

a causal clause:

(5) a. ... diz **que** porque la poblara alli brennio pora premia & danno de los Romanos **que** pusiera aquell nombre verona
(Gen Est;f171r1.14)
'S/He says that because Brennio had populated it to put pressure on and harm the Romans, that he had given it the name of Verona'

b. ... E tanto savio bello e largo portamento ver' de ciascuno facea, che tanti d'onne parti cavalieri trassero a llui, **che** per lo gran senno e valore suo e larghezza e per bona cavallaria che lui seguia, **che** XXVIII reami se soctomise. (Con, 21:150-151)
'that for his great sensibility, value and greatness and for the quality of his cavalry that followed him, that he conquered twenty nine kingdoms'

a hypothetical clause:

c. ... A queste novelle, si pensò ir re Pelleus **che**, se elli potesse tanto fare che Giason suo nepote volesse andare in quella isola per lo tosone conquistare, **che** mai non tornerebbe, e in tal maniera si diliverebbe di lui (Distr Tr; XDIV 1, pg 152, 21-25)
'... king Pelleus thought that, if he could do so that his nephew Giason wanted to go to that island to take that ..., that he would never come back, and so he could get rid of him'

a temporal adverb/clause:

d. ... fueron los desuiando de leuar los consigo en la batalla diziendo les **que** da quella uez **que** escusar los podien (Gen Est;179v2.19)
'And they managed to avoid to take them with into battle by saying that on that occasion that they could be excused'

e. ... e egli allora fue vie piue innamorato de lei che non iera dapprima, e amava sò fforte mente che a llui sì era tutta via viso **che** quando persona neuna la sguardasse, **che** inmantenente iglile togliesse (Tr Ricc; Cap 75,pg 149, 25-28)
'...and he loved so intensely that to him it was shown that, when nobody was watching, that immediately he would take them off her'

an object:

f. Et sobresto la demanda del Conde don Sancho era esta. **Que**. vi. castiellos **que** Almançor ganara de los xcristianos en otro tiempo estonces Yssem que gelos diesse (EstEspf106v2.37)
'And about this Count don Sancho's request was this: that Yssem the six castles that Almançor had won from the Christians long before that he gave them to him'

g. Sire, je te adjure par le vray Dieu **que** ta fille Tarsienne, **que** tu ne la donnes a mariage a autre que a moy (Apoll, f48b)
'Lord, I beg you in the name of the true God that your daughter Tarsienne you do not give her in marriage to anybody but me'

a subject:

h. ...& ordenaron assi **que** los germanos **que** fincassen en sus tierras..(Gen Est; f171v2)
'And they ordered thus that the Germans that they stayed in their land'

From the above examples it is immediately apparent that the mood restriction that regulated the appearance of *che2* in Tur does not apply to *QUE2*: in the embedded clause we find indicative, subjunctive and conditional verbs. The presence of the subjunctive is not very significant in itself, being due to the high proportion of hypothetical clauses that occupy the space between the two *QUE*. No link with modality can therefore be claimed. It seems that the only requirement on the realisation of *QUE2* is that there be phonetically realised syntactic material between the two *QUE*. Wanner (1995:421) interprets this construction in early Romance as a strategy to give prominence to a thematised phrase located between the two complementisers: '*La syntaxe médiévale des langues romanes permet une mise en relief dans la phrase subordonnée. Le procédé consiste en une antéposition de l'élément relevé à l'intérieur de la subordonnée (son thème) suivi d'une deuxième conjonction subordonnante apparemment*

superflue[5]*…'*. What Wanner describes as a 'preposition of the element inside the embedded clause' for discourse reasons is strongly reminiscent of the operation of left-dislocation. Combining this intuition with Rizzi's (1997) characterisation of left-dislocated elements, we claim that *QUE2* is the overt realisation of Top°, the head of the projection for left-dislocated phrases, which is lexicalised only when its Specifier position is filled: the presence of a phrase specified for [+Top] features triggers a [Spec, Head] agreement matching relation making *QUE2* visible. *QUE1* will simply be assumed to occupy Force°.

This analysis is readily applicable to those cases in which the thematised element is a direct object: in (5f) and (5g) the direct objects *vi. castiellos* 'six castles' and *ta fille Tarsienne* 'your daughter Tarsienne' are resumed, respectively, by the clitics *-los* 'them' and *la* 'her', making it clear that the direct objects are left-dislocated, and therefore occupy [Spec, Top]. As for the other elements, a small digression is needed.

Hypothetical clauses are analysed by von Fintel (1994:78ff) either as topical, namely expressing old information (more commonly), or as focus, expressing new information, the topic/focus status depending on their sentence-initial or sentence-final position respectively. This is exemplified in (6) and (7), in which the questions make it clear which portion of the answer is given information:

(6) What will you do if I give you the money?
a1. If you give me the money, I'll buy this house
a2.# I'll buy this house if you give me the money

(7) Under what conditions will you buy this house?
a1.# If you give me the money, I'll buy this house
a2. I'll buy this house if you give me the money

from von Fintel (1994:81)

The *if*-clause is preferred in sentence-initial position when it expresses given information, either already introduced in the discourse or known to both speaker and interlocutor, and in sentence-final position when it contributes new information[6]. The expression of old information is one of the defining characteristics of topicalised phrases, so that it is reasonable to claim that

[5] The medieval syntax of the Romance languages allows a 'giving-prominence' strategy in the embedded clause. The process consists in the pre-posing of the element inside the embedded clause (its theme) connected with an apparently superfluous second subordinating conjunction'.

[6] Cf. Munaro (2005) for the application of this classification to the introduction of two functional projections in the upper part of the left periphery: ConcessiveP and HypotheticalP.

when *if*-clauses occupy a sentence-initial position they are topicalised and fill [Spec, Top]. The idea that 'if-then' clauses form a class distinct from other types of conditional clauses is also supported by Haegeman (2003). Furthermore, instances of enclisis in Medieval Italian varieties, which respected loosely the verb-second restriction, as analysed by Benincà (1995), suggest that hypothetical, causal and in general all extra-sentential complements are topicalised when appearing sentence-initially:

(8) …e [quando il vide], raffigurol*lo*
'and when him-saw, recognised-him' (Schiaff 77, 17)
from Benincà (1995:336)

Accepting this evidence as conclusive and extending it to temporal and other types of clauses, we claim that when occupying a sentence-initial position (which includes all the cases in eR of clauses between *QUE1* and *QUE2*) an extra-sentential clausal complement is topicalised and occupies [Spec, Top].

Finally, let us turn to the remaining case, subjects. A possible indicator of the topicalised status of a subject is the presence, in those languages that have them, of a subject clitic. French is one of these languages: clitics co-occur with a pronominal or lexical left-dislocated subject.

(9) Pierre, il est parti
Pierre scl be.pr.3s leave.pple
'Pierre, he's left'

Children acquiring French as their first language occasionally produce sentences similar to those seen in the eR texts, in which the finite complementiser is repeated:

(10) a. Quand **que** les Indiens **qu'** i veulent l'attaquer
'When the Indians that they want to attack it'
b. Elle croyait **que** les loups **qu'**i les avaient mangés
'She thought that the wolfs that they had eaten them'
c. Quand **que** les Indiens **qu'**i veulent l'attaquer
'When that the Indians that they want to attack it'
from Rohers and Labelle (2003)

Interestingly, the element between the two complementisers is a subject and it co-occurs with a subject clitic, suggesting that it is not in its canonical position but is left-dislocated. Extending this conclusion to the eR data, we conclude that subjects appearing between the two *QUE* are also topicalised.

Concluding, we have shown that all the different elements that occupy the position between the two complementisers occupy a discourse-prominent position. We have interpreted this at the syntactic level by claiming that these elements occupy [Spec, Top]. *QUE2* can therefore be analysed as the overt realisation of Top°, made visible only when an element carrying a [+Top] feature enters a [Spec, Head] agreement relation with it. This was both a necessary and a sufficient reason for the triggering of *QUE2*.

We conclude the investigation of the COMP category turning to relative pronouns in Mar.

4. Relative pronoun alternation in Marebbano

Drawing on a Northern Italian variety we claim in this final section that [φ] features, too, can be expressed through an overt head at the CP level.
Mar, a variety of Ladin (a member of the Rhaeto-Romance family), has two relative pronouns, *che* and *co*, used for relativised objects and subjects respectively:

(11) a. La ëra **co** puzenëia les stighes è püra
the lady who clean.pr.3s the stairs be.pr.3s ill
'The lady who cleans the staircase is ill'
b.*La ëra **che** puzenëia les stighes è püra
c. La ëra **che** te ás encunté ennier è mia mëda
the lady who scl have.pr.2s meet.pple yesterday be.pr.3s my aunt
'The lady you met yesterday is my aunt'
d.*La ëra **co** te ás encunté ennier è mia mëda

Adopting the 'null operator' analysis for relative clauses (cf. Chomsky, 1980; 1981) according to which a null operator (Op henceforth) is fronted, and assuming different merging positions for subjects and objects, it could be argued that in the formation of relative clauses in Mar, extraction of an Op generated in [Spec, vP] triggers *co* while extraction of an Op generated post-verbally triggers *che*. Relativisation of subjects of unaccusative, reflexive, ergative and passive verbs, clearly shows that this is not the case:

(12) a. Sü amisc **co** gnará endoman dormiará atlò
his friends who arrive.fut.3p tomorrow sleep.fut.3p here
'His friends who are arriving tomorrow will sleep here'
b. La ëra **co** se ciara te spidl se mosöra en bel guant
the lady who rf look.pr.3s in mirror rf try.pr.3s a nice dress
'The lady who is looking at herself in the mirror is trying on a nice dress'

c. Le boché **co** s'è rot ea de mia
the vase that rf be.pr.3s break.pple be.imp.3s of my
besaóna
greatgrandma
'The vase that broke was my great grandma's'

d. I ciöfs **co** ea gnüs dês jö tla
the flowers that be.imp.3s come.pple give.pple down in the
rezeziun ea proprio por te
reception be.imp.3p really for you
'The flowers that had been delivered to the reception were really for you'

The examples in (12) clearly rule out that the initial position occupied by the Op has a part in the choice of the relative pronoun. Rather, they underline that it is the functional role played by the DP with which the Op is co-indexed that is the determining factor, i.e. whether it is the subject or the object of the clause. The crucial difference between the two is that the Op for the subject, but not the object, passes through [Spec, TP]. We could assume that all elements undergoing movement bear a 'checking history', which is then interpreted accordingly by the computational system. This makes sure that a subject is interpreted differently from an object by virtue of the positions it has moved through. This assumption would predict that *wh*-phrases, too, are subject to the same constraints. Following Kayne (1976, 1978) in assuming that the relative pronoun and the complementiser are the same lexical element, we would expect instances of *wh*-movement involving an overt complementiser to display the same alternation. This is, indeed, the case. In embedded questions, which require the *wh*-phrase to co-occur with the complementiser, the latter is realised as *co* if the moved *wh*-phrase is the subject of the embedded clause, as *che* if it is not[7].

(13) a. Dìjemo chê **co** s'à tut le cader
tell.imper.me-mo who that rf have.pr.3s take.pple the painting
'Tell me who has taken the painting!'

b. Dìjemo ci **che** Maria mangia
tell.imper.me-mo what that Maria eat.pr.3s
'Tell me what Mary eats!'

We assume that the relative pronoun fills the head of Force in the split-CP system; we further assume that both *co* and *che* are merged as the same

[7] Incidentally, the data in (14) also supports Kayne's (1976, 1978) unified analysis of relative pronoun and complementiser .

element [k] whose final form is determined through [Spec, Head] agreement with the Op raising into [Spec, Force][8].

The phenomenon discussed here closely resembles the *que/qui* alternation witnessed in French: just as Mar *che*, *que* co-occurs with a relativised object, and *qui* is triggered with a relativised subject, as Mar *co*. Nevertheless, the two relative pronouns in Mar do not correspond *tout court* with their French counterparts, as their availability is subject to person constraints. More specifically, *co* is only available to third person subjects, both singular and plural: the relative pronoun for first and second relativised subjects can only be *che*. This is particularly evident morphologically with the second person singular: in all the other cases the presence of a scl and the elision of the 'e' in *che* makes it less transparent that it is really *che* rather than a reduced version of *co*. Crucially, though, *co* makes the sentence ungrammatical[9]:

(14) a. Iu **ch'** i mangi dagnora plü de düc á ciamó fan
I who scleat.pr.1s always more of all have.pr.1s still hunger
'I, who am always eating more than everybody else, am still hungry'
a'.*Iu co i mangi dagnora plü de düć á ćiamó fan

b. Tö **che** te lies le foliet vigne dé sas dagnora döt
you who sclread.2s the newspaper every day know.2s always all
'You, who read the newspaper every day, always know everything'
b'.*Tö co te lies le foliet vigne dé sas dagnora döt

[8] An alternative analysis could be formulated on the lines of Rizzi's (2004) 'Criterial Freezing' approach, a constraint that prevents a phrase from moving further once it has met a criterion. We would need to assume a *pro* raising into [Spec, TP], and the relative Op carrying [+wh] features would target [Spec, Force]. By virtue of being co-indexed, the two would share their feature specification, so that the [φ] features could be 'visible' by the head of Force and through a [Spec, Head] agreement, they would trigger the appearance of *co*. This analysis and its consequences have not been yet fully investigated, so they are here left as a pure speculation.

[9] Support for the idea that in the examples in (14) we are not really witnessing a *co* being reduced to [k] can be found in the following data. While the 'e' of *che* can be elided, this cannot be done with the 'o' of *co*, not even when a cluster of vowels occurs:

(i) a. Al è sté propi la möta con les trëces **co à**
scl+L be.pr.3s be.pple just the girl with the plates who have.pr.3s
orü s' an jì
want.pple rf scl+pt go.inf
'It was exactly the girl with plates who wanted to leave'
b. *Ël c' è n bel ël è tres plën d' ëres
he who be.pr.3s a handsom man be.pr.3s very full of women
'He who is a good looking man is very popular with the ladies'

c. Nos **ch'** i laurun con i mituns capiun chisc
we who scl work.pr.1p with the children understand.pr.1p these
problems
problems
'We, who work with youngsters, understand these problems'
c'. *Nos co i laurun con i mituns capiun chisc problems
d. Os **ch'** i ponsëis demassa ne fajëis mai nia
you who scl think.pr.2p too much neg do.pr.2p never nothing
'You, who are thinking too much, never do a thing'
d' *Os co i ponsëis demassa ne fajëis mai nia

This person split could be accounted for in semantic terms. First and second person are distinct from third as they are specified for [+deictic] features: being directly involved in the discourse and being the parties present in the conversation, their reference is clearly and unequivocally identified. The referent of third person, on the other hand, always needs further information to be specified. This distinction resembles the opposition between two types of relative clauses, restrictive and non-restrictive (or appositive): it could be argued that when a first or second person are relativised they can only appear in non-restrictive relatives, since their referent, being already unequivocally identified, does not need any further specification. Consequently, *co* could be analysed as being the relative pronoun specified for some [restrictive] feature, only realised when introducing restrictive relative clauses. The following examples show that this is not the case: although the referent of the relativised element is uncontroversial and therefore the relative is of the non-restrictive type, the relative pronoun can only be *co*:

(15) a. La Talia, **co** à les leges dër rigoroses, prodüj
the Italy who have.pr.3s the laws very rigorous produce.pr.3s
le miù ere d' orì
the best oil of olive
'Italy, that has very strict laws, produces the best olive oil'
a'. *La Talia, **ch(e)'** à les leges dër rigoroses, prodüj le miù ere d'orì
b. La löna, **co** lomina ensnet dër sterscia, roda encër
the Moon who shine.pr.3s tonight very bright circle.pr.3s around
la tera
theEarth
'The Moon, that tonight shines very bright, revolves around the Earth'
b'. *La löna, **che** lomina ensnet dër sterscia, roda encër la tera

Let us now investigate an alternative, purely syntactic, analysis focusing on [φ] features. It has been claimed in a large body of literature starting from Benveniste (1966) (cf. Anagnostopoulou, 2003 more recently) that third person pronouns are 'determiner pronouns', in other words, they lack person features and are only specified for [number]. First and second person pronouns, on the other hand, are 'fully fledged' pronouns and are specified for both [person, number] features. Following this pronoun specification, we claim that *co* is triggered when an Op specified for [number] features, and lacking [person] features, reaches [Spec, Force]; *che*, on the other hand, surfaces when the element is specified for [person, number] features, irrespective of whether it has moved through [Spec, TP] or not. It thus seems that the [φ] feature specification overrides the checking history mentioned before: having moved through [Spec, TP] is a necessary but not sufficient condition to trigger *co*. We could view *co* as a relative pronoun underspecified for [person] features related to a subject[10].

The difference in person with respect to the choice of relative pronoun between Mar and French could be explained simply by assuming that while Mar *co* is sensitive to [φ] features, French *qui* is not, and it is purely dependent on the subject/object distinction. This is a tentative analysis rather than a definite conclusion, and further research is necessary in order to provide a more formal account for the person asymmetry witnessed in Mar and the way it differs from French.

Finally, referring back to the issue of reduplication of inflectional features, in Mar, too, we witness a reduced version of [φ] features, based on a [±] binary value, being expressed at the C level. From the limited evidence investigated here we could therefore conclude that the inflectional information that finds expression at the C level is a rudimentary version of the one encoded within the IP. More data is needed to provide a more complete answer to this question.

5. Concluding remarks

In this paper we have investigated the category traditionally labelled as COMP, drawing on some Northern Italian varieties, early Romance texts and French acquired as a first language. The different constructions examined have revealed that COMP performs a variety of tasks, a subordinating particle, a mood marker, a topic marker and an element specified for [φ]

[10] Given that *che* is the counterpart of the relative pronoun and complementiser in standard Italian, we would expect this to be the unmarked form: in other words, we would expect *co* to be the form encoding 'more' information, such as [person] features. This is not the case, as pointed out by an anonymous reviewer, and at this stage we don't have an answer to this observation.

features. These are properties of both the inflectional and complementiser domains, underlining the fact that IP information can also find expression, although in a reduced way as we have seen, at the C level. A clear quantification of exactly what and in what terms is repeated is still needed, and this will be the object of further research.

References

Anagnostopoulou, Elena. 2003. *The Syntax of Ditransitives: Evidence from Clitics*. Berlin, Mouton de Gruyter.

Benincà, Paola. 1995. "Complement Clitics in Medieval Romance: the Tobler-Mussafia Law". In A. Battye and I. Roberts (eds.) *Clause Structure and Language Change*. 325-344. New York/Oxford, Oxford University Press.

Benincà, Paola. 2001. "The position of Topic and Focus in the left periphery". In G. Cinque and G. Salvi (eds.) *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. 39-64. Amsterdam, Elsevier.

Benincà, Paola and Cecilia Poletto. 2004. "Topic, Focus and V2: defining the CP sublayers". In L. Rizzi (ed.) *The Structure of CP and IP – The Cartography of Syntactic Structures*, Vol. 3. 52-75. New York/Oxford, Oxford University Press.

Benveniste, Emile. 1966. *Problèmes de linguistique générale*. Paris, Gallimard.

den Besten, Hans. 1983. "On the Interaction of Root Transformations and Lexical Deletive Rules". In W. Abraham (ed.) *On the syntax of Westgermania*. 47-131. Amsterdam/Philadelphia, John Benjamins.

Chomsky, Noam. 1980. "On Binding". *Linguistic Inquiry* 11:1-46.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht, Foris.

Chomsky, Noam. 1986a. *Barriers*. Cambridge MA, MIT Press.

Chomsky, Noam. 1986b. *Knowledge of Language: Its Nature, Origin and Use*. New York, Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA, MIT Press.

Chomsky, Noam. 2002. *On Nature and Language*. Cambridge, Cambridge University Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. New York/Oxford, Oxford University Press.

Elliot, Jennifer R. 2000. "Realis and Irrealis: Forms and Concepts of the Grammaticalisation of Reality". *Linguistic Typology* 4:55-90.

Enç, Mürvet. 1987. "Anchoring Conditions for Tense". *Linguistic Inquiry* 18.4:633-657.

von Fintel, Kay. 1994. *Restrictions on Quantifier Domains*. Ph.D. dissertation, University of Massachusetts.

Frajzyngier, Zygmunt. 1995. "A Functional Theory of Complementizers". In J. Bybee and S. Fleischman (eds.) *Modality in Grammar and Discourse*. 473-502. Amsterdam/Philadelphia, John Benjamins.

Giorgi, Alessandra and Fabio Pianesi. 1997. *Tense and Aspect: from Semantics to Morphosyntax*. New York/Oxford, Oxford University Press.

Haegeman, Liliane. 2003. "Conditional clauses: external and internal syntax". *Mind and Language, Numéro spécial* 18.4:317-339.

Kayne, Richard S. 1976. "French relative *que*". In F. Hensey and M. Lujan (eds.) *Current Studies in Romance Linguistics*. Washington DC, Georgetown University Press.

Kayne, Richard S. 1978. "Le condizioni sul legamento, il collegamento dei clitici e lo spostamento a sinistra dei quantificatori". *Rivista di Grammatica Generativa* 3:147-171.

Keniston, Hayward. 1937. *The Syntax of Castilian Prose – The 16th century*. Chicago, The University of Chicago Press.

Munaro, Nicola. 2005. "Computational Puzzles of Conditional Clause Preposing". In A. M. Di Sciullo (ed.) *UG and External Systems*. 73-94. Amsterdam/Philadelphia, John Benjamins.

Palmer, Frank R. 2001. *Mood and Modality*. Second edition. Cambridge, Cambridge University Press.

Pesetsky, David. 1982. *Paths and Categories*. Ph.D. dissertation, MIT.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. New York/Oxford, Oxford University Press.

Rizzi, Luigi. 1997. "The fine structure of the left periphery". In L. Haegeman (ed.) *Elements of Grammar.* 281-337. Dordrecht, Kluwer.

Rizzi, Luigi. 2004. "On the Form of Chains: Criterial Positions and ECP Effects". Ms., Università di Siena.

Rohers, Dorian and Marie Labelle. 2003. "The Left Periphery in Child French: Evidence for a Simply-Split CP". In J. Quer et alii (eds.) *Romance Languages and Linguistic Theory 2001*. 279-294. Amsterdam/Philadelphia, John Benjamins.

Schiaffini, Alfredo. 1954. *Testi Fiorentini del Dugento e dei primi del Trecento*. Firenze, Sansoni.

von Stechow, Arnim. 1995. "On the Proper Treatment of Tense". Ms., University of Tübingen.

Stowell, Tim. 1982. "The Tense of Infinitives". *Linguistics Inquiry* 13.3:561-570.

Wanner, Dieter. 1995. “Les subordonnées à double complémentateur en roman medieval”. In G. Ruffino (ed.) *Atti del XXI Congresso Internazionale di Linguistica e Filologia romanza*, Section 1, 421-433.

Abbreviations of early Romance texts:
Con = *Conti di Antichi Cavalieri*, 13th-century Tuscan, OVI
DistrTr = *Da un libro della distruzione di Troia*, 13th-century Tuscan, OVI
TrRicc = *Il Tristano Riccardiano*, 13th-century Tuscan, OVI
Gen Est = *General estoria*, in Wanner (1995)
Apoll = *Le roman d'Apollonius de Tyr*, in Wanner (1995)
Schiaff = *Testi fiorentini del Dugento e dei primi del Trecento*, in Benincà (1995) [Schiaffini (1954)].

On some properties of subjects and topics

Luigi Rizzi
University of Siena

1. Introduction

Subjects bear certain interpretive similarities to topics, and yet the two notions are clearly distinct, both in their formal status, position, and interface properties. In this article I would like to identify some analogies and differences between subject and topic in the context of the general issue of the causes of phrasal movement. After a brief discussion of the "movement as last resort" guidelines, I would like to address the question of the interface effects related to movement to subject position. This will lead to a discussion of some interpretive similarities and differencies between subject and topic, and to the possible restatement of the EPP as a Subject Criterion. The consequences of this approach for the analysis of subject-object asymmetries induced by movement will be briefly mentioned. The article will be concluded by an analysis of the discourse conditions which must be met for a felicitous use of Clitic Left Dislocation topics in Italian.

2. The EPP and Movement as Last Resort

In its traditional form, the Extended Projection Principle (EPP), or, more precisely, the extended clause of the Projection Principle, is the statement that the subject position is an obligatory component of the form of clauses, regardless of whether the thematic grid of the verb contains a role for the filler of this position (Chomsky 1981).

(1) EPP: Clauses must have subjects.

Under this classical view, the EPP position is a potential thematic position, the site where external arguments receive their thematic role, and in this sense it directly falls under the Projection Principle; but the proviso is added (the extended clause (1)) that the position is formally obligatory even if no theta role is available for the external argument in the theta grid of the verb. In this case, the position is filled by an expletive, a non-referential pronominal element.

(2) a. There came a man
b. It seems that John left

Under the VP-internal Subject Hypothesis (Koopman and Sportiche 1991 and many related references) things change: the subject position ceases to be a potential thematic position, and becomes a target of movement in all clauses, a position which the thematic subject moves to after receiving its theta role in the verb phrase:

(3) John has [t seen Mary]

The link to the Projection Principle was severed under this approach, but the label EPP somehow survived to designate the obligatoriness of the position. The semantic drift of the label became more extreme in Minimalism, where EPP came to designate the property that any head may have of c-selecting a specifier, which may be targeted by movement (Chomsky 2000). We will not be concerned with this interpretation here (the EPP feature in this sense was replaced by the more perspicuous OCC (urrence) feature in Chomsky (2004)).

Once the EPP position, in the sense of (1), is seen as a systematic landing site for movement, last resort guidelines raise the question of the functional motivation associated to this kind of movement: why does movement to subject position take place? It is instructive in this respect to compare movement to subject to another much discussed type of plausible A-movement, the movement of the object to a position in the lower IP space to trigger participial agreement in French (Kayne 1989):

(4) a. La chaise que nous avons t repeint*e* t
the chair that we have repainted+Agr
'The chair that we have repainted'
b.*Nous avons la chaise repeint*e* t
we have the chair repainted+Agr
'We have repainted the chair'

An object moved to a higher position (to the relative complementizer system in (4a)) may trigger participial agreement in passing, but cannot stop in the participial agreement position (see (4b)). What precluded (4b)? It has been proposed that movement is a "last resort" operation, there is no "free" or truly optional movement (Chomsky 1986 and much subsequent work), a movement chain must be motivated by the fact that it satisfies some need which could not be satisfied otherwise. Reinhart (1995) and Fox (1997) identify such motivating factor in a requirement of the interface systems:

(5) Movement as last resort: A movement chain must determine some interface effect.

So, movement of the object in (4a) determines the interface effect of displacing the object to the relative clause complementizer system, where it can be interpreted as a relative operator (the chair x such that […x…]). On the other hand, the movement chain in (4b) determines no interface effect (apart from the triggering of morphological agreement on the participle; but evidently this morphological "effect" does not suffice, presumably because it does not satisfy a requirement of the moved element: we have here a residue of the notion "greed" of early versions of Minimalism). So, while the movement of the object through the participial agreement position on its way to the position of the relative operator is legitimate, the movement chain cannot legitimately terminate in a position unrelated to an interface effect in the intended sense, and (4b) is ruled out.

In the approach of Rizzi (2003), interpretively relevant positions like the left peripheral positions of relative and interrogative operators, topic and focus, etc. are criterial positions, positions defined by special, interpretively relevant features *Q, R(elative), TOP, FOC*, notated with capital labels (on the system of Criteria, see Rizzi 1991, 1997). These dedicated positions, and the features associated to them, signal to the interface systems that their specifiers must receive special interpretive properties, properties which Chomsky (2004) refers to as "properties of scope-discourse semantics": a relative or interrogative operator in its scope position, a topic, a focus, etc. Movement chains can involve intermediate steps, from the s-selection to the criterial position, as a consequence of the inherently local nature of movement. Again in the formalism of Rizzi (2003), intermediate movement is triggered by purely formal featural counterparts of the criterial features, notated with low case labels *q, r, top, foc*; in other systems, intermediate movement may be triggered by an unspecific peripheral feature (Chomsky 2005), or may be untriggered; all these systems have in common that the positions targeted by intermediate movement and the criterial positions are

formally distinct in a clear manner, and this is the property we need for a simple checking of the "movement as last resort" requirement. So, the movement of the relative operator to the left peripheral criterial position may involve an intermediate step to the Spec of a lower inflectional head, plausibly Asp, and the passage through this position is morphologically marked by participial agreement. Under a view adopting some version of principle (5), a movement chain (at least for certain phrasal movements) must terminate in a criterial position. This happens in (6), but not in (7), which is therefore excluded by the movement as last resort requirement, a conclusion which can be reached by local inspection. Whatever formal way one adopts to differentiate criterial and non-criterial positions, movement must terminate in a criterial position:

(6) La chaise Op C_R nous avons t Asp_r repeinte t

(7) *Nous avons la chaise Asp_r repeinte t

Object shift languages, involving movement of the object to a position akin to the one in (7), assign special interface properties to this position (specificity, and possibly other scope-discourse properties), so that the process complies with the last resort requirement. A process akin to object shift is possible in French when the object is a bare quantifier, as in (8a):

(8) a. Il a tout acheté t
he has all bought
'He has bought everything'
b. Il a acheté tout le terrain
he has bought all the land
'He has bought all the land'

Presumably here movement is motivated by the requirements of the bare quantifier as a "weak element", in the sense of Cardinaletti and Starke (1999), a requirement which, whatever its exact status (be it a PF, morphological, or LF requirement), suffices to comply with (5).

3. Movement to Subject Position and Last Resort.

Going back to subject movement, the comparison between (3) and (4b) leads to the conclusion that the subject position must be associated to some interface effect which makes it possible to terminate the chain there. What kind of interface effect?

A standard answer to this question involves the internal interface with morphology: subjects move in order to acquire/check nominative Case, and, more generally, to satisfy the properties of the Case-agreement system. So, the classical story goes, if in a SVO language nominative is assigned / checked in a Spec-head configuration, the subject must move to the Spec of Agr (or of whatever functional head bears the Phi features involved in this system) to have its Case assigned/checked, and at the same time to value the agreement features on the clausal head.

This classical approach may well be correct in part, but it could hardly be the whole story. A simple argument that the Case-agreement approach is not sufficient to cover all the cases of subject movement is provided by the existence of the quirky subject phenomenon, as Cardinaletti (2004) points out: in some languages, the canonical subject position can be filled by a nominal which bears an inherent (often dative) Case, and which does not trigger verbal agreement. For instance, with the so-called *piacere* class of psych-verbs in Italian, the unmarked order is Dative Experiencer - V - Nominative Theme (Belletti and Rizzi 1988), as in (9a), with the possibility of the marked order variant Nominative Theme - V - Dative Experiencer, as in (9b). Whatever order is chosen, the inflected verb agrees with the nominative argument:

(9) a. A Gianni piacciono queste idee
to Gianni please these ideas
b. Queste idee piacciono a Gianni
these ideas please to Gianni
'Gianni likes these ideas'

One type of evidence put forth by Belletti and Rizzi (op.cit.) for the conclusion that the dative experiencer is indeed in subject position in (9a), and not in topic, is that it does not interfere with movement, while a preposed (dative) topic determines at least a weak degradation on a movement chain which crosses it:

(10) a. Le idee che a Gianni piacciono di più sono queste
the ideas that to Gianni please most are these
'The ideas that Gianni likes most are these'
b.?(?) Le idee che a Gianni Maria raccomanda sono queste
the ideas that to Gianni Maria recommends are these
'The ideas that Gianni recommends to Maria are these'

The rich literature on the Icelandic equivalent of this construction provides a very varied array of evidence pointing to the same conclusion. See, e.g., Sigurdsson (2000).

Here, movement of the dative experiencer to subject position is not linked to any visible Case-agreement effect. It is sometimes assumed that the quirky subject bears an invisible structural Case, on top of the morphologically overt inherent Case, which motivates movement to subject position. This assumption is particularly implausible in languages like Italian, in which the Dative argument is prepositional, and a PP never is the bearer of any Case specification. Analogous difficulties for a Case approach to subject movement are raised by other cases in which the filler of the subject position is non involved in the Case-agreement system: locative inversion, and the inverse copular construction (Moro 1997), see Cardinaletti, op. cit., for detailed discussion.

For these reasons I will assume that reference to the requirements of the Case-agreement system is insufficiently general to motivate movement to subject position. This leads us to look into the possibility that some kind of interpretive requirement linked to the subject position may be the factor motivating movement to subject.

4. Subject interpretation

A salient interpretive property of preverbal subjects is that they can correspond to given information, while VP internal arguments, a direct object or a *by* phrase in passive, normally are focal, or in any event don't easily convey given information. Consider question-answer pairs like the following, in which the question introduces a referent which is taken up in the answer. The introduced referent can be the subject in an active (11A) or passive (12B) sentence, while it cannot be a direct object (12A) or a *by* phrase in passive (11B):

(11) Q: Che cosa è successo al camion?
'What happened to the truck?'
A: Il camion ha tamponato l'autobus
'The truck bumped into the bus'
B: # L'autobus è stato tamponato dal camion
'The but was bumped into by the truck'

(12) Q: Che cosa è successo all'autobus?
'What happened to the bus?'
A: # Il camion ha tamponato l'autobus
'The truck bumped into the bus'

B: L'autobus è stato tamponato dal camion
'The bus was bumped into by the truck'

The unmarked answers would have the introduced referent taken up by a null pronominal subject in these cases, i.e., ___ *è stato tamponato dal camion* (___ was bumped into by the truck) is the unmarked answer to question (12)Q. Still it is possible to reiterate the DP in subject position in the answer, as in (12B), etc., which sounds slightly redundant, and gives rise to a stylistic effect of insistence or reiteration of the discourse topic. The contrast with the reiteration in VP final position (yielding a non-felicitous result, as in (12A)) is clear.

The distinction also holds, though somewhat weakened, between subjects and other elements which are not VP final, as the object in a ditransitive construction:

(13) Q: A chi hanno dato il mio libro?
'To whom did they give my book?'
A: # Hanno dato il tuo libro a Gianni (OK Lo hanno dato a Gianni)
'They gave your book to Gianni' (OK They gave it to Gianni)

(14) Q: A chi è stato dato il mio libro?
'To whom was your book given?'
A: Il tuo libro è stato dato a Gianni
'Your book was given to Gianni'

Reiteration of the direct object does not sound felicitous in (13A) (the felicitous answer being the one which involves cliticization of the object), even though the deviance is less marked here than in the cases in which the relevant DP is in VP final position, as in (11B), (12A), the stronger deviance of the latter examples being presumably related to fact that a focal interpretation is necessarily associated with the VP-final element. Still, the contrast between (13A) and (14A) is clearly detectable: (14A) is felicitous, with the special stylistic effect induced by reiteration, as in (12B).

These facts could suggest the hypothesis that the EPP position is specialized for the expression of contextually given information. In this sense, its discourse function would be analogous to the discourse function of the topic, as has often been proposed. Notice, in fact, that the equivalents of (11A), (13A) are felicitous if the object becomes a topic in the Clitic Left Dislocation (see Cinque 1990), the typical topic construction used in the Romance languages:

(15) Q: Che cosa è successo all'autobus?
'What happened to the bus?'
A: L'autobus, il camion lo ha tamponato
'The bus, the truck bumped into it'

(16) Q: A chi hanno dato il mio libro?
'To whom did they give my book?'
A: Il tuo libro, lo hanno dato a Gianni
'Your book, they gave it to Gianni'

Here too, as in (11), (12), (14) for the subject, the overt expression of the Cl Left Dislocated element makes the structure slightly redundant, involving a special stylistic effect of reiteration.

In spite of this analogy, there is clear evidence that a full functional unification of subject and topic is not possible. First of all, preverbal subjects are fully felicitous when they express new information in out of the blue ("What happened?") contexts, contexts in which a (ClLD) topic is not felicitous:

(17) Q: Che cosa è successo?
'What happened?'
A: Un camion ha tamponato un autobus
'A truck bumped into a but'
A': Un autobus è stato tamponato da un camion
'A bus was bumped into by a truck'
A'': # Un autobus/l'autobus per Roma, un camion lo ha tamponato
'A bus/the bus for Rome, a truck bumped into it'
A''':# Un autobus/l'autobus per Roma, lo ha tamponato un camion
'A bus/the bus to Roma, bumped into it a truck'

An indefinite (new information) initial subject is fine in this context both in an active and a passive sentence, while a topic, whether definite or not (indefinite topics are possible in Italian in particular discourse contexts, on which see below) is not felicitous, whether the clause has an initial or final subject, as in A'', A'''.

These simple differences in appropriateness corroborate much well-known distributional evidence supporting the distinction between subject and topic position also in Null Subject Languages, for instance the fact that certain quantified DP's cannot occur as (ClLD) topics, whereas they are fully acceptable as subjects (Rizzi 1982, 1986, Cardinaletti 2004):

(18) a.*Nessuno, Piero lo ha visto
no one, Piero him has seen
b. Nessuno ha visto Piero
no one has seen Piero
'No one saw Piero'

This sharp fact is usually interpreted as showing that the ClLD position and the subject position should be kept distinct even in a Null Subject Language like Italian; it also shows that subjects and topics are not functionally equivalent, the former permitting a larger class of elements to fill the position than the latter.

Going back to subject interpretation, it should be noticed that selection of a single subject argument also takes place in cases in which the event involves two (or more) arguments which are both contextually given. In the following case, both arguments are introduced in the previous request of information, and the discourse may naturally continue by selecting either one as the starting point in the description of the event, in an active or a passive structure:

(19) Q: So che il camion e l'autobus hanno avuto un incidente, ma dimmi esattamente che cosa è successo.
'I know that the truck and the bus had an accident, but tell me exactly what happened'
A: Il camion ha tamponato l'autobus
'The truck bumped into the bus'
A': L'autobus è stato tamponato dal camion
'The bus was bumped into by the truck'

Examples (17) and (19) show that, quite independently from the articulation of the informational structure, one argument (which may be given information or not) is selected as the point of departure in the description of the event, which is described as being "about" that argument.

The choice of one particular argument as subject has the effect of making it prominent, in a sense which affects the following discourse: for instance, a *pro* subject in a following sentence can only pick up the previous subject (the "thema" in Calabrese's 1986 sense). So, consider a continuation like (20):

(20) … poi *pro* è ripartito
then *pro* is left
'…then he left'

If (20) is uttered after (19A), *pro* necessarily refers to the truck, while after (19A') it can only refer to the bus. The same selective effect is found if we consider (20) as a discourse continuation following (17A) and (17A'): *pro* always picks up the surface subject, the vehicle that left is, respectively, the truck and the bus. So, active and passive are not interchangeable in discourse, even in cases in which the interplay between new and given information seems irrelevant (because both arguments are new information, as in (17), or given information, as in (19)).

5. The Subject Criterion

A subject shares with a topic the prominence related to the fact that the described event is presented as being about that argument ("aboutness"); it differs from a topic (at least, a topic of the ClLD kind) in that it does not require the discourse-related property (expressed as Discourse-linking here, as in Pesetsky (1987), but see below for a more detailed analysis) which makes such topics infelicitous in out of the blue contexts. So, a topic involves aboutness and D-linking, while a subject involves pure aboutness:

(21) Top: [+ aboutness]
[+ D-linking]

(22) Subj: [+ aboutness]

We are now ready to restate the EPP as a Subject Criterion. Following in essence Cardinaletti (2004), I assume a Subj head in the high part of the IP system, higher than the head carrying Phi features and responsible for the Case-Agreement system (Cardinaletti in fact argues for a further distinction between a Subj and an EPP head, both distinct from the head which carries the agreement features: see below for discussion).

(23) … Subj … Phi …

Subj is endowed with nominal features which make it attract a nominal expression to its Spec. Such a featural characterisation must be broader than one merely involving a reference to Phi features, as the capacity of attraction that this position has extends beyond the capacity of the head responsible for the Case-agreement system: not only nominative subjects triggering verbal agreement may be attracted to Spec Subj, but also other elements, with a certain amount of cross-linguistic variability: quirky subjects, possibly locatives in the locative inversion construction, possibly nominal predicates in the inverse copular construction (Moro 1997, Cardinaletti 2004), as suggested in the previous discussion. I will just state the selective attraction of different kinds of nominal expressions by assuming a [+N] specification,

keeping in mind that a finer identification of the relevant nominal features may be needed.

So, the Subj head determines the subject-predicate articulation, much as Top determines the topic-comments articulation, and Focus determines the Focus-Presupposition articulation.

(24) [XP [Top YP]]

(25) [XP [Foc YP]]

(26) [XP [Subj YP]]

Under the criterial view, movement to the subject position is movement to a position dedicated to a special interpretive property of the scope-discourse kind ("aboutness"), which has reflexes on the organisation of discourse (as shown, e.g., by the interpretation of (20)); as such, it is akin to the other cases of criterial movement to the left periphery of the clause, and complies with the "movement as last resort" guidelines.

But how can the existence of expletives be reconciled with the criterial view? Clearly, in such cases it cannot be said that the event is presented as being about the expletive, which has no argumental status: in fact, no argument is chosen as the subject of predication in that case; this straightforward observation is often taken as a stumbling block against attempts to identify special interpretive properties of the subject position, and as arguing in favor of a purely formal approach to the EPP. Nevertheless, the interpretive peculiarities of the subject, supporting an interface role of the position, remain. How can these conflicting considerations be reconciled?

I would like to suggest that the Subj layer, at the junction of the CP and IP systems, shares certain properties of both. It shares with CP positions functioning as landing sites of movement the dedicated character to scope-discourse types of properties, hence the criterial character. On the other hand, Subj is an obligatory position, obligatoriness being a property that it shares with the backbone of the IP structure (Cinque 1999). So, while the left peripheral criterial heads of topic and focus are formally optional, and present in the structure only when discourse conditions and communicative intentions require them, Subj is always there, much as T and related functional heads. Expletives can then be seen as a way to reconcile these conflicting requirements. When communicative intentions, discourse conditions and the thematic properties of the predicate require a non-predicational sentence, an expletive formally complies with the requirements

of Subj, thus conveying the interpretation that the event is not presented as being about a particular argument.

That expletives may not be inconsistent with criterial configurations is also suggested by unquestionably criterial constructions of the A' type. Consider the so-called "partial wh movement" constructions, under an analysis like the one presented in McDaniel (1989) for German and other languages: in the following example, the substantive wh operator *welchen Mantel* sits in a lower C system, and the criterial position, the main C system in a direct question, is filled by the expletive-like wh element *was*:

(27) Was glaubst du welchen Mantel Jakob heute angezogen hat?
what believe you which coat Jakob today put-on has?
'Which coat do you believe that Jakob put on today?'

The use of expletives is more conspicuous and cross-linguistically stable in Spec-Subj, due to the obligatoriness of the position, but it is important to observe that the possibility of an expletive fulfilling a criterion is attested elsewhere (see also Benincà's (2005) discussion of the distribution of *sì* in Old Italian, which she interprets as an expletive filler of the focus position in the V-2 construction). Pursuing the parallel with T, the vacuous satisfaction of the Subject Criterion in presentational structures may be akin to the use of the unmarked tense, present, in statements expressing logical or mathematical truths: such truths are atemporal, but the obligatoriness of the T position enforces the choice of a tense form, the unmarked present form, rather than permitting the expression of an untensed structure.

A salient additional difference between movement to subject and to topic is the highly local character of the former. Movement to subject is basically constrained to affect the closest nominal, the fundamental locality property of A-chain, while movement to topic is much freer, basically the freest form of A'-chain. If Subj attracts a [+N] element, it will always involve the closest nominal, and strict locality is expected (both the dative and nominative arguments should therefore count as "the closest nominal" in cases like (9a-b), whence the possible word-order alternation; I leave open here how this "equidistance" effect can be captured, but see Rizzi (2003) for discussion). As for topics, the special freedom of this construction can be made to follow from the typology of A' features and the cartographic assumptions introduced in Rizzi (2004).

A final comment concerns the relation between the proposed criterial interpretation of the EPP and the cartography of the IP. Recent cartographic studies show that more IP-initial subject positions must be postulated than (23) suggests. Cardinaletti (2004) argues that at least an additional EPP

position must be assumed, where expletives, *pro* and (certain) weak pronouns stop, whereas lexical subjects may move to a distinct and higher Subject-of-predication position; Shlonsky (2000) provides a very refined typology of subject positions in Semitic, etc. I will not be able to address the refinements required by these results here, but I would like to note that a finer cartography can be made compatible with the current approach if it is assumed that the EPP head (our Subj) involves criterial satisfaction of "pure aboutness", while other positions may be specialized for particular kinds of subjects (with lexical content, etc.). This can be made compatible with Criterial Freezing through the mechanism of head movement, which can create a complex head composed of distinct criterial heads, giving rise to a Spec-head configuration in which different criteria can be satisfied simultaneously, a mechanism which is independently required for multiple satisfaction of A'-criteria (Rizzi 2003, fn. 8).

6. Subject Criterion, Criterial Freezing and ECP Effects

The criterial analysis of the EPP has implications for the subject-object asymmetries traditionally ascribed to the ECP. These implications are explored in detail elsewhere (Rizzi 2003, Rizzi and Shlonsky in prep.), and will only be touched upon here. Rizzi (2003) argues for a principle, dubbed "Criterial Freezing", which has the effect of freezing in place an element meeting a criterion: for instance, the wh element meeting the Q Criterion in an embedded C system in an indirect question (28a) is unavailable to further syntactic movement, e.g., cannot be moved to a higher focus/clefting position (28c), while focalization in situ is fine (28b), and also pied piping of the whole indirect question (28d):

(28) a. Mi domando [per chi C votare]
'I wonder for whom to vote'
b. Mi domando [PER CHI C votare], non contro chi
'I wonder for whom to vote, not against whom'
c.* E' per chi che mi domando [___ C votare], non contro chi
'It is for whom that I wonder ___ to vote, not against whom'
d. E' [per chi C votare] che mi domando ___, non contro chi
'It is for whom to vote that I wonder ___, not against whom'

It should be noted that (28c) is not trivially excluded by the violation of the selectional requirement of the verb *domandarsi* (wonder), which selects an indirect question: under the copy theory of traces, the embedded C system contains a silent occurrence of the wh phrase *per chi*, which should be

sufficient to satisfy the verb's selectional requirement. A formal principle like Criterial Freezing thus seems to be needed.

So, criterial satisfaction appears to terminate a chain, much as s-selection normally initiates a chain. Apparently, as soon as a scope-discourse criterion is satisfied, the chain is sent to the interface systems, and the element satisfying the criterion is unavailable to further syntactic computation.

If the subject position is criterial, we expect subjects to be unmovable in the normal case, as a consequence of Criterial Freezing. This immediately derives the familiar subject-object asymmetry of the kind illustrated by the following pair in French, the major empirical effect of the Empty Category Principle (ECP) of Chomsky (1981):

(29) a.* Qui crois-tu que t viendra?
'Who do you believe that will come?'
b. Qui crois-tu que Marie va rencontrer t?
'Who do you believe that Marie will meet?'

The thematic subject *qui* must fulfil the Subject Criterion in (29a), but then it cannot be moved further under Criterial Freezing. The object in (29b) remains freely movable, as there is no Object Criterion (no EPP for objects, in the classical sense). The asymmetry is thus explained.

From the viewpoint of this approach, the strategies of subject extraction that languages use typically amount to finding a device to fulfil the Subject Criterion without requiring the thematic subject to move to that position, so that the thematic subject remains available for further movement. Consider for instance the absence of that-trace effects in a Null Subject Language like Italian:

(30) Chi credi che verrà?
'Who do you think that will come?'

As argued for in Rizzi (1982) and much subsequent work, there is evidence that the wh element *chi* is not extracted from the clause-initial subject position (the criterial position, in our current terms), but from a lower position, while the initial position is filled by expletive *pro*:

(31) Chi credi che [*pro* verrà t]?

In the current terms, *pro* formally satisfies the Subject Criterion, thus permitting the thematic subject to remain available for further syntactic movement (but not necessarily from the "inverted" position, which is

dedicated to the focus interpretation, see Belletti 2001, 2004 and related references). Different languages use variants of this strategy, or other devices finalized to make the thematic subject available to movement and extraction. See Rizzi and Shlonsky (in prep.) for the discussion of many cases.

7. Indefinite topics, specificity and partitivity

In the previous sections we have characterized topics as necessarily involving a connection to the discourse, or D(iscourse)-linking. This accounts for the fact that topics cannot felicitously occur in out of the blue contexts. A problem for this characterization is raised by the fact that Italian permits indefinite topics. In particular discourse conditions, the following examples are fully natural:

(32) a. Un libro così interessante, non lo avevo mai letto
a book so interesting, not it had never read
'Such an interesting book, I had never read'
b. Una soluzione, la dobbiamo trovare anche stavolta
a solution, it must find also this time
'A solution, we must find this time as well'
c. Un articolo ben scritto, tutti lo leggono sempre volentieri
an article well written, everybody it read always willingly
'A well-written article, everybody reads it always willingly'

Not only does Italian allow indefinite left-dislocated elements: even in cases in which the indefinite element is clearly non-specific, the construction remains felicitous (in particular contexts, on which see below). For instance, a non-specific indefinite modified by a relative clause is formally marked by the subjunctive mood in the relative clause, and clitic left dislocation of such a phrase remains possible:

(33) Una segretaria che sappia tenere la contabilità del dipartimento, non
a secretary who can keep the budget of the department, not
riesco proprio a trovarla
manage really to find-her
'A secretary who could keep the department budget, I really can't find her'

And of course, Clitic Left Dislocation may involve specific indefinites, as in the following example, in which the specific interpretation is enforced by the tag:

(34) Una buona cosa, Gianni l'ha fatta: ha smesso di fumare
one good thing, Gianni it has done: has given-up of smoking
'A good thing, Gianni did it: he quit smoking'

The possibility of indefinite (specific and non-specific) topics seems to question the assumption that topicality is linked to contextually given information: a DP already introduced in discourse may be reiterated, but it typically is introduced by the definite determiner ("A man came to my house. The man carried a heavy suitcase,…"). So, a sharper characterisation of the notion of discourse-linking involved is necessary (a similar conclusion is reached in Starke (2001) in connection with the selective sensitivity of different kinds of D-linked wh phrases to island phenomena).

Notice that a closely related language, French, which also possesses a kind of Clitic Left Dislocation construction, seems to disallow indefinite topics, whatever their status with respect to specificity (native speakers' judgments on examples like the following range from marginal to impossible; I have reported here the most restrictive set of judgments):

(35) a.* Un livre si intéressant, je ne l'avais jamais lu
'Such an interesting book, I had never read (it)'
b.* Une solution, il faudra la trouver cette fois aussi
'A solution, we must find (it) this time as well'
c.* Une secrétaire qui sache tenir la comptabilité du département, je n'arrive pas à la trouver
'A secretary who could keep the department budget, I really can't find (her)'
d.* Une bonne chose, Jean l'a faite: il a arrêté de fumer
'A good thing, Gianni did (it): he quit smoking'

So, the discourse conditions licensing ClLD in French are more restrictive, and transparent: the referent of the topic must have been introduced in previous discourse, so that it is taken up by a definite DP. Dominique Sportiche points out that some of these examples improve if the genitive-partitive clitic *en* (of it, of this kind) is used instead of the definite accusative clitic *la/le*:

(36) ? Une secrétaire qui sache tenir la comptabilité du département, je n'arrive pas à en trouver
'A secretary who could keep the department budget, I really can't find '

This suggests that the definiteness restriction may be a function of the clitic doubler, rather than an inherent property of the topic position in French.

In Italian things are more complex, but the construction is not unconstrained: we have already seen that ClLD cannot be used felicitously in out of the blue context, and this extends to the cases of an indefinite topic. Some kind of connection to the discourse background is needed, but the link is more subtle.

Let us consider the issue in more detail, starting with the discussion of a "real" example taken from a magazine. The article starts with the following text:

(37) a. Quanto costa all'ambiente un telefono cellulare?
'How much does a cell phone cost to the environment?'
b. … domande … cruciali per ecologisti, economisti, e politici…
'…questions … crucial for ecologists, economists, and politicians'
c. Una risposta la fornisce la Life cycle impact assessment (LCIA),…
'An answer, the LCIA provides it,…'
(Internazionale, 18-24/02/05, p. 79)

Sentence (37c) involves an indefinite ClLD topic *Una risposta* ("An answer"). Clearly, the felicity of this text crucially depends on the fact that certain questions have been introduced in the b sentence, which makes the set of possible answers contextually salient. In fact, the text could not have started with an indefinite ClLD (actually, with a ClLD *tout court*): an article could start with sentence (38a), but not with (38b):

(38) a. La LCIA fornisce una risposta a domande cruciali
'The LCIA provides an answer to crucial questions'
b.#Una/la risposta a domande cruciali la fornisce la LCIA
'An/the answer to crucial questions the LCIA provides it (as initial sentence of a text)'

The connection to the previous discourse context is necessary for a felicitous ClLD, but it can be very subtle and indirect. Consider the following two dialogues (e.g., say, between a father and his son preparing a university exam):

(39) A: Mi sembra che ieri non hai fatto granché per preparare l'esame…
'It seems to me that yesterday you did not much to prepare the exam…'

B: Beh, un libro, l'ho letto…
'Well, a book, I read it…'

(40) A: Mi sembra che ieri non hai fatto granché per preparare l'esame…
B: Beh, ho letto un libro…
'Well, I read a book…'

The two dialogues are identical, except that the indefinite direct object is in situ in (40B), and clitic left dislocated in (39B). Now, while (40B) implies nothing about the connections between the book and the exam (there could be no connection at all, the answer could be explicitly paraphrased as "It's true that yesterday I did nothing for the preparation of the exam, but I was engaged in some other activity, I read a book…"), (39B) strongly invites an interpretation in which the book is part of the exam's program, so that a full paraphrase would be "it's not true that I didn't do much for the preparation of the exam yesterday, I read one of the books in the exam's program". So, in order to be licit the indefinite topic must be interpreted as involving a sort of implicit partitive form picking out an element from a set established in discourse, the set of readings for the preparation of the exam.

Consider now the following two dialogues, involving a non-specific indefinite DP "Una brava segretaria" ("A good secretary"):

(41) Q: Come mai Gianni è così nervoso?
'Why is Gianni so nervous?'
A: Cerca da molto tempo una brava segretaria, ma non la trova
'He's been looking for a good secretary for a while, but he can't find her'
A': # Una brava segretaria, la cerca da molto tempo, ma non la trova
'A good secretary, he has been looking for her for a while, but he can't find her'

(42) Q: Gianni ha poi trovato qualcuno che tenga la contabilità del dipartimento?
'Did Gianni find somebody for the departemental bookeping in the end?'
A: Macché! Una brava segretaria, la cerca da molto tempo, ma non la trova
'Not at all! A good secretary, he's been looking for her for a long time, but he can't find her'

In (42Q), the role of taking care of the department budget is introduced, so that in the following sentence (42A) this functional role can be referred to through an indefinite non specific topic "una brava segretaria". No such contextual justification is provided by (41Q), so the indefinite non specific object can be introduced clause-internally, as in (41A), but not in topic position, as in (41A').

These examples clearly show that specificity, in the usual interpretation of the term, is not involved in the licensing of topics in Italian. Rather, the relevant notion appears to be partitivity, defined by Ionin (2003) as follows:

(43) If a DP is [+partitive], it denotes an individual which is a member of a set introduced by previous discourse (cf. Enç 1991, Diesing 1992)

This notion should be carefully distinguished, Ionin argues, from the notion of specificity, which she defines as follows:

(44) If a DP is [+specific], the speakers intends to refer to a unique individual in the set denoted by the NP, and considers this individual to possess some noteworthy property.

This terminology is not used consistently in the literature (for instance, Enç uses the term "specific" in a broader way including Ionin's definition of "partitive" (43), Diesing refers to the "presuppositional" vs. non-presuppositional" distinction, partitivity being an overt case of presuppositionality, etc.; also related to these distinctions is Starke's (2001) distinction between "range-based" and "specificity-based" presuppositionality, introduced to capture certain selective effects in weak island sensitivity, with range-based presuppositionality essentially corresponding to partitivity), but the conceptual distinction is quite clear: in Ionin's terms, partitivity expresses set membership, while specificity expresses the speaker's intent to refer. Only the former appears to be relevant for the licensing of topics in Italian, as the discussed examples suggest.

8. Conclusion

The subject position is a target of movement and, under "movement as last resort" guidelines, movement to subject must be motivated by the satisfaction of some interface requirement. Morpho-syntactic properties of the Case-agreement system are a plausible candidate, but there are cases of movement to subject, primarily the quirky subject phenomenon, which involve elements not involved in the functioning of the Case-agreement system. These phenomena cast doubts on the generality of an approach to

subject movement based on this system. As an alternative, the possibility is worth exploring that the motivating interface effect may be of an interpretive nature. This has led us to consider the interpretive properties linked to the subject position. Subjects share with topics the fact that they are arguments selected as the point of departure in the description of the event, which is presented as "being about" the subject argument. In addition to that, topics also require some connection to the discourse background, a requirement that subjects are exempted from. The identification of an interpretive element of the scope-discourse kind connected to the subject position makes it possible to state a Subject Criterion, which the moved argument satisfies. Under Criterial Freezing, this approach can derive the unmovability of subjects in certain environments, hence offer an alternative to the classical ECP analysis of subject-object asymmetries. As for topics, Italian allows indefinite, and also non-specific indefinite topics, provided that they satisfy a subtle condition of connection to the discourse background, which we have assimilated to Ionin's partitivity condition.

References

Belletti, Adriana. 2001. "Inversion as focalization". In A. Hulk and J.Y. Pollock (eds.) *Inversion in Romance*. 60-90. New York/Oxford, Oxford University Press.

Belletti, Adriana. 2004. "Aspects of the Low IP Area". In L. Rizzi (ed.) *The Structure of CP and IP – The Cartography of Syntactic Structures*. 16-51. New York/Oxford, Oxford University Press.

Belletti, Adriana (ed.). 2004. *Structures and Beyond - The Cartography of Syntactic Structures*. New York/Oxford, Oxford University Press.

Belletti, Adriana and Luigi Rizzi. 1988. "Psych-verbs and Theta Theory". *Natural Language and Linguistic Theory* 6:291-352.

Benincà, Paola. 2005. "A Detailed Map of the Left Periphery of Medieval Romance". In R. Zanuttini et alii (eds.) *Negation, Tense and Clausal Architecture.* Georgetown University Press.

Calabrese, Andrea. 1986. "Some Properties of the Italian Pronominal System: An Analysis Based on the Notion of Thema as Subject of Predication". In H. Stammerjohann (ed.) *Tema-Rema in Italiano.* 25-36. Tuebingen, Gunter Narr Verlag.

Cardinaletti, Anna and Michal Starke. 1999. "The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns". In H. van Riemsdijk (ed.) *Clitics in the Languages of Europe.*145-233. Berlin, Mouton de Gruyter.

Cardinaletti, Anna. 2004. "Towards a Cartography of Syntactic Positions". In L. Rizzi (ed.) *The Structure of CP and IP – The Cartography of*

Syntactic Structures. 115-165. New York/Oxford, Oxford University Press.
Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht, Foris Publications.
Chomsky, Noam. 1986. *Knowledge of Language.* New York, Praeger.
Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge MA, MIT Press.
Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework". In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by Step – Essays in Minimalist Syntax in Honor of Howard Lasnik*. 89-156. Cambridge MA, MIT Press.
Chomsky, Noam. 2004. "Beyond Explanatory Adequacy". In A. Belletti (ed.) *Structures and Beyond - The Cartography of Syntactic Structures*. 106-131. New York/Oxford, Oxford University Press.
Chomsky, Noam. 2005. "On Phases". Ms., MIT.
Cinque, Guglielmo. 1990. *Types of A' Dependencies.* Cambridge MA, MIT Press.
Cinque, Guglielmo. 1999. *Adverbs and Functional Heads*. New York/ Oxford, Oxford University Press.
Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass, MIT Press.
Enç, Mürvet. 1991. "The Semantics of Specificity". *Linguistic Inquiry* 22:1-25.
Fox, Danny. 1997. *Economy and Semantic Interpretation.* Cambridge MA, MIT Press.
Ionin, Tania. 2003. *Article Semantics in Second Language Acquisition*, Ph.D. dissertation, MIT.
Kayne, Richard. 1989. "Facets of Romance Past Participle Agreement". In P. Benincà (ed.) *Dialect Variation and the Theory of Grammar.* 85-103. Dordrecht, Foris Publications.
Koopman, Hilda and Dominique Sportiche. 1991. "The Position of Subjects". *Lingua* 85:211-258.
McDaniel, Dana. 1989. "Partial and Multiple Wh-Movement". *Natural Language and Linguistic Theory* 7.4:565-604.
Moro, Andrea. 1997. *The Raising of Predicates: Predicative noun phrases and the theory of clause structure.* Cambridge, Cambridge University Press.
Pesetsky, David. 1987. "Wh in Situ: Movement and Unselective Binding". In E. Reuland and A. Ter Meulen (eds.) *The Representation of (In)definiteness*. 98-129. Cambridge MA, MIT Press.
Reinhart, Tanya. 1995. "Interface Strategies", Ms., University of Utrecht.
Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht, Foris Publications.

Rizzi, Luigi. 1986. "On the Status of Subject Clitics in Romance". In O. Jaeggli and C. Silva-Corvalán (eds.) *Studies in Romance Linguistics*. 391-420. Dordrecht, Foris Publications.

Rizzi, Luigi. 1991. "Residual Verb Second and the Wh Criterion". *Geneva Working Papers on Formal and Computational Linguistics* (republished in Rizzi 2000).

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In L. Haegeman (ed.) *Elements of Grammar.* 281-337. Dordrecht, Kluwer (republished in Rizzi 2000).

Rizzi, Luigi. 2000. *Comparative Syntax and Language Acquisition.* London, Routledge.

Rizzi, Luigi. 2003. "On the Form of Chains: Criterial Positions and ECP Effects". To appear in L. Cheng and N. Corver (eds.) *On Wh Movement.* Cambridge MA, MIT Press.

Rizzi, Luigi. 2004. "Locality and Left Periphery". In A. Belletti (ed.) *Structures and Beyond - The Cartography of Syntactic Structures.* 223-251. New York/Oxford, Oxford University Press.

Rizzi, Luigi (ed.). 2004. *The Structure of CP and IP – The Cartography of Syntactic Structures.* New York/Oxford, Oxford University Press.

Rizzi, Luigi and Ur Shlonsky. In prep. "Strategies of Subject Extraction". Ms., University of Siena - University of Geneva.

Shlonsky, Ur. 2000. "Subject positions and copular constructions". In H. Bennis, M. Everaert and E. Reuland (eds.) *Interface Strategies.*325-347. Amsterdam, Royal Netherlands Academy of Arts and Sciences.

Sigurðsson, Hallðor. 2000. "The Locus of Case and Agreement". *Working Papers in Scandinavian Syntax* 65:65-108.

Starke, Michal. 2001. *Merge Dissolves into Move*, Ph.D. dissertation, University of Geneva.

The syntax of argument and modifier genitives

Patricia Cabredo Hofherr and Carmen Dobrovie-Sorin
UMR 7023/Paris 8 and UMR 7110 /Paris 7

1. Introduction

The analysis of Saxon genitives poses interesting questions concerning both their syntax and the semantic composition underlying their interpretation.

Syntactically, a salient property of prenominal Saxon genitives (and of other genitive constructions cross-linguistically) is their incompatibility with a lexical determiner. Following one of the proposals put forward by Abney (1987) and in line with the generalised DP hypothesis, much subsequent work has posited a null determiner in noun phrases with a prenominal Saxon genitive (*John's house*). In what follows we will argue in favour of a Bare Phrase Structure analysis that dispenses with a null determiner.

Concerning the semantic interpretation of Saxon genitives two issues arise: (i) how is the semantic composition obtained for prenominal Saxon genitives and (ii) what relation is there between the prenominal, postnominal and predicative uses of the Saxon genitive (*John's house, a house of John's* and *this house is John's*).

In order to unify the semantic analysis of prenominal, postnominal and predicative Saxon genitives a recent proposal by Partee and Borschev (2003) derives the semantics of the prenominal Saxon genitive from the semantics of the postnominal genitive. Partee and Borschev further argue that prenominal Saxon genitives (*Mary's book*) can be assigned two distinct semantic analyses, being either arguments (type e) or modifiers (type <e,t>).

We will argue that prenominal and postnominal Saxon genitives have to receive distinct semantic and syntactic analyses despite the common

morphological marking since there are substantial semantic and syntactic differences between the two constructions.

At the same time we reject the modifier/argument ambiguity for prenominal Saxon genitives proposed by Partee and Borschev (2003). More specifically, we will defend an analysis that treats prenominal Saxon genitives uniformly as semantic arguments. We further show that in the nominal domain the syntax-semantics mapping for syntactic arguments is transparent: syntactic arguments are mapped onto semantic arguments. In contrast, no direct correspondence exists between *thematic* arguments (pertaining to the level of syntacticized argument structure related to the Lexicon) and *syntactic* arguments. Under this analysis, the mapping between thematic arguments and semantic arguments is always mediated by the syntax.

The central hypothesis of the present analysis is that prenominal Saxon genitives (and more generally genitives occupying the Specifier position of maximal nominal projections) are interpreted by a rule of semantic composition where the head noun denotes a function from individuals to individuals (type <e,e>). Given this rule, we can derive the syntactic constraint that the genitive DP in Spec cannot co-occur with a projecting determiner on the head noun.

2. Thematic, syntactic and semantic arguments

For the analysis of prenominal Saxon genitives, the different notions of "argument" corresponding to the levels of argument structure, syntax and semantic representation have to be clearly distinguished.

Thematic arguments are given by the positions in the theta-grid associated with a lexical item. Syntactic arguments in the noun phrase are the maximal nominal projections in A-positions, where we take the A-positions to be the specifier and the complement positions. Semantic arguments are expressions of type <e> (individuals) or <et,t> (generalised quantifiers).

(1) a. **thematic** arguments: positions in the theta-grid associated with the lexical item
b. **syntactic** arguments in the noun phrase: N_{max} in A-position (specifier and complement of the noun phrase)
c. **semantic** arguments: of type <e> or <et,t>.

In Saxon genitives the thematic arguments of the head noun need not coincide with the semantic arguments. Following Barker (1991, 1995) and Jensen and Vikner (1994) we assume that Saxon genitives require the head

noun to be relational (for a different implementation of the same intuition see Partee 1983/1995).

Saxon genitives coerce lexically object-denoting head nouns such as *bike* into relational nouns ($bike_{rel}$), denoting the set of pairs of individuals (x,y) such that y is a bike that entertains an underspecified relation R_{gen} with x:

(2) a. $[[bike_{rel}]] = \lambda x \lambda y\ [R_{gen}\ (x,y)$ and bike $(y)]$
object-denoting noun type-shifted to a relational noun
b. $[[sister]] = \lambda x \lambda y\ [sister\ (x,y)]$
relational noun: no type-shift

The Saxon genitive is interpreted as an argument of a relation, regardless of whether the relation is underlying in the lexical representation of the head noun or obtained by coercion. In (3a), the denotation **m** of the expression *Mary* is one of the *semantic* arguments of the coerced relation $bike_{rel}$, although Mary is not the *thematic* argument of the object denoting head noun *bike*. In (3b), **m** is a semantic as well as thematic argument of the relational head noun *sister*.

(3) a. *Mary's bike* = $\lambda x \lambda y\ [R_{gen}\ (x,y)$ and bike $(y)]$ (**m**)
b. *Mary's sister* = $\lambda x \lambda y\ [sister\ (x,y)]$ (**m**)

The preceding examples show that the semantic arguments may be coerced by the syntactic context and need not correspond to thematic arguments of the head noun. The syntactic representation may not only add a semantic argument by coercion, it may also limit the projection of the thematic arguments since the mapping rules between argument-structure and syntax depend on the syntax of the NP, which differs across languages. Compare English, which has an N-complement position, with Romanian where the N' level is not projected (see Ghomeishi 1997 for Persian): unlike English, Romanian cannot realise two arguments of a noun simultaneously by a syntactic NP argument of the noun.

(4) a. his translation of the Iliad
b. the enemy's destruction of the city

(5) a. *traducerea lui a Iliadei (Romanian)
b. *distrugerea orasului (a) dusmanului

The thematic properties of the head noun may constrain the syntactic representation (in particular the projection of complements) but do not

constrain the semantic representation directly. English has thematic restrictions on the N-complement position, which is realised as a postnominal *of-DP*-constituent (Barker 1995). Unlike the Saxon genitive, the *of-DP* in N-complement position does not trigger the type shift of an object-denoting head noun to a relational noun; the N-complement position can only be filled by lexical arguments, as shown by the contrast in (6):

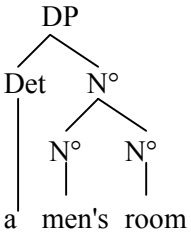
(6) a. a son of John — argument of a relational noun
b. the presentation of John — theme of a nominalization
c. a picture of John — theme of a picture noun
d.*a book of John — * not argument-taking noun

3. Differences between prenominal and postnominal Saxon genitives

As already mentioned above, the analysis proposed by Partee and Borschev (2003) unifies the semantic analysis of prenominal and postnominal Saxon genitives (*John's house* vs *a house of John's*). We argue that postnominal Saxon genitives differ syntactically and semantically from prenominal Saxon genitives and therefore have to be treated separately.

In what follows we will use the term *prenominal Saxon genitives* to refer exclusively to maximal nominal projections occupying the specifier position of the maximal projection of the head noun (7a). This construction has to be kept apart from modificational 's-genitives illustrated in (7b) for which the prenominal 's-marked constituents are not maximal N projections. Modificational genitives are analysed as modifiers that adjoin to N°-constituents (Woisetschlaeger 1983, Quirk et al 1985, Munn 1995).

(7) a. John's *big* room
b. a *big* men's room
c.



Prenominal Saxon genitives differ from postnominal *of DP's*-constituents in several respects.

First, prenominal Saxon genitives are incompatible with any determiner on the head N while *of DP's* may co-occur with a lexical determiner on the head noun.

(8) a. John's mother
b. my sister's house
c. a neighbour's daughter

(9) a. *John's **the/this/a** house
b. *the neighbour's/the neighbours' **some** houses

(10) a. **this/a** house **of John's**
b. **some** houses **of** the neighbour's

Secondly, (in)definiteness spread only occurs with prenominal Saxon genitives: the overall N^{max} is interpreted as definite or indefinite depending on whether the Saxon genitive itself is definite or indefinite (Jackendoff 1974) this is not true for postnominal Saxon genitives (see (12)):

(11) a. There is a man/*the man/??John in the garden.
b. There is a man's dog/*the man's dog/??John's dog in the garden.

(12) There is a student of John's waiting in the hall.

Finally, the genitive relation with post-nominal *of DP's* is semantically much more restricted than with prenominal Saxon genitives (Barker 1995, Storto 2000). Saxon genitives allow "free interpretations" while postnominal of-DP's do not (compare (13a) and (13b)):

(13) Yesterday John and Paul were attacked by (different) groups of dogs.
a. ... unfortunately, John's dogs were rabid.
b. # unfortunately, **some/two/many** dogs of John's were rabid.

Prenominal Saxon genitives can freely express any kind of contextually determined relation (such as *attack* in (13)) postnominal of-DP's-constituents are limited to possessive/control relations such as *own, be entrusted, take care of*, (see (14)):

(14) Yesterday, John and Paul were entrusted (different) groups of dogs.
… unfortunately, **some/two/many** dogs of John's were rabid.

We conclude from the differences between prenominal and postnominal Saxon genitives reviewed here that the two constructions should be analysed separately.

4. From DPs to NPs: The Syntax of N-projections containing a Specifier

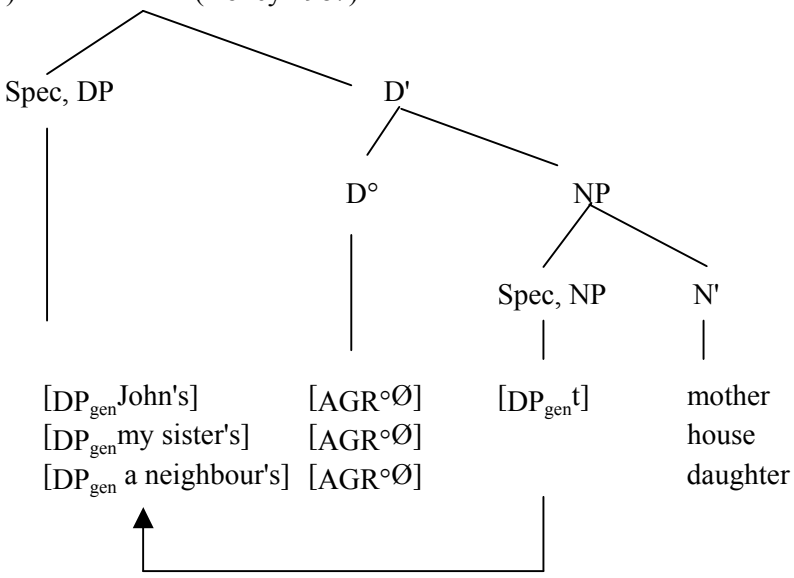
In what follows, we will show that genitives in specifier position are incompatible with a syntactically projecting determiner. We reject the hypothesis of a null determiner with prenominal Saxon genitives and propose a syntactic analysis for Saxon genitives that extends to Hebrew Construct State Nominals and Romanian synthetic genitives.

4.1 *Against null D°*

The current analyses agree that prenominal Saxon genitives of category DP are uniformly syntactic arguments: they always occupy the Spec position. The central problem of any analysis is to account for the complementary distribution of prenominal Saxon genitives and overt determiners (see (9) above).

Following proposals by Abney (1987) and Szabolcsi (1983), prenominal Saxon genitives have been analyzed as occupying the Spec, DP position, with the D° position filled by a null determiner.

(15) DP (Abney 1987)



The null determiner is postulated for two main reasons. First, the determiner is needed to conform to the DP-hypothesis postulating that maximal nominal projections are uniformly DPs. Secondly, the null determiner has been used to explain the uniqueness effect with prenominal Saxon genitives illustrated in (16) (Barker 1991, Partee 1983/1997, Jensen and Vikner 1994).

(16) I saw John's house = I saw **the** house that John owns /built etc
not: I saw **a/some** house that John owns /built etc

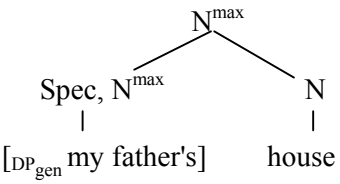
However, assuming a semantically contentful null determiner in order to explain the uniqueness effect leads to a paradox: the null D° cannot be analyzed as occupying the D° position of the overall DP, since the overall DP is not necessarily definite as we have seen in example (11) above.

The hypothesis of a (semantically empty) purely syntactic null D° is equally problematic. The hypothetical null D° is not attested independently and is restricted to co-occur with a Saxon genitive. Furthermore, there is no principled reason that would force the null D° to occur with the prenominal Saxon genitives while all lexical determiners are excluded in the same environment.

Following Dobrovie-Sorin (2000a), we therefore reject the hypothesis of a null D° and propose a Bare Phrase Structure account of prenominal Saxon genitives. Since Bare Phrase Structure does not require D° to be projected, we will assume that the head N associated with a prenominal Saxon genitive has no D°. Since noun phrases containing prenominal Saxon genitives may occupy any syntactic argument position, the overall projection has to be a maximal projection of N (N^{max}).

Prenominal Saxon genitives occupy the Spec position of N^{max}, the maximal projection of N itself.

(17)



In Bare Phrase Structure (Chomsky 1994) Spec, N^{max} is defined contextually as the XP sister to a non-maximal N-projection that is immediately dominated by N^{max}.

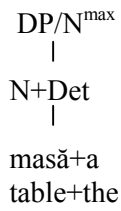
Under the present assumptions extended maximal projections of N are not necessarily DPs but may be either DP, NP or bare N (subject to interpretability constraints).

4.2 *Romanian synthetic genitives*

Dobrovie-Sorin (1987), Grosu (1988, 1994) and Cornilescu (1993), among many others, assumed a rule of N-to-D, which was suggested by the suffixal status of the Romanian definite article. Note however that the analysis relying on N-to-D was related to the GB model, which required N's and D's to be generated in distinct positions, even if the determiner was an affix.

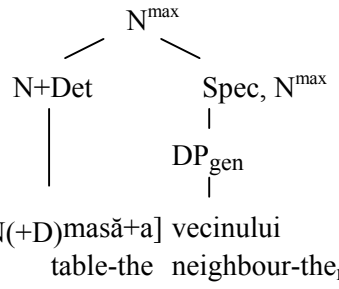
In a Bare Phrase Structure approach the noun with the suffixal definite determiner can be directly merged (as an inflected word) under a single syntactic node N+D, the maximal projection of which can be DP or N^{max}, depending on whether D° or N° projects:

(18)



N-to-D movement can be dispensed with and synthetic genitives target the Spec, N^{max} position. We thus obtain a syntactic representation that is comparable to that of noun phrases containing prenominal Saxon genitives, linear word order set aside:

(19)



4.3 *Crosslinguistic generalization*

The syntactic analysis in terms of Bare Phrase Structure allows a cross-linguistic characterisation concerning languages that allow two types of genitives.

(20)	**synthetic genitives** (no/restricted determiner)	**analytic genitives** (with determiner on the head N)
English	Saxon genitives	*of*-genitives
Modern Hebrew	Construct State associates	*Sel*-genitives
Romanian	*al*-less genitives	*al*-genitives

Longobardi (1996) proposes that the defining characteristic of synthetic genitives (which he groups under the heading of *Construct State nominals*) is the presence of N-to-D raising. Within the analysis proposed here, synthetic genitives can be characterised by their syntactic position:

(21) In languages with alternating genitives, synthetic DP-genitives occupy **Spec, N^{max}**.

We will show below that projecting Spec, N^{max} forces the determiner of the head noun to be either absent (as in English and Hebrew) or filled with a suffixal definite article (as in Romanian).

5. The semantic composition of genitive specifiers and the constraint on determiners

Given the syntactic analysis proposed above, we now need to understand the constraint imposed on the determiner of the head noun. We have to account for two generalisations. The positive generalization is that a nominal projection embedding Spec, N^{max} can have the semantic type of arguments (type <<e,t>,t> or type e) although it is not governed by a determiner. The answer to this puzzle will be that genitives in Spec, N^{max} are semantically analyzed as arguments of a function from individuals to individuals (type <e,e>), which returns the individual denoted by the overall possessive expression. The negative generalization is that Determiners (other than a suffixal definite article) are incompatible with the projection of Spec, DP. This generalization follows from the semantic type of canonical Determiners.

5.1 *The relational analysis*

Before we present the semantic analysis proposed in Dobrovie-Sorin (2000a, 2002) we will briefly review the current semantic analysis.

As already mentioned above, the analyses proposed by Barker (1991, 1995), Partee (1983 /1997) and Jensen and Vikner (1994) all assume that the semantic analysis of Saxon genitives relies on a relation, which is either

provided by the lexical properties of the head N (see relational nouns such as *son, middle*, etc.) or contributed by the structure itself. We repeat here the denotations assumed for the nouns *bike* and *sister* in the context of a Saxon genitive:

(22) a. [[bike]] $=_{\cdot}$ λxλy [R_{gen} (x,y) & bike (y)]
b. [[sister]] $=_{\cdot}$ λxλy [sister (x,y)]

For possessive expressions of the type *John's bike* or *John's sister* the denotations in (22a-b) are applied to the individual denoted by *John*, notated **j**:

(23) a. [[John's bike]] $=_{\cdot}$ λxλy [R_{gen} (x,y) $\wedge_{\cdot}$ bike (y)] (**j**)
$=_{\cdot}$ λy [R_{gen} (**j**,y) ∧ bike (y)]
b. [[John's sister]] $=_{\cdot}$ λxλy [sister (x,y)] (**j**)
$=_{\cdot}$ λy [sister (j,y)]

The expressions in (23a-b) denote sets of individuals (type <e,t>): the set of individuals that are bikes and entertain an underspecified relation R_{gen} with John, and the set of individuals that entertain the sister-relation with John, respectively. Because these expressions are of type <e,t>, they must combine with determiners in order to yield a generalized-quantifier type of denotation. Since no determiner is overtly realized, the existence of a null element must be postulated. As we have seen above, the postulation of a null determiner poses substantial empirical and theoretical problems. Note furthermore that the relational analysis predicts that synthetic genitives should be compatible with any kind of determiner. In order to avoid this overgeneration, Partee (1983, 1997) and Barker (1991, 1995) are forced to stipulate a further constraint: the null determiner governing the head noun can only be a definite article (for Partee), or a possessive article (for Barker).

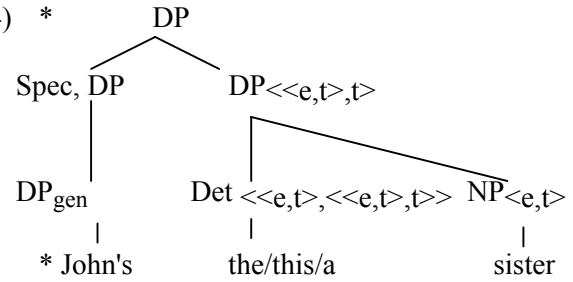
Vikner and Jensen (2002) differ from the other relational analyses in that they link uniqueness to the 's-morpheme. This is problematic, as they note themselves, since uniqueness is limited to the prenominal Saxon genitives and does not apply to *of-DP's*.

5.2 *Determiners block the projection of a Spec, DP position*

The fact that canonical determiners (i.e., determiners that apply to properties and yield generalized quantifiers or individuals) block the projection of Spec,DP follows from their semantic type: by applying Det° (type <<e,t>, <<e,t>,t>) to NP (type <e,t>) we obtain a constituent that already has the

semantic type of an argument (type <<e,t>,t>), and as such can further expand only via adjunction. But prenominal Saxon genitives, Construct State Nominals and Romanian synthetic genitives cannot be adjuncts.

(24) *

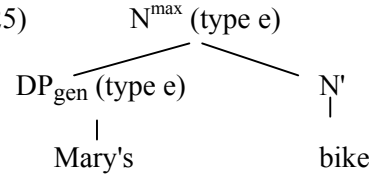


When the head N is governed by a lexical Determiner, the genitive can be (i) a N°-modifier (ii) a N°-complement, or (iii) a DP-adjunct.

5.3 *The Semantic Composition of N^{max} constituents with a filled Spec, N^{max}*

Let us now consider the semantic compostion for maximal nominal projections containing a specifier. Combining the syntactic representation above and the semantic types of the nodes that are known we obtain the following picture:

(25)



Dobrovie-Sorin (2000a,b) proposes the rule of semantic composition in (26), which applies to maximal nominal projections that contain a Spec constituent:

(26) If the Specifier of N^{max} is filled, the head N denotes a function f of type <e,e>, which applies to the individual denoted by the DP in Spec, N^{max} and yields the individual denoted by the overall N^{max}.

According to (26), the noun *bike* in (25) is interpreted as a function $\mathbf{f_{bike}}$ of type <e,e>. The interpretation of the head noun as a function in (26) is triggered by the presence of a Specifier of N^{max}, and is independent of the lexical properties of the head noun:

(27) a. f_{mother} = λx ιy [mother (x,y)] (functional noun)
b. f_{sister} (x) = λx ιy [sister (x,y)] (relational noun)
c. f_{bike} (x) = λx ιy [R_{gen}(x,y) ∧ bike (y)] (object-denoting noun)

By applying the function to the individual denoted by the genitive specifier we obtain the (contextually) unique individual that stands in the relevant relation with it:

(28) a. [[Mary's mother]] = λx ιy [mother (x,y)] (m)
= ιy [mother (m,y)]
b. [[Mary's sister]] = λx ιy [sister (x,y)] (m)
= ιy [sister (m,y)]
c. [[Mary's bike]] = λx ιy [R_{gen}(x,y) ∧ bike (y)] (m)
= ιy [R_{gen}(m,y) ∧ bike (y)]

When applied to plural head nouns, the value of the function is a plural individual (group) notated by a capital letter below:

(29) [[Mary's houses]] = λx ιY [R_{gen} (x,Y) ∧ houses (Y)] (m)
= ιY [R_{gen} (m,Y) ∧ houses (Y)]

The overall N^{max} denotes the maximal group of houses that is associated to Mary by the underspecified R_{gen}.

For the interpretation of QPs occurring in the Spec, N^{max} position we assume Quantifier Raising (see e.g. Barker 1995):

(30) a. Each student's car will be checked.
b. each x (x is a student) ιy (car (y) ∧R_{gen} (x,y)) [y will be checked]

Relational analyses like Vikner and Jensen (2002) assume that in the context of a Saxon genitive, the denotation of object-denoting nouns is shifted to a relational noun relational (type <e ,<e,t>>).

Under the present analysis, in the context of a genitive specifier, the head N is shifted to a function (type <e,e>). By applying this type of function to the individual denoted by the genitive specifier we obtain the individual denoted by the overall N^{max} constituent. No determiner is needed for the semantic composition, which correlates with the syntactic lack of overt determiners in possessives built with Saxon genitives. Under relational analyses like Vikner and Jensen (2002) where it is assumed that in the context of a Saxon genitive, the denotation of object-denoting nouns is

shifted to a relational noun relational (type <e ,<e,t>>), a null definite article must be postulated.

5.4 *(In)definiteness Spread*

As we mentioned above, the overall N^{max} is interpreted as definite or indefinite depending on whether the Saxon genitive itself is definite or indefinite (Jackendoff 1974).

Under the present analysis this (in)definiteness spread phenomenon is a consequence of the functional analysis:

(31) [[Mary's house]] $= \lambda x \ \iota y$ [house (y) $\wedge$ R_{gen} (x,y)] (m)
$= \iota y$ [house (y) $\wedge$ R_{gen} (m,y)]

In (31), the *house-of* function applies to a constant individual (denoted by *Mary*), and therefore yields a constant individual, hence the definite interpretation.

(32) [[a neighbour's house]]
$= \lambda x \ \iota y$ [house (y) $\wedge$ R_{gen} (x,y)] ([[a neighbour]])
$= \iota y$ [house (y) $\wedge$ R_{gen} (z,y) $\wedge$ neighbour(z)]

In (32) the *house-of* function applies to the free variable *z* contributed by the indefinite genitive NP, *a neighbour*. The overall NP has an "indefinite" meaning, because the value of the *house-of* function co-varies with the values of the free variable z introduced by the indefinite to which it applies.

5.5 *Definite possessives*

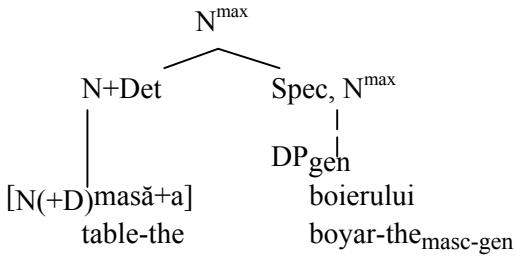
We now have to address the facts of Romanian that show that the functional analysis triggered by Spec, N^{max} is compatible (under certain conditions) with a definite article on the head N.

We propose that in this case, the definite article only indicates that the head noun denotes a function of type <e,e> (Löbner 1985): it does not have the semantic type of canonical determiners and it does not project (or the overall DP would have a definite interpretation). The definite article is the only determiner that is semantically compatible with a genitive specifier since it need not project in the syntax.

Romanian has a transparent syntax: the definite article is suffixed on the head noun and the noun+det complex forms a single morphological word that

can be merged in a single position (exactly as required by the semantic composition of genitive specifiers).

(33)

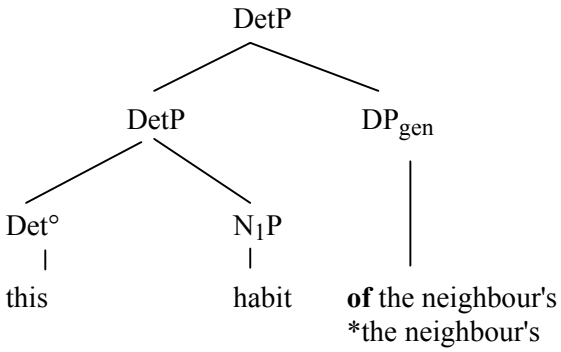


The semantic analysis proposed here explains why a (suffixal) definite article is *possible* with genitive specifiers, but not why it is *obligatory* in Romanian. One possible hypothesis is that a determiner is needed in order for morphological case to be realized, since morphological case cannot appear on bare nouns in Romanian. However, the only determiner that compatible with the semantic composition required by genitive specifiers is the suffixal definite article.

6. Genitive modifiers of category DP

Genitive modifiers of category DP such as *of-DP's*-constituents occupy DP-adjunct positions (see (34)). These adjunct genitives are interpreted as arguments of a *relation* introduced by the genitive marking; in this they differ from genitive specifiers, which introduce a *function*. The *of-DP's*-constituent in an example like *a house of John's* has the semantic type of predicates (type <e,t>) and is interpreted like a relative clause corresponding to *that is John's*. This analysis is supported by the fact that both Saxon genitives in predicate position and *of-DP's*-constituents do not allow contextual readings that are available for pronominal Saxon genitives (see discussion above).

(34)



The predicate corresponding to the *of-DP's*-constituent is obtained by abstracting over one of the argument positions of the (coerced) relation introduced by the genitive marking as in (35):

(35) [[of the neighbour's]] $=_{\lambda}$ λy [R_{gen} (the neighbour, *y*)]

We assume that the semantic composition of *of-DP's*-genitives is comparable to the semantic composition of relative clauses. As the semantic composition of relative clauses is controversial, however, we will not further investigate the exact semantic composition for *of-DP's*-genitives here.

7. Conclusion

The constraint that genitive specifiers impose on determiners is due to the rule by which they are semantically composed: the syntactic position they occupy (Spec, N^{max}) forces the interpretation of the head noun as a function. The lack of determiners and a suffixal definite article are the only two legitimate possibilities. The choice of one or the other option cannot be explained by the semantic analysis. It depends on the syntactic and morphological properties of a given language, in particular (i) the type of determiner (free morpheme, clitic or affix), (ii) the position (post- or pre-nominal) of the genitive specifier; (iii) constraints on morphological Case.

References

Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. Ph.D. dissertation, MIT.

Barker, Christopher. 1991. *Possessive Descriptions*. Ph.D. dissertation, University of California at Santa Cruz (published 1995 by CSLI Publications).

Chomsky, Noam. 1994. *Bare Phrase Structure*. Occasional Papers in Linguistics 5, MIT.

Cornilescu, Alexandra. 1993. "Notes on the Structure of Romanian DP and the Assignment of the Genitive Case". *University of Venice Working Papers in Linguistics* 3.2.

Dobrovie-Sorin, Carmen. 1987. "A propos de la structure du groupe nominal en roumain". *Rivista di grammatica generativa* 12:123-152.

Dobrovie-Sorin, Carmen. 2000a. "(In)definiteness Spread: from Romanian Genitives to Hebrew Construct State Nominals". In V. Motapanyane (ed.) *Comparative Studies in Romanian Syntax*. Amsterdam/Philadelphia, John Benjamins.

Dobrovie-Sorin, Carmen. 2000b. “De la syntaxe à l'interprétation, de Milner (1982) à Milner (1995): le génitif”. In G. Lardreau (ed.) *Cahier Jean Claude Milner*. Paris, Verdier.

Dobrovie-Sorin, Carmen. 2002. “From DPs to NPs: A Bare Phrase Structure Account of Genitives”. In M. Coene and Y. D'Hulst (eds.) *From NP to DP. Volume 2: The expression of possession in noun phrases*. Amsterdam/Philadelphia, John Benjamins.

Ghomeshi, Jila. 1997. “Non-projecting nouns and the EZAFE Construction in Persian”. *Natural Language and Linguistic Theory* 15:729-788.

Grosu, Alexander. 1988. “On the Distribution of Genitive Phrases in Romanian”. *Linguistics* 26:931-949.

Grosu, Alexander. 1994. *Three Studies in Locality and Case*. London and New York, Routledge.

Jackendoff, Ray. 1974. *An Introduction to the X-bar Convention*. Ms., Bloomington, Indiana, Indiana University Linguistics Club.

Jensen, Per Anker and Carl Vikner. 1994. “Lexical knowledge and the semantic analysis of Danish genitive constructions”. In S.L. Hansen and H. Wegener (eds.) *Topics in Knowledge-based NLP Systems*. 37-55. Copenhagen, Samfundslitteratur.

Löbner, Sebastian. 1985. “Definites”. *Journal of Semantics* 4:279-326.

Longobardi, Giuseppe. 1996. "The Syntax of N-raising: a minimalist theory", Ms., University of Venice.

Munn, Alan. 1998. “The Possessor that Stayed Close to Home”, In V. Samiian *(ed.) Proceedings of WECOL* 24.181-195.

Partee, Barbara. 1983/1997. “Uniformity vs versatility: The genitive, a case study.” Appendix to Th. Janssen 1997. “Compositionality”. In J. van Benthem and A. ter Meulen (eds.) *The handbook of logic and language*. 464-470. New York, Elsevier.

Partee, Barbara and Borschev, Vladimir. 2003. “Genitives, relational nouns and argument-modifier ambiguity”. In E. Lang, C. Maienborn and C. Fabricius-Hansen (eds.) *Modifying adjuncts (Interface Explorations 4)*, 67-112. Berlin, Mouton de Gruyter.

Quirk, Randolph, Sidney Greenbaum, Geoffry Leach and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London, Longman.

Storto, Gianluca. 2000. “On the structure of indefinite possessives”. *Proceedings of SALT X*.

Szabolcsi, Anna. 1983. “The Possessor that Ran away from Home”. *The Linguistic Review* 3:69-102.

Vikner, Carl and Per Anker Jensen. 2002. “A semantic analysis of the English genitive. Interaction of lexical and formal semantics”. *Studia Linguistica* 56:191-226.

Woisetschlaeger, Erich. 1983. “On the question of definiteness in ‘an old man's book’”. *Linguistic Inquiry* 14:137-154.

An anti-locality approach to Greek demonstratives

Kleanthes K. Grohmann and Phoevos Panagiotidis
University of Cyprus - Cyprus College

1. Introduction

Many languages show a great deal of variation in their nominal system regarding ordering relations among nouns, adjectives, determiners, and so on. We focus on the Modern Greek (henceforth, Greek) DP and examine in particular the different positions demonstratives can occupy. The basic contrast under investigation will be the one exemplified by the a- and b-examples in (1) through (3), where the demonstrative *aft-* 'this/these' may occupy a DP-initial position (preceding everything else) or a DP-medial one (between adjective and noun), in each case "doubled" by an apparent determiner in the guise of the definite article (glossed as *ART*):

(1) a. afta ta nea fenomena
b. ta nea afta fenomena
ART.NOM.PL.NEUT *new.NOM.PL.NEUT* *this.NOM.PL.NEUT* *phenome-non.NOM.PL.NEUT*
'these new phenomena'

(2) a. afti i orea gineka
b. i orea afti gineka
ART.NOM.SG.FEM *pretty.NOM.SG.FEM* *this.NOM.SG.FEM* *woman.NOM.SG.FEM*
'this pretty woman'

(3) a. aftos o kalos andras
b. o kalos aftos andras
ART.NOM.SG.MASC *nice.NOM.SG.MASC* *this.NOM.SG.MASC* *man.NOM.SG.MASC*
'this nice man'

We can formulate the main issues of interest for our study as (I1) and (I2):

(I1) the connection between demonstratives and doubling articles
(I2) (anti-)locality restrictions on their positions and operations

In particular, we will construe a novel connection between demonstrative elements and apparent definite articles. Throughout this paper, the class of demonstrative elements is understood to include both an overt demonstrative pronoun and a covert, phonetically empty demonstrative operator (but see section 5); we will refer to the article as *doubling* the demonstrative. We believe one may establish this connection from the mechanisms and operations involved in the derivation of complex nominal expressions. Our language of investigation is Greek, but the proposal can, and possibly must, be extended to other languages — at least those that also exhibit such "doubling" strategies. In a nutshell, the novel connection concerns the derivational insertion of a determiner (the apparent article) as the result of an anti-local configuration involving the demonstrative (operator) in the course of the derivation and is couched within the *Anti-Locality Hypothesis* developed recently by Grohmann (2000, 2003). As such, this study constitutes a further application of the framework originally proposed for clausal syntax to the nominal domain (cf. Grohmann and Haegeman 2003, Ticio 2003). By so doing, this study (i) corroborates the claim that the Anti-Locality Hypothesis is a more general condition on the computational system of human language, and (ii) further supports the *Clausal DP-Hypothesis* that considers the structure and derivation of DP as the nominal equivalent of the structures and derivations found within clausal syntax (see e.g. Abney 1987, Ritter 1991, and Panagiotidis 2002).

2. Demonstrative Issues within the Greek DP

Nominal expressions in Greek have the general surface structure [$_{DP}$ D > A > N > DP$_{GEN}$]. All nominal elements agree for φ-features (gender, number) and Case (cf. (1)-(3)). We argue that the Greek DP makes heavy use of nominal positions to the left of the thematic noun-position (N). The co-occurrence of

demonstrative, whether DP-initial as in (4) or post-adjectival as in (5), and article is initially explained along the lines of Panagiotidis (2000).[1]

(4) afta ta nea fenomena [DEM > D > A > N]
these ART new phenomena
'these new phenomena'

(5) ta nea afta fenomena [D > A > DEM > N]
ART new these phenomena
'these new phenomena'

First, it has to be noted that doubling the demonstrative with the article is obligatory in Greek, as the ungrammaticality of (6a) shows. Second, demonstrative obligatorily precedes article when the two are adjacent (6b).

(6) a.*afta (nea) fenomena
these new phenomena
'these new phenomena'
b.*ta afta (nea) fenomena
ART these new phenomena
'these new phenomena'

As Panagiotidis shows, the structure in (4) is the result of demonstrative raising to the nominal left periphery in order to give the DP a deictic interpretation (7); when failing to do so, a locally generated covert operator *OP* takes its place (8), resulting in an anaphoric reading for the DP. We will generally adopt this approach (but see section 5).

(7) afta ta nea ~~afta~~ fenomena
these ART new phenomena

(8) *OP* ta nea afta fenomena
ART new these phenomena

The idea that demonstratives raise from a lower position within the DP has been proposed for Spanish by Brugè (1996) and Bernstein (1997), for

[1] For presentation purposes, we refrain from explicit morphological analysis in glosses, simply pointing out that there is agreement as described above. Throughout, the apparent definite article will be glossed as *ART*, for reasons that will become clear shortly; all other conventions are standard, unless noted otherwise.

example; see Giusti (1997) for an overview incorporating data from Romanian, (Irish) Gaelic, and (Modern) Greek. As for the base-generated position of the demonstrative *afta* 'these' (nominative/accusative plural neuter), we can note that it must certainly be within the agreement layer of the DP (see section 3.2), as (8) above illustrates: the demonstrative in base-position shows up between the noun and the agreeing adjective.

We further follow Panagiotidis' and others' arguments that the orders observed cannot result from remnant movement: *ta* 'the' and *nea* 'new' do not form one constituent, so that determiner and adjective could never move as one alone, and too many unmotivated steps would have to be assumed in order to align *ta* and *nea* on the right edge and move a remnant category (such as high N-raising beyond its thematic base-position, for which there is hardly any evidence in Greek; see also Alexiadou and Stavrou 1998).

However, the demonstrative *afta* 'these' and the determiner, the apparent definite article *ta* 'the' (an apparentness to which we return), can co-occur and may even be adjacent (in fact, adjacency is preferred, but with the demonstrative obligatorily preceding the article). For lack of a better term, we refer to this co-occurrence as *doubling* for reasons that will unfold in section 4. Our analysis will thus account for the doubling pattern observed in Greek (and possibly other languages; cf. Grohmann and Panagiotidis 2004), and it will also have something to say about the adjacency with the demonstrative.

Panagiotidis (2000) adopts the basic analysis of two demonstrative positions related through movement (high occurrence as the result of moving from a lower position), but throws up the question why the demonstrative *may* move. For starters, call the two observed occurrences of the demonstrative as exemplified by (4) and (5) DEM_{HI} and DEM_{LO}, respectively. The answer he reaches (cf. Manolessou and Panagiotidis 1999), and which we will adopt, is that DEM_{HI} receives a *deictic interpretation* (which we take to be encoded in the nominal left periphery, as one would expect if the Clausal DP-Hypothesis, presented in section 3.2 below, holds). DEM_{LO}, on the other hand, is restricted to a *discourse-anaphoric function* (where OP is hence an anaphoric demonstrative operator). (9) briefly illustrates (from Panagiotidis 2000:723):

(9) Context: A customer at the butcher's, pointing to a pork joint.

a.	# Thelo		to			butaki	afto.
b.	# Thelo		to	apaho	afto	butaki.	
c.	Thelo	afto	to	apaho		butaki.	
	I-want	*this*	*ART*	*lean*		*joint*	

'I want this lean joint.'

As these examples suggest, only the DP-initial position is genuinely deictic. The post-adjectival (or even -nominal) occurrence must pick up a reference already introduced in the discourse and cannot introduce a new topic.

Panagiotidis' analysis invokes two additional ingredients, the Demonstrative-Criterion and a demonstrative article. The Demonstrative-Criterion as formulated by Panagiotidis (exploring a suggestion from Campbell 1996) requires a Spec-head relation between two demonstrative features: a demonstrative head (determiner) must enter into a local relationship with a demonstrative specifier (operator). (The notation [+TH] is used for this feature, on analogy with [+WH] of the *Wh*-criterion, the ur-criterion.) Regarding the demonstrative article, this would constitute a third type of determiner alongside the regular definite article and what might be called the expletive use of the article (originally proposed in Roussou and Tsimpli 1994).

(10) *Demonstrative Criterion* (Panagiotidis 2000:724)
(i) A [+TH] determiner has a [+TH] specifier.
(ii) A [+TH] operator specifies a [+TH] determiner.

We will first argue against any criterion-based approach to syntactic structures (in section 3.1) and then propose an analysis which doesn't posit a new type of article (in section 4 — with Greek articles, among other issues, addressed further in section 5). To get things started, we present the Anti-Locality Hypothesis and the Clausal DP-Hypothesis as necessary background for our study.

3. Clausal and Nominal Structures

The Anti-Locality Hypothesis is an attempt to capture the intuition that licit movement must not only be restricted in terms of an upper bound, but also of a lower bound: movement must cross a minimum distance in order to be well-formed. The relevant metric for measuring distance is expressed in terms of derivational sub-components (henceforth Prolific Domains), which span information-relevant related projections (in some sense, possibly similar to "extended projections" in the sense of Grimshaw 1991, 2003). Grohmann (2000, 2003) identifies the classic tripartition of clause structure as the three Prolific Domains at the clausal level:[2] the *Thematic or Θ-*

[2] On the "classic tripartition" — i.e. [CP [IP [VP]]] of Chomsky (1986) — see e.g. Larson (1988) on VP-shells, Pollock (1989) on split Infl, Rizzi (1997) on split Comp, and Platzack (2001) on the three-tiered clause roughly along the lines pursued here. In addition, the recent

Domain (basically VP-shells: *v*P and VP), the *Agreement- or Φ-Domain* (split Infl: hosting IP/TP, and others), and the *Discourse- or Ω-Domain* (split Comp: topics, foci, operators, etc.)

Since Abney (1987), attempts have been made to formulate the Clausal DP-Hypothesis, exploring the observation that D seems to largely mimic the role of C in the nominal layer (see also Szabolcsi 1983, Horrocks and Stavrou 1987, Ritter 1991, and others, for which Haegeman 2001 and Bernstein 2001 provide overviews). As such, one might expect that a tripartition in terms of Prolific Domains could be mirrored in the nominal domain; after all, the nominal system displays both thematic and agreement properties (suggesting Θ- and Φ-Domain), and if D is the nominal C, there should be also properties reflecting the Ω-Domain.

3.1. *The Anti-Locality Hypothesis*

Under the guiding minimalist desideratum that the structure of the grammar be determined by (virtual) conceptual necessity (Chomsky 1993, 1995), much of the GB-machinery should be reconsidered, in particular restrictions on the computation that are not motivated by Bare Output Conditions (see e.g. Hornstein 2001: chap. 1, Grohmann 2003: chap. 2, Hornstein, Nunes, and Grohmann, in press: chap. 2). We might thus ask whether the ungrammaticality of (11a-c) could receive an alternative explanation to standard approaches, which commonly invoke filters of sorts (such as the Theta Criterion, Case Filter, various Affect Criteria, etc.):

(11) a. *John likes. (cf. *John likes himself.*)
b. *Him kissed her. (cf. *He kissed her.*)
c. *Who, Mary detests? (cf. *Who does Mary detest?*)

Assume that (12a-c) are appropriate representations of the derivations corresponding to (11a-c) at the relevant points under the copy theory of movement:[3]

(12) a. #[$_{vP}$ John v^0 [$_{VP}$ likes ~~John~~]]
b. #[$_{TP}$ him T^0 [$_{AgrOP}$ ~~him~~ $AgrO^0$ [$_{vP}$ softly [$_{vP}$ ~~him~~ v^0 [$_{VP}$ …]]]]]

volumes on the "cartography of structure" provide further relevant material: Cinque (2002), Rizzi (2004), and Belletti (2004).

[3] See e.g. Chomsky (1995), Nunes (1995, 2004), Hornstein (2001), and Hornstein, Nunes, and Grohmann (in press) on formulating the copy theory. Throughout this paper, lower copies are represented in ~~strikethrough~~ and structural ill-formedness is indicated by the hash mark '#'.

c. #[$_{TopP}$ who Top^0 [$_{FocP}$ ~~who~~ Foc^0 [$_{TP}$ Mary T^0 detests … (~~who~~)]]]

We can observe that the derivations in (12) are all ill-formed, so one would need to say something else to rule them out, if we follow the path just mentioned, that restrictions on the computation which do not follow from Bare Output Conditions are not allowed. A starting point for a purely syntactic explanation for this ungrammaticality would be the hypothesis in (13), generalizing ideas from e.g. Bošković (1994) or Murasugi and Saito (1995):

(13) *Anti-Locality Hypothesis* (Grohmann 2003:26)
Movement must not be too local.

In structural terms, "too local" or *anti-local* describes a dependency between two contextually related positions. We take contextual information (as relevant for anti-locality) to be encoded in all lexical and functional heads that build up a derivation. In order to capture this intuition in structural terms, we introduce the notion of a *Prolific Domain*:

(14) *Prolific Domain* (Grohmann 2000:58)
A Prolific Domain is a contextually defined part of the computational system, which (i) provides the interfaces with the information relevant to the context and (ii) consists of internal structure, interacting with derivational operations.

Following earlier conceptions of the clause (e.g. Chomsky 1986) and much current research on the finer structure of these projections (see Cinque 1999 or the cartography-volumes cited above for review and references), a presumably natural implementation of contextual information would be a clausal tripartition, a formal split of the clause into three Prolific Domains: a Theta-, an Agreement-, and a Discourse-Domain (cf. Platzack 2001 for a related proposal). Following Grohmann (2000, 2003), we refer to these as (i) the *Θ-Domain* (that part of the derivation where thematic relations are created; "VP-shells"), (ii) the *Φ-Domain* (where agreement/inflectional properties are licensed; "split Infl"), and (iii) the *Ω-Domain* (establishing discourse information; "split Comp").

Further following Grohmann's framework, we adopt a dynamic approach to the computational system of human language in terms of cyclic Spell Out, namely one which allows the operation Spell Out to apply more than once (Uriagereka 1999, Chomsky 2000). Each Prolific Domain forms a part of the derivation where Spell Out applies and the information contained within gets

shipped to the PF- (and possibly LF-) interface component. One minimalist criterion that all conditions, operations, and principles must abide by is that they follow from Bare Output Conditions (Chomsky 1995). With the abolishment of the GB-levels of D- and S-structure, many of the standard conditions do not follow from Bare Output Conditions (cf. discussion around (11)). We can now formulate a single condition that does, the CDE:

(15) *Condition on Domain Exclusivity* (*CDE*; Grohmann 2003:78)
For a given Prolific Domain ΠΔ, an object O in the phrase-marker must receive an exclusive interpretation at the interfaces, unless duplicity of O yields a drastic effect on the output of that ΠΔ.

Further details aside, the CDE applies to all and only XP-dependencies within a Prolific Domain (but it allows head movement, as head movement changes the PF-matrix of the two heads; see Grohmann 2003:79-80, 2004).

A further prediction of the CDE is that if a dependency between two XPs within one Prolific Domain involves two different PF-matrices (the phonological shape of a linguistic expression), the dependency should be well-formed. An interesting and reasonably clear-cut instance of this is a type of left dislocation in German, often labeled "contrastive" left dislocation:

(16) a. [*Seinen*$_i$ *Vater*], *den* mag jeder$_i$ Junge.
his.ACC father RP.ACC likes every boy
'His father, every boy likes.'
b. [$_{CP}$ seinen Vater C^0 [$_{TopP}$ *den* mag-Top^0 [$_{TP}$ jeder Junge T^0…]]]

The left-dislocated XP and the resumptive pronoun RP (morphologically, a weak demonstrative pronoun), with which it agrees in Case and φ-features, are in the same Prolific Domain (Ω-Domain). Moreover, (16) allows a bound variable reading and aside from such absence of weak crossover effects, contrastive left dislocation displays other signs of reconstruction (such as presence of Condition A effects, absence of Condition C effects, or idiom chunks; see Grohmann 2003:149-152 for discussion and references).

All this and more (such as possibilities of embedding or multiple left dislocation) stands in sharp contrast to hanging topic left dislocation:

(17) a. [*Sein*$_i$ *Vater*], jeder$_{*i/k}$ Junge mag *den/ihn*.
his.NOM father every boy likes RP/him.ACC
'His father, every boy likes him.'
b. [$_{CP}$ sein Vater [$_{CP}$ C^0 [$_{TP}$ jeder Junge mag-T^0 den/ihn…]]]

Hanging topics characteristically show up in nominative Case only and the "correct" Case-marking shows up on the RP. The RP may be expressed by either the weak demonstrative (as in contrastive left dislocation) or the personal pronominal form; moreover, it may appear in the same position as in contrastive left dislocation (not shown here; cf. Grohmann 2003:144) or show up in the base-position, thus suggesting further that it is the RP which is selected by the predicate and inserted into the derivation.

The obvious analysis made possible by the Anti-Locality Hypothesis (now understood as per (13)-(15) above) is to derive contrastive left dislocation in terms of a (movement) dependency between the left-dislocated XP and the RP, while hanging topics are generated in their surface position, as in standard analyses. By the CDE, this movement can be understood as the result of Copy Spell Out ('➲'), changing the PF-matrix of the lower of the two copies that are in the same Prolific Domain (where TopP and CP are both part of the Ω-Domain):

(18) $[_{CP}$ seinen Vater C^0 $[_{TopP}$ ~~s.V.~~ ➲ den mag-Top^0 $[_{TP}$ … ~~s.V.~~ …]]]

We can understand the operation Copy Spell Out to be a repair strategy that applies at a given Prolific Domain as the result of a PF-legibility violation. At the point when a Prolific Domain is formed, PF sees two identical copies of one linguistic expression and cannot deal with them. Spelling out the lower copies provides the "drastic effect" required by the CDE.

If RPs in contrastive left dislocation can be reasonably analyzed as a derivational result, rather than fully lexical items part of the numeration/lexical array, two relevant questions arise: (i) Do we find other instances of resumption that could be analyzed as Copy Spell Out? (ii) Do we find other occurrences of pronouns that could be understood as resumption?

Given the clausal tripartition into Prolific Domains, the CDE and the operation Copy Spell Out as sketched above, one could indeed envision another set of "resumptive" elements, namely grammatical formatives inserted to legitimize a dependency whose members would otherwise be too close to be licensed. A pronoun-qua-grammatical-formatives view has recently been integrated into a derivational approach for local anaphors by Hornstein (2001). As relevant for the Anti-Locality Hypothesis, Grohmann (2003) suggests that reflexives may be employed to legitimize a too-close dependency.

To briefly illustrate with a relevant structure touched on above, take (19), where *v*P and VP form one Prolific Domain (namely, the Θ-Domain):

(19) a. John$_i$ likes himself$_i$.
b. [$_{TP}$ John T^0 [$_{vP}$ ~~John~~ v^0 [$_{VP}$ likes ~~John~~ ➲ himself]]]

If on the right track (see Grohmann's and Hornstein's works for further discussion and references), the common characterization of the distribution of RPs — that they get inserted when the distance between two positions in a dependency would otherwise be too far to be licensed legitimately (on standard "upper-bound" accounts of locality) — can be extended (e.g. Boeckx 2003a, Boeckx and Grohmann 2004a, 2004b). Note that it does not matter that *John* in (19) eventually ends up in a higher Prolific Domain (SpecTP in the Φ-Domain). Prolific Domains serve as cyclic, dynamic points at which Spell Out (to PF) applies (but see Grohmann 2005). That is, at the point where *John* has moved *v*P-internally (resulting in two copies in the Θ-Domain) and Spell Out applies to the Θ-Domain, the CDE kicks in.

We now have (at least, theoretical) reasons to believe that some instances of resumption may take place derivationally, namely in an anti-local relationship, when the distance between two positions is too close. In other words, modifying a Last Resort approach to resumption (Shlonsky 1992), one type of RP is inserted into a structure from which movement cannot take place because the distance between the two positions is too far in a sense ("standard locality"), another when the distance is too close ("anti-locality").

3.2. *The Clausal DP-Hypothesis*

Our ultimate goal in this study is to apply the Anti-Locality Hypothesis to Greek nominal structures and derive the article analogously to other spelled out grammatical formatives. Such an approach is intricately connected to a partition of the nominal layer akin to the one we have sketched for the clausal layer. Let us thus look more closely at DP-structure.

One obvious similarity between nominal and clausal constructions concerns left dislocation. The following examples illustrate the fact that left dislocation may also apply within DPs in German ((20a) from Grohmann and Haegeman 2003:51):

(20) a. [Über [Kanzler Schröder]$_i$ dem$_i$ seine$_i$ Fehler] haben wir geredet.
about Chancellor Schröder RP his mistakes have we talked
'About Chancellor Schröder's mistakes, we talked.'
b. [Peter$_i$ dem$_i$ sein$_i$ lockeres Mundwerk]
Peter RP his loose gab
'Peter's quick tongue'

Assume that the analysis presented briefly above is indeed the correct analysis for left dislocation structures (see also Boeckx and Grohmann 2004b). Then the relevance of this type of "nominal left dislocation" as shown in (20) is as follows: first, if left dislocation involves Copy Spell Out in the clausal layer, it should also do so in the nominal layer; second, if Copy Spell Out in the clausal layer is due to satisfying the CDE (viz. Prolific Domains), the nominal layer should also be sensitive to the CDE (i.e. have Prolific Domains). This intuition was first explored by Ticio (2003) and served as background assumption underlying Grohmann and Haegeman's (2003) work on prenominal possessive doubling (extended in Grohmann 2003: chap. 6), and it is the one we will be working with, thus hopefully furthering our understanding of the architecture of the nominal layer.

Ever since the formulation of Abney's (1987) DP-Hypothesis and Ritter's (1991) suggestion of e.g. an agreement-related Num(ber)P within DP, much evidence has been collected to align the nominal DP-structure to the clausal CP-structure, where D^0 plays the "nominal role" of C^0, so to speak (see e.g. Bernstein 2001 and Haegeman 2001 for recent reviews). Replacing "NumP" by a more general "AgrP" (and do the same with "TP/IP"), the following picture emerges; the parallel between the two structures is completed once we hypothesize a light noun *n* (Radford 1999, Adger 2003:266-269):

(21) a. clausal structure: CP > AgrP > *v*P
b. nominal structure: DP > AgrP > *n*P

If *v*P denotes the domain of thematic relations, AgrP of agreement properties, and C/DP of discourse information (all as understood throughout), a first approximation would thus be to assign the same Prolific Domains, as illustrated in (22):

(22) a. $CP_{\Omega\Delta} > AgrP_{\Phi\Delta} > vP_{\Theta\Delta}$
b. $DP_{\Omega\Delta} > AgrP_{\Phi\Delta} > nP_{\Theta\Delta}$

Such a tripartite composition of DP is in principle widely employed, and as such suggests that we would find the same (type of) Prolific Domains here as well, just as with the tripartite composition of CP (the clause). And just as these functional projections have been finer articulated in the clausal layer, so have they in the nominal layer (see references above).

We are not so much concerned with identifying various positions within the nominal layer (i.e. categorial labels and specific projections) as we are with the relational ordering of projections and the consequences for the syntax of demonstrative expressions. One important assumption we are

making concerns the phrase-structural status of the major players involved: following Stavrou and Horrocks (1989), we take demonstratives to be maximal projections within an articulated DP, alongside adjectives, as opposed to heading their own projection; this is shown in (23).

(23) $[_{DP}$ Spec D^0 $[_{AgrP}$ Spec Agr^0 $[_{NP}$ Spec N^0]]]

Since demonstratives (as well as adjectives) agree with the noun (as mentioned in section 2), it is reasonable to assume that their merging site lies somewhere within the Φ-Domain, signaled in (23) by AgrP. The next section will deal with a concentrated exploration of "DP" in (23). (To repeat, we ask the reader to disregard exact phrasal identification, which we provide here just for convenience, where adjectives would presumably either be adjoined to AgrP, or its relevant articulation, such as NumP, or occupy its specifier.)

Now that we can conceptually motivate Prolific Domains in the nominal layer, let us see whether we can empirically support them the same way we have done in the clause, i.e. in terms of the CDE. Our testing case is the Greek DP. One defined goal of this paper is thus a strengthening of the Anti-Locality Hypothesis by demonstrating a more general application of both tripartite structure in terms of Prolific Domains and anti-locality effects.

4. The Anti-Locality of Demonstrative Operator-Movement

Let us begin by formulating the main problems with Panagiotidis (2000):

(P1) If all Affect Criteria should be dispensed with (section 3.1), so should the Demonstrative-Criterion (independently of other shortcomings).
(P2) The purported demonstrative article can't be motivated (no morphological distinction or properties) nor does it do anything (and the demonstrative is still present).

Looking at some (un)grammatical positional variations displayed in the examples below, we further argue that the Ω-Domain in Greek is made up of (at least) three positions:

(Ω1) a topic position in (24) below preceding the demonstrative (cf. the unacceptability of (25) below);
(Ω2) the position of the demonstrative itself (DEM_{HI} in (4), possibly encoding deictic force viz. a focus projection FocP);
(Ω3) the position of the article (our DEM_{LO}-position; cf. Rizzi's (1997) Fin/lowest C-head, quite possibly DP proper).

Concentrating on the nominal left periphery for the obvious reason that demonstrative (overt or null OP) and article appear in this part of the nominal expression, (24) and (25) are two relevant structures that exemplify exactly these three positions:

(24) [tis epohis] afta ta fenomena [~~tis epohis~~]
the.GEN age.GEN these ART phenomena
'these phenomena of our times'

(25) * afta [tis epohis] ta fenomena [~~tis epohis~~]
these the.GEN age.GEN ART phenomena
'these phenomena of our times'

Our analysis runs as follows. The demonstrative, coming from the Φ-Domain (7), lands first in the 'article'-position before moving to its surface position; since the second movement is too local (within the Ω-Domain), the violating copy is spelled out in the form of the article, fully agreeing with the demonstrative in number, gender, and Case. Similarly, the empty operator *OP*, moving too locally from/through the position of the article, also leaves behind a spelled out copy, the article. (26) summarizes the relevant steps of the derivation in which the article is derivationally introduced by the rescuing strategy Copy Spell Out.

(26) a. [$_{\Omega\Delta}$ … afta … [~~afta~~ ➲ ta [$_{\Phi\Delta}$ nea [~~afta~~ [$_{\Theta\Delta}$ fenomena]]]]]
b. [$_{\Omega\Delta}$ … *OP* … [~~*OP*~~ ➲ ta [$_{\Phi\Delta}$ nea [afta [$_{\Theta\Delta}$ fenomena]]]]]

As noted in the discussion of example (9) above, the low position of the demonstrative *afta*, as in (26b), has an anaphoric reading. The relevant derivation involves an empty operator *OP* forming a chain with the demonstrative *afta*, which remains in situ. The interpretive effect of the *OP…afta*-chain is possibly part of a more general pattern for elements appearing in high and low positions, i.e. in Ω- and Φ-Domain respectively: high elements (in the Ω-Domain) receive a 'strong' interpretation, whereas low elements (in the Φ-Domain) receive a 'weak' one. In the case of demonstratives, the strong interpretation corresponds to a deictic reading, while the weak interpretation is restricted to discourse anaphoricity.

A parallel example from the nominal domain would be Turkish *bir* 'one':

(27) *Turkish* (Ayşe Gürel, p.c.)
a. bir/iki güzel kuş
one/two beautiful bird
'one / two beautiful bird(s)'
b. güzel bir kuş
beautiful a bird
'a beautiful bird'

The element *bir* has a numeral interpretation when it appears in (or possibly moves to) the Ω-Domain of the DP, the position of other, ordinary numerals, as illustrated in (27a). When it shows up in the lower (possibly base) position, within the Φ-Domain (as is the case with the Greek demonstrative in (8)), it can only receive a weak interpretation, one akin to that of English *a*, as the gloss of (27b) indicates. Without getting into proposing a detailed analysis for *bir* here (possibly a "Diesing-effect"), we would like to believe that similar examples indicate the weak interpretive option of an *OP*...X-chain, where X is an overt demonstrative in its Φ-Domain (base) position.

We are now in a position where we can possibly pinpoint further the architecture of the "DP"-part, the left-peripheral nominal Ω-Domain. The existence of a Fin-position in the DP (cf. (Ω3) above) is motivated by well-known arguments regarding the need to anchor the constituent in discourse, whether this be 'referential', 'anaphoric' or a variable.[4] Hence, SpecFinP, the position occupied by the Copy Spell Out of the demonstrative, is no other than the Determiner-position "D" as we know it. The Focus-position above it also provides information crucial to the interpretation of the demonstrative, hence its (anti-local) overt movement to SpecFocP in Greek. This is so because of the focal character of deixis, which foregrounds information by associating it with a point in (conceptual) space; this of course includes time.

If these considerations are on the right track, we have considerable support to identify the two left-peripheral positions suggested in our analysis above as the specifiers of FinP and FocP respectively. They also strongly suggest that the nominal left periphery makes available two operator-positions, a low operator (SpecFinP) and a high operator (SpecFocP).

[4] See Longobardi (1994) for some such arguments. Other relevant references include Uriagereka (1996), who investigates the syntax and semantics of possessive constructions, and Castillo (2001), a cross-linguistic study devoted to the grammar of content-container relations and other issues in thematic relations and displacement phenomena within the nominal layer.

5. A Note on Consequences and Cross-Linguistic Extensions

With the proposal out on the table, we want to zoom in on three points of interest: (i) the nature of the apparent article in Greek demonstrative doubling and (ii) an ontology of movement dependencies and the application of Copy Spell Out (including a brief discussion on empty or null operators).

Ad (i): We mentioned in section 2 above that Greek has been argued to make available two types of determiners, the 'regular' definite article and an occurrence that Roussou and Tsimpli (1994) dubbed 'expletive' article (though we hasten to add that this is not the focus of their discussion). Panagiotidis (2000) adds to these two types of articles a third variant, the 'demonstrative' article. Giannakidou and Stavrou (1999) argue in detail for an intensionality-operator analysis of these expletive articles, which we tentatively adopt. At this point, we are not so much concerned about the inventory of the species "article" as we are with the more general issue at hand. Hence we only wish to point out that demonstratives do not combine with the expletive version of the article, wherever there is a morphological distinction thereof from the definite one. That much is made clear by Panagiotidis himself (Table 1 is adapted from Panagiotidis 2000:731):

Table 1: Types of Articles

	definite article	ι-operator/expletive	demonstrative article
N.Greek	*u skilus* 'the dog'	*i Yans* '(the) Yannis'	*aftos u / *i skilus* 'this dog'
Catalan	*el gos* 'the dog'	*en Joan* '(the) Joan'	*el / *en gos aquest* 'this dog'

Nevertheless, unlike Panagiotidis' original proposal, we don't need to assume a third type of article as the demonstrative's companion: it is neither a definite (second column in Table 1) nor an expletive article (or ι-operator, third column), nor is it a demonstrative article (fourth column) — it is a grammatical formative: the result of Copy Spell Out of the demonstrative, specified exactly for the relevant φ-features. As far as grammatical formatives go, homophony with the article seems to be a perfect candidate: it's a minimally pronounceable form with just the (Case and φ-feature) agreement markings needed.

To make our proposal explicit, we are indeed suggesting that the article doubling the demonstrative is not an independently merged expression present in the numeration, but a purely grammatical formative, inserted into the structure for PF-reasons (CDE). The article is the spelled out copy of an anti-locally moved demonstrative — and we mark it as such in all our

examples by glossing it as *ART* throughout.

Ad (ii): We are now faced with an important outcome, if our proposal (demonstrative doubling qua Copy Spell Out) and the framework in which it is framed (Anti-Locality Hypothesis) are of any interest: *any* syntactic object underlies the PF-condition of the CDE, even phonetically null material. (This was already observed in Chomsky and Lasnik 1978 on the presence of traces and the role for *wanna*-contraction, a comparison we will not develop here any further.) Recall that simple trace/copy-deletion does not suffice to satisfy the "drastic effect on the output" required by PF. This becomes very apparent if we look at displaced empty elements, such as null operators. The postulation of the CDE that an anti-local dependency may be "rescued" by changing the PF-matrix of the lower copy (or member of a chain), on the other hand, now receives further strengthening, since this is something that can be observed in dependencies involving empty elements.

This reasoning yields the following ontology of movement dependencies:

(28) a. [XP … [… ~~XP~~ …]]
b. [OP … [… XP …]]
c. [XP … ~~XP~~ ➲ YP…]
d. [OP … ~~OP~~ ➲ YP…]

This said, there arises an immediate problem with our adoption of a null operator OP and the "guiding minimalist desideratum" presented briefly in section 3.1 above. In line with the program sketched most clearly in Hornstein (2001), an adoption of minimalist strategies one may call "rigorous minimalism" (Grohmann 2003), there should be no room in the grammar for such null elements. If it turns out that a minimalist approach should indeed dispense with theory-internal constructs, such as empty operators, we would need to find an alternative for the analysis of demonstrative doubling sketched above. Hornstein suggested movement-alternatives for the phenomena investigated in his work and thus circumvented the need to posit OP (such as relativization or *tough*-constructions).

We will leave this issue aside for the time being, simply pointing to work by Cedric Boeckx (see Boeckx 2003a, 2003b, Boeckx and Grohmann 2004a), where a movement-analysis can be envisioned if it targets a sub-part of the demonstrative in base position. There could be an additional layer on top of the demonstrative itself which undergoes the relevant movement into the left periphery (and then within the Ω-Domain to yield Copy Spell Out) — possibly a notational variant in terms of outcome, but not in terms of mechanics and additional stipulations.

What an OP-less analysis might yield, however, is the fact that Greek makes available two readings, the deictic interpretation with overt demonstrative high up preceding the article and the discourse-anaphoric interpretation with the article preceding the overt demonstrative. If the null element is not an operator, there is no need for it to anti-locally move to the same position within CP. Rather, one could imagine that the null (non-operator) element (whatever its exact nature) targets a topic position (SpecTopP), which would be in line with the anaphoric, "given" nature of the structure's interpretation. Such an analysis would also reduce the number of "operator-positions" in the clause/nominal layer.

These issues are addressed in detail in concurrent research (Grohmann and Panagiotidis 2005), where we also consider cross-linguistic implications of the analysis and apply it to linguistically diverse languages such as Spanish, Macedonian, Romanian, or Beiruti Arabic, and discuss structures that, at face value, might look similar but in fact are not (in Irish Gaelic and Maori).

6. Conclusion

In this paper we looked at demonstrative constructions in Modern Greek nominal structures and applied the Anti-Locality Hypothesis to the Greek DP. Our main proposal is that:

(i) nominals have the same Prolific Domains as clauses,
(ii) the Condition of Exclusivity applies to nominals as well, and
(iii) thus anti-locality is a general condition on the grammar (clauses and nominals).

In particular, we analyze the Greek doubling article in nominal demonstrative structures as the Copy Spell Out of an anti-locally moved demonstrative. This analysis accounts for the curious co-occurrence of demonstrative and article: the latter is but a grammatical formative.

References

Abney, Stephen. 1987. *The English Noun Phrase in Its Sentential Aspect*. Ph.D. dissertation, Cambridge MA, MIT

Adger, David. 2003. *Core Syntax*. Oxford, Oxford University Press.

Alexiadou, Artemis and Melita Stavrou. 1998. "(A)symmetries in DPs and Clauses: Evidence from Derived Nominals". *The Linguistic Review* 15:257-276.

Anagnostopoulou, Elena. 1994. *Clitic dependencies in Modern Greek*. Ph.D.

dissertation, University of Salzburg.
Belletti, Adriana (ed.). 2004. *Structures and Beyond: The Cartography of Syntactic Structures, Vol. 3*. Oxford, Oxford University Press.
Bernstein, Judy B. 1997. "Demonstratives and Reinforcers in Romance and Germanic Languages". *Lingua* 102:87-113.
Bernstein, Judy B. 2001. "The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain". In M. R. Baltin and C. Collins, (eds.) *The Handbook of Contemporary Syntactic Theory*. 536-561. Oxford, Blackwell.
Boeckx, Cedric. 2003a. *Islands and Chains*. Amsterdam/Philadelphia, John Benjamins.
Boeckx, Cedric. 2003b. "(In)direct Binding". *Syntax* 6.3:213-236.
Boeckx, Cedric and Kleanthes K. Grohmann. 2004a. "Sub-Move: Towards a Unified Account of Scrambling and D-Linking". In D. Adger, B. Plunkett, and G. Tsoulas (eds.) *Peripheral Positions*. 241-257. Oxford, Oxford University Press.
Boeckx, Cedric and Kleanthes K. Grohmann. 2004b. "Left Dislocation in Germanic". In Werner Abraham (ed.) *Focus on Germanic Typology*. 131-144. Berlin, Akademie-Verlag.
Bošković, Željko. 1994. "D-Structure, Theta-Criterion, and Movement into Theta-Positions". *Linguistic Analysis* 24:247-286.
Brugè, Laura. 1996. "Demonstrative Movement in Spanish: A Comparative Approach". *University of Venice Working Papers in Linguistics* 6:1-53.
Campbell, Richard. 1996. "Specificity Operators in SpecDP". *Studia Linguistica* 50:161-188.
Castillo, Juan Carlos. 2001. *Thematic Relations within Nouns*. Ph.D. dissertation, University of Maryland, College Park.
Chomsky, Noam. 1986. *Barriers*. Cambridge MA, MIT Press.
Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In K. Hale and S. J. Keyser (eds.) *The View from Building 20. Essays in Honor of Sylvain Bromberger*. 1-52. Cambridge MA, MIT Press.[Reprinted in Chomsky 1995, 167-217.]
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA, MIT Press.
Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework". In R. Martin, D. Michaels, and J. Uriagereka (eds.) *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. 89-155. Cambridge MA, MIT Press.
Chomsky, Noam and Howard Lasnik. 1978. "A Remark on Contraction". *Linguistic Inquiry* 9:268-274.
Cinque, Guglielmo. 1994. "On the Evidence for Partial N-Movement in the

Romance DP". In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (eds.) *Paths towards Universal Grammar. Studies in Honor of Richard S. Kayne*. 85-110. Washington, DC: Georgetown University Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford, Oxford University Press.

Cinque, Guglielmo (ed.). 2002. *The Structure of IP and DP: The Cartography of Syntactic Structures, Vol. 1*. Oxford, Oxford University Press.

Giannakidou, Anastasia and Melita Stavrou. 1999. "Nominalization and Ellipsis in the Greek DP". *The Linguistic Review* 16:295-331.

Giusti, Giuliana. 1997. "The Categorial Status of Determiners". In L. Haegeman (ed.) *The New Comparative Syntax*. 95-123. London, Longman.

Grimshaw, Jane. 1991. *Extended Projection*. Ms., Brandeis University.

Grimshaw, Jane. 2003. *Words and Structure*. Stanford, CA, CSLI Publications.

Grohmann, Kleanthes K. 2000. *Prolific Peripheries: A Radical View from the Left*. Ph.D. dissertation, University of Maryland, College Park.

Grohmann, Kleanthes K. 2003. *Prolific Domains. On the Anti-Locality of Movement Dependencies*. Amsterdam/Philadelphia, John Benjamins.

Grohmann, Kleanthes K. 2005. "The Road to PF". Paper presented at the *17th International Symposium on Theoretical and Applied Linguistics*. Aristotle University Thessaloniki, 15-17 April 2005.

Grohmann, Kleanthes K. and Liliane Haegeman. 2003. "Resuming Reflexives". *Proceedings of the 19th Scandinavian Conference of Linguistics*. *Nordlyd* 31.1:46-62.

Grohmann, Kleanthes K. and Phoevos Panagiotidis. 2004. "Demonstrative Doubling in Greek". In P. Chandra, T. Fujii, U. Soltan, and M. Yoshida, (eds.) *University of Maryland Working Papers in Linguistics* 13:109-134.

Grohmann, Kleanthes K. and Phoevos Panagiotidis. 2005. "Demonstrative Doubling". Unpublished manuscript, University of Cyprus and Cyprus College, Nicosia. [expected to be re-submitted: June 2005]

Haegeman, Liliane. 2001. "DP-Periphery and Clausal Periphery: Possessor Doubling in West Flemish". In D. Adger, C. de Cat, and G. Tsoulas, (eds.) *Peripheries: Syntactic Edges and Their Effects*. 211-240. Dordrecht, Kluwer.

Hale, Kenneth and Samuel Jay Keyser. 1993. "On Argument Structure and the Lexical Expression of Grammatical Relations". In K. Hale and S. J. Keyser (eds.) *The View from Building 20. Essays in Honor of Sylvain Bromberger*. 53-110. Cambridge MA, MIT Press.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford, Blackwell.

Hornstein, Norbert, Jairo Nunes, and Kleanthes K. Grohmann. In press. *Understanding Minimalism*. Cambridge, Cambridge University Press.

Horrocks, Geoffrey and Melita Stavrou. 1987. "Bounding Theory and Greek Syntax: Evidence for Wh-Movement in NP". *Journal of Linguistics* 23:79-108.

Longobardi, Giuseppe. 1994. "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form". *Linguistic Inquiry* 25:609-665.

Manolessou, Io and Phoevos Panagiotidis. 1999. "Ta dhiktika tis ellinikis te i shetikes epipokles" [Greek Demonstratives and the Relevant Complications]. In S. L. Tsoxatzidis, M. Koutita, and A. Haralambopoulos (eds.) *Studies in Greek Linguistics: Proceedings of the 19th Annual Meeting of the Department of Linguistics, Faculty of Philosophy*. 199-212. Thessaloniki, Aristotle University of Thessaloniki.

Murasugi, Keiko and Mamoru Saito. 1995. "Adjunction and Cyclicity". In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia, eds. *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. 302-317. Stanford CA, CSLI Publications.

Nunes, Jairo. 1995. *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*. Ph.D. dissertation, University of Maryland, College Park.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge MA, MIT Press.

Panagiotidis, Phoevos. 2000. "Demonstrative Determiners and Operators: The Case of Greek". *Lingua* 110:717-742.

Panagiotidis, Phoevos. 2002. *Pronouns, Clitics and Empty Nouns: 'Pronominality' and Licensing in Syntax*. Amsterdam/Philadelphia, John Benjamins.

Platzack, Christer. 2001. "Multiple Interfaces". In U. Nikanne and E. van der Zee (eds.) *Cognitive Interfaces. Constraints on Linking Cognitive Information*. 21-53. Oxford, Oxford University Press.

Pollock, Jean-Yves. 1989. "Verb Movement, UG and the Structure of IP". *Linguistic Inquiry* 20:365-424.

Radford, Andrew. 1999. "NP Shells". Unpublished manuscript, University of Essex, Colchester.

Ritter, Elisabeth. 1991. "Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew". In S. Rothstein (ed.) *Perspectives on Phrase Structure. Syntax and Semantics 26*. 37-62. New York NY, Academic Press.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In L.

Haegeman (ed.) *Elements of Grammar: Handbook of Generative Syntax*. 281-337. Dordrecht, Kluwer.

Rizzi, Luigi. 2004. *The Structure of CP and IP: The Cartography of Syntactic Structures, Vol. 2*. Oxford, Oxford University Press.

Roussou, Anna and Ianthi-Maria Tsimpli. 1994. "On the Interaction of Case and Definiteness in Modern Greek". In I. Philippaki-Warburton, K. Nicolaidis, and M. Sifianou (eds.) *Themes in Greek Linguistics*. 69-76. Amsterdam/Philadelphia, John Benjamins.

Shlonsky, Ur. 1992. "Resumptive Pronouns as a Last Resort". *Linguistic Inquiry* 23:443-468.

Stavrou, Melita and Geoffrey Horrocks. 1989. "Enklitikes ke dhiktikes antonimies mesa stin OF" [Enclitic and Demonstrative Pronouns inside NP]. In *Proceedings of the 10th Annual Meeting of the Department of Linguistics, Faculty of Philosophy*. 225-243. Thessaloniki: Aristotle University of Thessaloniki.

Szabolcsi, Anna. 1994. "The Noun Phrase". In F. Kiefer and K. É. Kiss, eds. *The Syntactic Structure of Hungarian. Syntax and Semantics 27*. 179-274. New York NY, Academic Press.

Ticio, Emma. 2003. *On the Structure of DPs.* Ph.D. dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1996. "From Being to Having". *University of Maryland Working Papers in Linguistics* 4:152-172.

Uriagereka, Juan. 1999. "Multiple Spell Out". In S. D. Epstein and N. Hornstein, eds. *Working Minimalism*. 251-282. Cambridge MA: MIT Press.

Deriving relative and factive clauses

Enoch O. Aboh
University of Amsterdam

1. Introduction

The traditional analysis for relative clauses suggests that these are comparable to DP-modifiers that are adjoined to the noun phrase. It is therefore commonly assumed that the relative clause involves a CP that right adjoins to the DP. In addition, spec,CP of the relative clause hosts a null operator that is co-indexed with the relative head noun. In terms of this analysis, a sentence like (1a) is assigned the partial representation in (1b).

(1) a. The man that Mary invited.
b. [$_{DP}$ the man$_i$ [$_{DP}$ [$_{CP}$ Op$_i$ [$_{C°}$ that [$_{IP}$ Mary invited t$_i$]]]]]

Under Kayne (1994), however, the relative clause involves a complementation structure similar to that in (2), where the proposition (IP) merges as the complement of the relative complementizer under C to form CP, whose specifier hosts the relative noun that raises there. The CP with the relative head noun in its specifier then merges with D, which hosts the determiner to form the DP, that is, the relative clause.

(2) [$_{DP}$ [$_{D}$ the [$_{CP}$ man$_i$ [$_{C°}$ that [$_{IP}$ Mary will invite t$_i$]]]]]

The question naturally arises how to choose between these two competing analyses. This paper discusses new data from Gbe (Kwa)[1] that can be interpreted as evidence for the complementation analysis under (2). I further

[1] See Capo (1991), Aboh (2004a) and references cited there for discussion on the Gbe as a cluster of Kwa languages.

show that the complementation analysis explains the contrast between relative and factive clauses.

The sentences under (3) and (4) illustrate the contrast between relative clauses and factive clauses in Gungbe. The example (3a) shows that, in this language, the NP (or the extended projection of N) may precede the specificity marker glossed here as Det, and the number marker, glossed as Num. I assume that these markers manifest the determiner system of Gungbe, that is, they are D-type elements, the nominal counterparts of C-type elements (see section 2 and references cited there for discussion). The specificity marker refers to strongly D-linked elements in the sense of Pesetsky (1987), and is translated here as "the aforementioned…" As for the parentheses in (3a), they indicate that the Gungbe D-type elements need not be realized (simultaneously), and the language allows for bare noun phrases. Sentence (3b) shows that the nominal modifiers, such as adjectives, numerals and demonstratives, must follow the noun, but precede the D-type elements. Note from example (3c) that, like noun modifiers, the Gungbe relative clause is sandwiched between the head noun and the D-type elements that are set off to the right (i.e. NP-[relative clause]-D, see Aboh 2004a for detailed discussion on the DP in Gungbe).

(3) a. Kòfí wè xɔ̀ àgásá (lɔ́) (lɛ́)
Kofi Foc buy crab Det Num
'KOFI bought the [aforementioned] crabs'

b. Kòfí wè xɔ̀ àgásá [ɖàxó àtɔ̀n éhè] lɔ́ lɛ́
Kofi Foc buy crab big Nral Dem Det Num
'KOFI bought these [aforementioned] three big crabs'

c. Kòfí wè xɔ̀ àgásá ɖàxó [ɖě mí wlé] lɔ́ lɛ́
Kofi Foc buy crab big that$_{[Rel]}$ 1pl catch Det Num
KOFI bought the [aforementioned] big crabs that we caught'

At first sight, these examples support the adjunction approach because the noun modifiers and the relative CP clause share the same space within the noun phrase. Yet, in situations where a noun immediately precedes the D-type elements and surfaces to the left of the relative complementizer (i.e. NP-D-[CP clause]), the sentence is assigned a factive reading, as in (4).

(4) Àgásá ɖàxó lɔ́ lɛ́ [ɖě mí wlé] vɛ́ ná Kòfí
crab big Det Num that$_{[Rel]}$ 1pl catch hurt for Kofi
'The fact that we caught the aforementioned big crabs hurt Kofi'
*'The aforementioned big crabs that we caught hurt Kofi'

Since this factive meaning is not available in the English example (1a), or in the Gungbe example (3c), I propose to describe the Gungbe factive constructions as sentences where a full DP immediately precedes the relative marker *ɖě* and the resulting sentence is interpreted factively. These sentences are translated here by their English equivalent introduced by '*the fact that*…', which are traditionally regarded as clausal complements to nouns. I argue in this paper that these constructions are not complements to the head nouns, but truncated relative clauses (see section 3.2 and subsequent). Note from sentence (5) that factive clauses differ in interpretation from relative clauses. Indeed, if factive and relative clauses were identical, this sentence would be a contradiction because the first part would mean that the crab that *Kòfí* caught was good, while the second part would imply that the very same crab was not good (see Collins 1994 for discussion).

(5) Àgásá lɔ́ [ɖě Kòfí wlé] nyɔ́n,
crab Det that$_{[Rel]}$ Kofi catch good
àmɔ́n àgásá lɔ́ kpàkpà má nyɔ́n
but crab Det itself Neg good
'The fact that Kofi caught the crab was a good thing but the crab (itself) wasn't good/sweet'

Assuming there is a direct correlation between the relative clause reading assigned to the sequence NP-[CP clause]-D, and the factive reading associated with the structure NP-D-[CP clause] in (3–4), there seems to be no obvious explanation for such contrast under the adjunction analysis. On the contrary, the complementation view offers a possible analysis that captures the Gungbe data in a straightforward manner and appears to handle this contrast without further stipulation. The reason for this is that the complementation analysis implies two layers: the CP level and the DP level. In what follows I argue that it is the presence of this DP level (or the lack thereof) that determines the restrictive relative versus factive reading found in Gbe. Before getting on to this discussion, the following section presents the architecture of the D-system that is assumed in this paper.

2. The DP-Structure

Following work on the syntax of DPs (e.g. Abney 1987; Szabolcsi 1994; Ritter 1995; Siloni 1997; Carstens 2000; Bernstein 2001a, b; Longobardi 1994, 2001; Panagiotidis 2000, Aboh 2002, 2004a, b, among others), I assume that the nominal structure is parallel to the clausal structure in that it involves a predicate layer (i.e. NP), a functional layer (i.e. ΣP) that encodes

inflectional specifications (e.g. agreement and definiteness), and a left periphery (i.e. DP). Extending Rizzi's (1997) split-C hypothesis to the nominal domain, I propose that the D-system splits into distinct projections that encode number, focus, topic, as represented in (6) (see Giusti 1996, and Aboh 2004a, b for discussion).

(6) [$_{DP}$ [$_{D}$ [$_{TopP}$ [$_{Top}$ lɔ́ [$_{FocP}$ [$_{Foc}$ tɛ́ [$_{NumP}$ [$_{Num}$ lɛ́ [$_{ΣP}$ [$_{NP}$]]]]]]]]]

Under this hypothesis, D is a subordinator comparable to the clausal C. It heads the highest projection of the D-system, whose specifier provides an escape hatch for extraction out of the DP as is the case in the Hungarian dative extraction in (7) (Szabolcsi 1994). As the highest projection of the D-system, DP represents the interface between the noun phrase (or the nominal predicate) and the discourse.

(7) [Péter-nek] mindenki csak [a kalap-já -t] látta
Peter-DAT everyone only the hat-Poss-3sg-ACC saw
'As for Peter, everyone saw only his hat (e.g. no one saw his coat)'
(Szabolcsi 1994: 205)

Num°, on the other hand, delimits the nominal left periphery downward as the interface between the nominal left periphery and the inflectional domain. It encodes number (i.e. [±number]) as well as nominal agreement (e.g. definiteness) features that match those of the nominal inflectional domain. This can be seen in the Gungbe examples under (8). These examples show that while a bare noun is interpreted as indefinite, definite or generic, depending on the context, as in (8a), a noun plus the number marker is necessarily definite, as in (8b).

(8) a. Mì sà àkwékwè átɔ́n ná mì [Indefinite]
2pl sell banana five for 1sg
'Sell me five bananas'
b. Mì sà àkwékwè átɔ́n lɛ́ ná mì [Definite]
2pl sell banana five Num for 1sg
'Sell me the five bananas'

However, the examples under (9) show that the complex [noun + numeral modifier] *àkwékwè átɔ́n* 'five banana' cannot be marked as specific in the absence of the number marker. Recall from examples (3a) that the specificity marker and the number marker need not co-occur.

(9) a. Mì sà àkwékwè átɔ́n lɔ́ lɛ́ ná mì
2pl sell banana five Det Num for 1sg
'Sell me the aforementioned five bananas'
b.*Mì sà àkwékwè átɔ́n lɔ́ -- ná mì
2pl sell banana five Det for 1sg
'Sell me the aforementioned five bananas'

It appears from this description that the number marker *lɛ́* encodes definiteness (8b) and establishes a concord between the plurality expressed in the nominal inflectional domain and the nominal left periphery (9a). Assuming that phrase structure builds in a bottom-up fashion, I interpret the impossibility of (9b) as evidence that the head Num must merge before D. This is indirect evidence for the hierarchy DP > NumP assumed throughout this paper (see Aboh 2004a for discussion).

Under the split-D hypothesis, I further propose that the specificity marker *lɔ́* in (10a) expresses the topic head, while the nominal question particle *tɛ́* in (10b) is comparable to the Gungbe clausal focus marker *wɛ̀*, and therefore encodes the nominal focus head (Aboh 2004b).

(10) a. Kɔ̀kú mɔn távɔ̀ títán bò ɖɔ̀
Koku see table first and say
émì ná xɔ̀ távò lɔ́
3sg Fut buy table Det
'Koku saw the first table and said that he would buy that specific table'
b. [Távò xɔ́xɔ́ tɛ́] wɛ̀ Kòfí xɔ̀?
table old Q Foc Kofi buy
'Which old table did Kofi buy?'

Example (11) further shows that the topic-focus articulation can project (assuming that *ɖɛ̌* and the specificity marker *lɔ́* compete for the same position Top).

(11) [Távò xɔ́xɔ́ ɖɛ̌ tɛ́ lɛ́] wɛ̀ Kòfí xɔ̀?
table old Top$_{[particle]}$ Q Numb Foc Kofi buy
'Which one of the old tables did Kofi buy?'

Assuming representation (6) above, I conclude that the NP- *lɔ́*/*ɖé-lɛ́* order, which is typical of Gbe, derives from movement of the nominal inflectional domain as a whole (i.e. ΣP) to spec,NumP and spec,TopP (ignoring the focus

projection), in order to check the features number and specificity, as illustrated in (12).

(12) $[_{DP}\,[_{D}\,[_{TopP}$ távò $[_{Top}$ lɔ́ $[_{FocP}[\,_{Foc}$ té $[_{NumP}$ $t_{távò}$ $[_{Num}$ lé $[_{\Sigma P}\,[_{NP}\,t_{távò}]]]]]]]]$

The representation under (12) therefore suggests that the Gungbe DP involves the nominal counterpart of predicate/proposition fronting because the inflectional domain ΣP that includes the noun phrase must front to spec,NumP and spec,TopP, due to number and specificity licensing. The following sections discuss to what extent such fronting strategy may apply to relative clauses and what the implications are with regard to the presence or absence of fronting rules for the described relative versus factive asymmetry.

3. The derivation of relative and factive clauses in Gungbe

As I briefly discussed in the introductory section, Gungbe manifests a relative versus factive clause asymmetry whereby the former reading is assigned to the sequence NP-[CP-clause]-D, while the latter is assigned to the sequence NP-D-[CP-clause]. I first discuss relative clauses.

3.1. *Relative clauses*

Gungbe manifests the types of relative clauses in (13). Example (13a) illustrates what could be referred to as a bare noun relative clause (i.e. a bare noun immediately precedes the relative marker). In this context, the head noun is interpreted as definite, thanks to the relative clause. We can therefore describe the Gungbe relative sentences as restrictive relative clauses (de Vries 2002). In (13b), however, the relativised noun is interpreted as specific due to the presence of the specificity marker to the right edge. Similarly the example (13c) illustrates a relative clause where the head noun is plural and the plural marker occurs to the right edge. Finally, the sentence under (13d) indicates that when the relative noun is plural and specific, both the specificity marker and the number marker must surface to the right edge.

(13) a. Kòfí xɔ̀ àgásá [ɖě mí wlé]
Kofi buy crab $that_{[Rel]}$ 1pl catch
'Kofi bought the crab that we caught'

b. Kòfí xɔ̀ àgásá [ɖě mí wlé] lɔ́
Kofi buy crab $that_{[Rel]}$ 1pl catch Det

‘Kofi bought the [aforementioned] crab that we caught’

c. Kòfí xɔ̀ àgásá [ɖě mí wlé] lɛ́
Kofi buy crab that$_{[Rel]}$ 1pl catch Num
‘Kofi bought the crabs that we caught’

d. Kòfí xɔ̀ àgásá [ɖě mí wlé] lɔ́ lɛ́
Kofi buy crab that$_{[Rel]}$ 1pl catch Det Num
‘Kofi bought the [aforementioned] crabs that we caught’

Assuming the derivation in (12), an analysis of the Gungbe relatives in terms of adjunction would be that the relative clause right adjoins to ΣP, and this complex phrase pied-pipes to spec,NumP and spec,TopP, as shown in (14).

(14)

[$_{DP}$[$_{D°}$[$_{TopP}$[$_{Top°}$ lɔ́ [$_{NumP}$[$_{Num°}$ lɛ́ [$_{ΣP}$ àgásá$_i$ [$_{ΣP}$[$_{CP}$ Op$_i$ [$_{C°}$ɖě [$_{IP}$ mí wlé t$_i$]]]]]]]]]]

Setting aside issues about free adjunction rules (see Kayne 1994), such an analysis does not seem to carry over to the distribution of the D-type elements with regard to the head noun, and the implications of the discussed distributive facts with regard to the relative versus factive asymmetry as illustrated in (15) and schematized in (16).

(15) a. Àgásá ɖàxó [ɖě mí sà] ná Kòfí lɔ́ lɛ́
crab big that$_{[Rel]}$ 1pl sell to Kofi Det Num
‘the aforementioned big crabs that we sold to Kofi’
*‘the fact that we sold the [aforementioned] big crabs to Kofi’

b. Àgásá ɖàxó lɔ́ lɛ́ [ɖě mí wlé] vɛ́ ná Kòfí
crab big Det Num that$_{[Rel]}$ 1pl catch hurt for Kofi
‘The fact that we caught the aforementioned big crabs hurt Kofi’
*‘The [aforementioned] big crabs that we caught hurt Kofi’

(16) a. NP-[CP clause]-lɔ́-lɛ́ = restrictive relative clause
b. NP- lɔ́-lɛ́-[CP clause] = factive clause

The only difference between (16a–b) is the placement of the relative CP clause vis-à-vis the determiners. Under the adjunction analysis, a possible solution that comes to mind would be to treat the asymmetry in (16) as evidence that the CP clause may adjoin at different sites prior to movement: at the ΣP level in (16a) or at the DP level in (16b). In the literature, however (e.g. Toribio 1992, de Vries 2002), such distinction is related to the

restrictive versus appositive reading, but not to the relative versus factive clause asymmetry that Gungbe manifests.[2] Similarly, that certain adverbs may adjoin to different sites within the clause (e.g. VP versus IP) is traditionally linked to scope effects, but not to the type of change in meaning that we face here, where sequence (16a) equals to a restrictive relative clause (with the relative CP) acting semantically as a modifier of the noun, while in (16b), the clause relates to the event that is expressed by the verb inside what in (16a) appears as the relative CP. Put differently, while (16a) has to do with the relative noun, (16b) has to do with the event head, that is, the verb.

Building on this, I propose that the contrast described in (16), as well as the restrictive nature of Gungbe relative clauses can be accounted for if we adopt Kayne's (1994) complementation analysis for relative clauses.[3] Under the split-D hypothesis, I further propose that relative clauses manifest a structure where Num° encoded by the number marker merges with a CP clause headed by the relative complementizer as its complement to form NumP. The latter merges with Top°, encoded by the specificity marker to form TopP, which merges with D to form DP. In terms of Kayne (1994) the relativised noun must move in overt syntax to spec,CP, presumably to check the strong features under C. Building on this, I further argue that, in the Gbe languages, the CP-clause headed by the relative marker ɖě (with the NP raised into its specifier) must raise to spec,NumP and spec,TopP as an instance of predicate fronting in order to check the features [specific, plural]. Under this approach, a sentence like (17) is derived as in (18).

(17) Kòfí xɔ̀ àgásá [ɖě mí wlé] lɔ́ lɛ́
Kofi buy crab that$_{[Rel]}$ 1pl catch Det Num
'Kofi bought the [aforementioned] crabs that we caught'

(18) [$_{DP}$[$_{D°}$[$_{TopP}$[$_{Top°}$ lɔ́ [$_{NumP}$ t$_{CP}$ [$_{Num°}$ lɛ́ [$_{CP}$ àgásá$_i$ [$_{C°}$ ɖě [$_{IP}$ mí wlé t$_i$]]]]]]]]]

This would mean that, being a complement of D, the relative clause is subject to the same raising constraint as the inflectional nominal domain ΣP as illustrated in (12) (see Aboh 2002, 2004a). Note also that predicate or proposition fronting in (18) is last resort generalized pied-piping process. The operation seeks to raise the features of the NP in spec,CP (i.e. the relative head) but the whole CP is pied-piped for convergence.

[2] Gungbe does not have appositive relative clauses.
[3] See also Bianchi (1995), de Vries (2002) and references cited there for discussion.

There are open questions as to the categorial status of the constituent in spec,CP (i.e. whether it is an NP or a DP). Under Kayne's (1994) approach, this constituent is a bare NP or some of its extended projection, but not a DP. Alternatively, De Vries (2002) proposes that the relative head starts out as a full relative DP, whose head hosts the relative pronoun when there is one.[4] This would mean that a German relative clause such as (19a) could be partially represented as in (19b) whereby the NP moves DP-internally to the specifier of the relative DP prior to raising to spec,CP (De Vries 2002: 123). In this framework, the strong [wh] features of the relative pronoun (e.g. *which*, *whom*) under D trigger movement of the relative DP to spec,CP. In addition, the outer D is still needed because it allows the relative head noun to ultimately associate with a determiner, so that its ϕ-features can be checked and the whole relative clause can function as a proper argument.

(19) a. Ich fürchte den Herrn der eine Pistole trägt
I fear the.Acc gentleman.Acc who.Nom one gun carries
b. $[_{DP}\ [_{D}$ den $[_{CP}\ [_{DP\text{-}rel}$ Herrn$_h$ der $t_h]_i\ [_{C^\circ} \varnothing [_{IP}\ t_i$ eine Pistole trägt]]]]]

The empirical motivation for this view is that the gap inside the CP clause represents an argument position (i.e. a legitimate position for argument DPs), and the trace of the relative noun is parallel to a variable (i.e. a DP-trace) in many respects (e.g. binding or control properties, licensing of parasitic gaps, case marking, weak island effect). Finally, in some languages (e.g. Southern Dutch dialects) the relative NP or PP may be realized alongside with the relative pronoun and the relative complementizer (den Besten 1998).

(20) a. De stoelen die (dat) kapot zijn
Det chairs which Comp broken are
'The chairs which are broken' (den Besten 1998)
b. De man an wie dat ik het gegeven heb
Det man to whom Comp I it given have
'The man to whom I have given it' (den Besten p.c.)

Building on this, De Vries (2002) further suggests that in languages where there is no relative pronoun (e.g. Gungbe), or else in 'that-relatives' the head D is empty.

Given that Gungbe manifests the relative complementizer *dǐ* 'that' only, there is no empirical ground in this language for the IP-internal relative DP hypothesis for relative clauses. However, I conjecture in what follows that

[4] See also Bianchi (1995) for discussion.

the presence of a relative DP in spec,CP that may or may not interact with the outer D provides us with a possible account for the relative versus factive reading discussed above. Put differently, I suggest that factive clauses are truncated relative clauses whereby the truncated CP clause (i.e. without the outer DP) hosts a factive (or an event) DP in spec,CP. However, the proposal made here departs slightly from De Vries's (2002) own work in that I'm assuming (following Szabolcsi 1994, Longobardi 1994 among others) that the outer D is an argument introducer, which is why it can select (or nominalise) a clause in various languages (e.g. the Gungbe clausal determiner, see example 25).

3.2. *Factive clauses*

The sentences under (15) and the descriptions in (16) indicate that relative clauses and factive clauses differ with respect to both the interpretation and the positioning of the CP clause. The following examples illustrate further differences between the two constructions. The sentences under (21) indicate that, unlike relative clauses, factive clauses may involve two different structures without a change in meaning. In (21a) the DP object is left adjacent to the relative complementizer and the sentence is read as factive. In (21b), however, it is the verb that fronts, leaving a copy IP-internally. Yet, the sentences (21a) and (21b) have the same meaning even though they may differ pragmatically. In (21a), the speaker is more concerned with referring to specific crabs that were caught, while in (21b) s/he puts some emphasis on the event of catching those crabs.

(21) a. Àgásá lɔ́ lɛ́ [ɖě mí wlé] vɛ́ ná Kòfí
crab Det Num that$_{[Rel]}$ 1pl catch hurt for Kofi
'The fact that we caught the crabs hurt Kofi'

b. **Wlé** [ɖě mí **wlé**] Àgásá lɔ́ lɛ́ vɛ́ ná Kòfí
catch that$_{[Rel]}$ 1pl catch crab Det Num hurt for Kofi
'The fact that we CAUGHT the crabs hurt Kofi'

I will not discuss these subtleties here. The point I want to make though is that the parallel between (21a) and (21b) does not arise in relative clauses.

(22) a. Àgásá [ɖě mí wlé] lɔ́ lɛ́ víví
crab that$_{[Rel]}$ 1pl catch Det Num sweet
'The crab that we caught were sweet'

b. Wlé [ɖě mí wlé] Àgásá lɔ́ lɛ́ víví
catch that$_{[Rel]}$ 1pl catch crab Det Num sweet
*'The crabs that we caught were sweet'
'The fact that we caught the crabs was great [i.e. we enjoyed catching the crabs]'

These facts point to the interesting possibility that even though the factive reading in (21a) is achieved by fronting an internal DP argument, the intended meaning (e.g. in 21b) is that of 'the event of catching crabs'. Put differently, the fronted internal DP argument relates to the event expressed by the verb *wlé* 'catch'. That this might be the right characterization is further suggested by the fact that factive constructions manifest a type of external versus internal argument asymmetry. The sentences under (23a) and (23b) indicate that internal arguments (e.g. theme, goal, beneficiary) may be fronted in factive constructions, while factive constructions with fronted external arguments (e.g. agents) are excluded or at least considerably degraded (see also Collins 1994).[5]

(23) a. [Dáwè lɔ́]$_i$ [ɖě náwè lɔ́ kàn wé xlán t$_i$]
man Det that$_{[Rel]}$ woman Det write letter to
vɛ́ ná mì gbáú
hurt for me a lot
'The fact that the woman wrote a letter to **the man** hurt me a lot'

b. [Wé lɔ́]$_i$ [ɖě náwè lɔ́ kàn t$_i$ xlán dáwè lɔ́]
letter Det that$_{[Rel]}$ woman Det write to man Det
vɛ́ ná mì gbáú
hurt for me a lot
'The fact that the woman wrote **a letter** to the man hurt me a lot'

c. ??/*[Náwè lɔ́]$_i$ [ɖě t$_i$ kàn wé xlán mì]
woman Det that$_{[Rel]}$ write letter to me
vɛ́ ná mì gbáú
hurt for me a lot
'The fact that **the woman** wrote to me hurt me a lot'

Relative clauses, on the other hand, show no such effect: all arguments can be relativised, and there is no subject versus object asymmetry.

[5] The boldface in the translation is meant to show that the fronted elements appear to have a more prominent discourse status, but I leave this question open pending new pragmatic research in Gbe.

(24) a. Dáwè$_i$ [ɖě náwè lɔ́ kàn wé xlán t$_i$] lɔ́ wá
man that$_{[Rel]}$ woman Det write letter to Det come
'The aforementioned man that the woman wrote a letter to came'

b. Wé$_i$ [ɖě náwè lɔ́ kàn t$_i$ xlán dáwè lɔ́] lɛ́ ɖíè
letter that$_{[Rel]}$ woman Det write to man Det Num here
'Here are the letters that the woman wrote to the man'

c. Náwè$_i$ [ɖě t$_i$ kàn wé xlán dáwè lɔ́] lɛ́ wá
man that$_{[Rel]}$ write letter to man Det Num come
'The women who wrote a letter to the man came'

I interpret these facts as evidence that both the fronted internal DP argument *àgásá* 'crab' and the verb *wlé* 'catch' in (21a–b) share some event feature.[6] That internal arguments (but not external arguments) typically tend to manifest such feature is not surprising. For instance, several works on aspect licensing (e.g. Aktionsart), event structure, and argument structure have shown the role of internal arguments in determining certain event types (e.g. Tenny 1987, Dowty 1991, Torrego 1998, Basilico 1998).

Building on this, I further propose that, unlike the relative DP, whose D bears a strong nominal [wh] feature that forces movement of the relativised DP argument to spec,CP (see De Vries 2002), the D head of a factive DP is endowed with a strong [wh] event feature that may trigger movement of an event DP to spec,CP.[7] That the Gungbe D (i.e. of the internal argument) can encode such feature is indirectly supported by the presence in this language of a clausal determiner that is homophonous with the nominal determiner. In example (25), the clausal determiner *lɔ́* occurs sentence-finally and indicates that the event being referred to is pre-established in discourse (see Aboh 2004a and references cited there for discussion).

(25) Đé Kòfí hɔ̀n **lɔ́** kpácá dó mì káká
as Kofi flee Det$_{CL}$ surprise at me a lot
bɔ̀ má nyɔ́n nú ɖě ná ɖɔ̀
and 1sg-Neg know thing that 1sg-Fut say
'As Kofi fled, I was so surprised I didn't know what to say'

Assuming that the DP in spec,CP may also bear an event [wh] feature, I therefore conclude that the D head of a relativised DP may bear an argument-related [wh] feature (i.e, a nominal feature), or an event-related

[6] More research is needed to determine the type of event feature involved here, but I refer the interested reader to the cited references for discussion.

[7] I'm using the label [wh] in the sense of a strong C-type feature that is checked against the relative feature under C, and expressed by the relative complementizer.

[wh] feature (i.e. a verbal feature). As a first description, I label the former as [wh_A] and the latter as [wh_E]. These features trigger the relativisation of the argument or the event such that, a relative DP with the feature [wh_A] binds a relevant argument in the clause, while a relative DP with the feature [wh_E] binds the event. Given the parallel in (21), I further propose that the event head (i.e. the verb) may bear the [wh_E] feature as well.[8] This amounts to saying that factive constructions are types of event relative clauses.

With this in mind, let us go back to De Vries' (2002) analysis sketched in (19) and let us assume, as before, that the outer DP layer is needed in relative clauses in order for the whole relative clause to function as an argument. If this is indeed the case, there is the possibility that a factive clause (i.e. an event relative clause) that embeds an event DP in spec,CP, but lacks the outer DP layer converges, just as any CP clause with some relevant material in spec,CP (e.g. *wh*-questions in matrix clauses). Put differently, the outer DP does not and therefore cannot merge in such event relative clauses because the latter do not qualify as arguments. This leads me to conclude that relative and factive constructions differ in that the latter don't involve the DP layer typical of arguments and relative clauses.

This analysis captures both the surface similarities and differences between relative and factive clauses. Unlike the relative clauses analyzed in (18), the factive constructions cannot involve D-type elements to the right edge of the clause because there is no outer DP involving a TopP such that the whole factive clause would move to spec,TopP as in relative clauses with a specific noun phrase. Contrast, for example, the factive (26a) to the relative (26b).

(26) a.*Àgásá **lɔ́** **lɛ́** [ɖě mí wlé] **lɔ́** **lɛ́** vɛ́ ná Kòfí
crab Det Num that$_{[Rel]}$ 1pl catch Det Num hurt for Kofi
b. Àgásá [ɖě mí wlé] **lɔ́** **lɛ́** nyɔ́n ná Kòfí
crab that$_{[Rel]}$ 1pl catch Det Num good for Kofi
'The aforementioned crabs that we caught are good for Kofi' [i.e. he will like them].

Factive constructions and relative clauses are parallel, however, with regard to the relative C. The latter has strong relative [$wh_{A/E}$] features expressed by *ɖě* that must be checked before spell-out. In factive constructions, this requirement is met by raising the event DP in spec,CP, as in (27b).

[8] In this regard, the interested reader is referred to Aboh (2004a) where it is shown that the focus feature is distributed over arguments, adjuncts, and verbs. These categories are attracted to the focus phrase (spec,FocP, or Foc) to check their focus features.

(27) a. Àgásá lɔ́ lɛ́ [ɖě mí wlé] vɛ́ ná Kòfí
crab Det Num that$_{[Rel]}$ 1pl catch hurt for Kofi
'The fact that we caught the crabs hurt Kofi'
b. [$_{CP}$ [$_{DP}$ àgásá lɔ́ lɛ́]$_i$ [$_{C°}$ ɖě [$_{IP}$ mí wlé t$_i$]]]…

Even though the proposed analysis suggests that relative clauses and factive clauses share a common part (i.e. the CP clause) it does not imply that factive clauses are headless relatives. In the representations (18) and (27b), for instance, both the relative clause and the factive clause have a relevant DP (i.e. DP$_{[whA]}$/DP$_{[whE]}$) in spec,CP. These constructions do differ structurally because relatives involve an outer DP layer, but factives do not. This analysis is compatible with the fact that factives are more restricted in their distribution than relatives. The former mainly occur in subject positions while the latter occur in virtually any argument position. In this regard, the ungrammatical sentence (28a) indicates that, like CP-clauses (28b) but unlike relatives (28c), factives cannot occur in the pre-verbal object position in Gungbe progressives.[9]

(28) a.*Ùn tò [àgásá lɔ́ lɛ́ ɖě mí wlé] jrɛ́ ná Kòfî
1sg Prog crab Det Num that$_{[Rel]}$ 1pl catch report for Kofi
'I'm telling Kofi about the fact that we caught the crabs'
b.*Ùn tò [ɖɔ̀ mí wlé àgásá lɔ́ lɛ́] ɖɔ̀ ná Kòfî
1sg Prog that$_{[Comp]}$ 1pl catch crab Det Num tell for Kofi
'I'm telling Kofi that we caught the crabs'
c. Ùn tò [àgásá ɖě mí wlé] lɔ́ lɛ́ jrɛ́ ná Kòfî
1sg Prog crab that$_{[Rel]}$ 1pl catch Det Num report for Kofi
'I'm telling Kofi about the aforementioned crabs that we caught'

The proposed analysis is also compatible with the factive constructions in (29a), whereby the event head (i.e. the verb) left adjoins to the relative complementizer leaving a copy in the IP-internal position. Examples (29b-c) indicate that the event head and the fronted factive noun exclude each other. Accordingly, the factive DP and the event head V fulfill the same requirement in checking the strong [wh$_E$] features of the relative complementizer.

[9] See Aboh (2004a) and references cited there for the discussion of CP-clauses in Gbe.

(29) a. Wlé [ɖě mí wlé àgásá ɖàxó lɔ́ lɛ́] vɛ́ ná Kòfí
catch that$_{[Rel]}$ 1pl catch crab big Det Num hurt for Kofi
'The fact that we CAUGHT the [aforementioned] big crabs hurt Kofi'

b.*àgásá ɖàxó wlé [ɖě mí wlé] lɔ́ lɛ́ vɛ́ ná Kòfí
crab big catch that$_{[Rel]}$ 1pl catch Det Num hurt for Kofi

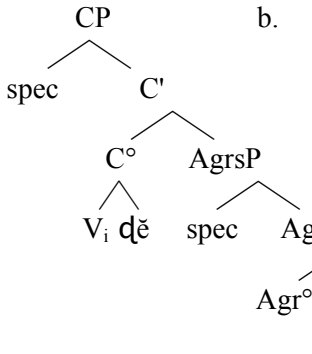
c.*àgásá ɖàxó lɔ́ lɛ́ wlé [ɖě mí wlé] vɛ́ ná Kòfí
crab big Det Num catch that$_{[Rel]}$ 1pl catch hurt for Kòfí

The copy strategy adopted in factive constructions (e.g. 29a) is reminiscent of the copy strategy observed in verb focus constructions whereby the fronted verb leaves a copy IP-internally, as illustrated in (30).

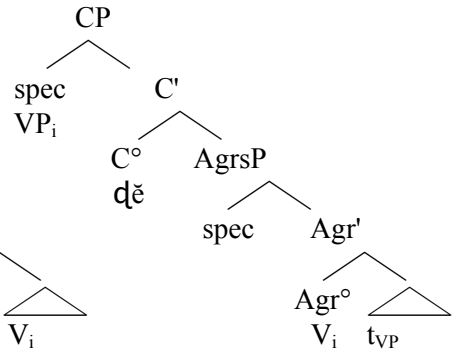
(30) Ùn sè ɖɔ̀ xɔ̀ (wè) Kòfí xɔ̀ àgásá lɔ́
1sg hear that buy Foc Kofi buy crab Det
'I heard that Kofi BOUGHT the [aforementioned] crab'

Under the analysis developed here, and building on the similarities between factives involving event head fronting (29a) and focused verbs constructions, where the fronted verb leaves a copy IP-internally (30), two scenarios arise, as described in (31a) and (31b). In (31a) the factive verb adjoins to the relative complementizer, leaving a copy IP-internally. In (31b), however, a remnant VP moves to spec,CP that is spelt out in concord with the lower V, as indicated by the index. Note that, in (31b), the lower V does not spell out the trace of the fronted VP, but represents the lexical verb that has moved to some intermediate INFL position (Koopman 1999, Aboh 2004a).

(31) a.



b.



Choosing between these two options requires a discussion that goes beyond the scope of this paper and I leave the matter for further research. Note, however, that these issues could not have been raised if we were to adopt the

adjunction analysis to relative/factive clauses. I now move to the Gungbe equivalents of relative and factive clauses in Romance and Germanic.

3.3. *Relative and factive clauses in Romance and Germanic*

Building on the discussion in Gbe, I propose that in languages, which do not manifest predicate fronting (i.e. CP-movement to spec,TopP), specificity within D is checked thanks to the presence of an operator in spec,TopP within the outer DP. Consider, for instance, the English and French relative clauses in (32) where the head noun is interpreted as specific (Aboh 2002). In (32a) spec,TopP is filled by a null operator unlike (32b) where the demonstrative *cette* is moved to spec,TopP to check the specificity feature (see Bianchi 1995, De Vries 2002 and references cited there for discussion on relative clauses in Romance and Germanic).

(32) a. $[_{DP}[_{TopP}\ Op_i\ [_{D^\circ}$ the $[_{NumP}\ [_{CP}$ table$_i$ $[_{C^\circ}$ that $[_{IP}$ I bought t$_i$]]]]]]]…
b. $[_{DP}[_{TopP}$ Cette$_i$ $[_{D^\circ}[_{NumP}[_{CP}\ [t_i$ table] $[_{C^\circ}$ que $[_{IP}$ j'ai achetée t$_i$]]]]]]]…

It appears from this analysis that the Gungbe-type languages and the French- or English-type languages differ because CP-to-spec,TopP movement must apply before spell-out in the former but not in the latter.

Assuming this is the right characterization, I propose that in Romance and Germanic factives where an event DP does not raise to spec,CP, the factive reading can only be achieved by inserting a factive expletive DP (i.e. *the fact*) in spec,CP that binds the event head. This factive expletive noun can also be null as shown in examples (33a–b).

(33) a. <u>The fact</u> that John came worried me/that John came worried me
b. <u>Le fait</u> que Jean soit venu m'embête/que Jean soit venu m'embête

The sentences under (33) are represented in (34).

(34) a. $[_{CP}\ [_{DP}$ the fact/ ∅] $[_{C^\circ}$ that $[_{IP}$ John came]]] worried me a lot
b. $[_{CP}$ le fait/Ø $[_{C^\circ}$ que [Jean soit venu]]] m'embête

Note that if the sequences in (33) were true relative clauses, we would expect the relative clause reading. But the absence of the outer D-system above C precludes such reading here, also restricting as such the distribution of these factives. As mentioned previously, factive constructions tend to occur in subject position mainly, and when they occur in object position, they are selected by certain verbs only. It is not clear yet whether this

selection restriction is semantic or syntactic, but no such constraint seems to hold for relatives. In example (35a), the factive sentence is felicitous in subject position, but not in object position (35b). The relative sentences under (35c–d) are not sensitive to this effect.

(35) a. Le fait que Jean soit venu à la réunion décrit bien l'ambiance au sein du groupe.
b.*Je décrit le fait que Jean soit venu à la réunion
c. Je décrit l'homme qui est venu à la réunion
d. L'homme qui est venu à la réunion décrit la situation au sein du groupe

That the expletive DP can be null in both English and French lends further support to this analysis. Indeed, headed relatives introduced by a null head are not allowed in these languages, hence the ungrammatical example (36).

(36) *[$_{DP}$ [$_{CP}$ Ø$_{[les\ crabes]}$ [$_{C°}$ que [$_{IP}$ nous avons acheté t$_{[les\ crabes]}$]]]]

4. Some further implications

The proposed analysis has further implications as to the characterization of the so-called sentence complements to nouns, and to the contexts of *that*-deletion.

4.1. *Expletive factive versus referential factive*

Assuming the proposed analysis is on the right track, I conjecture that the so-called clausal complements to nouns as in (37a) are actually types of factive constructions, where a referential DP merges in spec,CP as in (37b).

(37) a. The rumour that Jacques Chirac likes beer annoys the party people
b. [$_{CP}$ The rumour [$_{C°}$ that [Jacques Chirac likes beer]]] annoys the party people

For the discussion sake, I assume that the factive DP first merges in spec,CP similarly to the factive expletive. Be it so, one can make the following observations:

1. Referential factives differ from expletive factives because the former can be modified unlike the latter.

(38) a. The persistent rumour that Jacques Chirac likes beer annoys the party people
b.*The persistent fact that Jacques Chirac likes beer annoys the party people

2. Referential factives allow equative constructions unlike expletive factives.

(39) a. The rumour that Jacques Chirac likes beer is a well-known rumour around here
b.*The fact that Jacques Chirac likes beer is a well-known fact around here

3. Referential factives allow inversion unlike expletive factives.

(40) a. Jacques Chirac loves beer says the rumour
b.*Jacques Chirac loves beer says the fact

Under the assumption that referential factives and expletive factives share the same underlying structure, these facts indicate that a refinement is needed to account for their differences. A possibility that comes to mind is that the observed differences could be linked to the (traditional) contrast between expletive DPs and referential DPs. I leave this matter for further research.

4.2. *A unified approach for that-deletion*

An interesting, though seemingly unrelated, aspect of this analysis is that it may provide the basis for a unified account for *that*-deletion as illustrated by the sentences under (41).

(41) a. The man that/Ø I saw in Amsterdam
b. I propose that/Ø you come to Amsterdam
c. The fact that/*Ø John came to Amsterdam

Under the adjunction analysis to relative clauses and the traditional analysis of (41c) as sentence complement to the noun *fact*, two different scenarios need to be imagined to account for *that*-deletion in certain adjunct structures as in (41a), and in certain complementation contexts only (41b-c), but not others (41c). In terms of the approach developed here, however, the generalization seems to be the following:

(42) *That*-deletion is licensed in complementation context, only.

Under the analysis of factive constructions as truncated relatives, the impossibility of deletion in (41c) is straightforward. What we have here is an instance of spec-head relationship where the factive constituent *the fact* is in spec,CP. Accordingly, *that*-deletion is not allowed. Under Pesetsky's (1996) analysis, where the null C is an affix that must raise to the matrix V (i.e. C-to-V movement) this would mean that, in relative clauses (41a), null C is possible due to C-to-D movement. Put differently, the embedded C incorporates into the D of the outer DP-layer. Such incorporation is obviously impossible in factive and related clauses where there is no outer DP layer. While supporting Pesetsky's analysis to some extent, (42) has an advantage that it unifies *that*-deletion phenomena in relative clauses and embedded clauses without invoking affix hopping (i.e. an instantiation of Morphological Merger or PF Merger) as recently proposed by Boskovic & Lasnik (2003). Obviously, more needs to be said before we reach a full understanding of *that*-deletion in English. However, the analysis outlined here appears very promising and I hope to return to these facts in future work.

5. Conclusion

On the assumption that the D-system represents the nominal left periphery that provides room for topic and focus nominals, this paper shows that the distribution of the specificity marker (i.e. the nominal topic marker) and the number marker in Gungbe provides evidence for Kayne's (1994) analysis of relative clauses where D selects a CP clause as its complement and the relative noun raises to spec,CP. Under this approach, the Gungbe sequence Noun-[relative clause]-Det-Num results from DP-raising to spec,CP followed by movement of the relative clause (i.e. CP) to spec,NumP and spec,TopP, in order to check the features [number] and [specific], respectively.

In terms of the proposed analysis the semantic contrast between the relative clauses and the factive constructions is accounted for by suggesting that factive clauses do not project the outer D-system typical of relative clauses. Factive constructions, instead, are simple CP-clauses, where spec,CP hosts an event DP that has moved out of the embedded IP. This would mean that factive constructions are truncated relative clauses because the top projections DP and NumP have been pilled off. This analysis extends to Romance and Germanic, where it is argued that the factive reading is achieved by inserting an expletive factive DP in spec,CP. In this regard, it is argued that the so-called sentence complement to nouns represent types of

factive clauses where a referential DP merges in spec,CP. The proposed analysis has far reaching consequences as to the analysis of sentence complement to nouns as sub-types of factive constructions. Finally, this analysis suggests that *that*-deletion occurs in complementation context only.

References

Abney, Paul S. 1987. *The English Noun Phrase in its Sentential Aspect.* Ph.D. dissertation, MIT.

Aboh, O. Enoch. 2002. "Morphosyntaxe de la péripherie gauche nominale." In Anne Zribi-Hertz and Anne Daladier (eds.), *Recherches linguistiques de Vincennes* 31:9-26, La syntaxe de la definitude 2002. Paris, PUV.

Aboh, O. Enoch. 2004a. *The Morphosyntax of complement-head sequences* : *Clause structure and word order patterns in Kwa*. New York/Oxford, Oxford University Press.

Aboh, O. Enoch. 2004b. "Topic and Focus within D". In L Cornips & J. Doetjes (eds.), *Linguistics in The Netherlands* 21:1-12.

Bernstein, Judy. 1997. "Demonstratives and Reinforcers in Romance and Germanic languages." *Lingua* 102:87-113.

Bernstein, Judy. 2001a. "Focusing the 'Right' Way in Romance Determiner Phrases." *Probus* 13.1:1-29.

Bernstein, Judy. 2001b. "The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain." In M. Baltin and C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*. Oxford, Blackwell.

Bianchi, Valentina. 1995. *Consequences of Antisymmetry for the Syntax of Headed Relative Clauses.* Ph.D. dissertation. Pisa, Scuola Normale Superiore.

Boškovic, Željko and Howard Lasnik. 2003. "On the Distribution of Null Complementizers". *Linguistic Inquiry* 34:527-546.

Campbell, Richard. 1996. "Specificity Operators in Spec,DP." *Studia Linguistica* 2:161-188.

Carstens, M. Vicky. 2000. "Concord in Minimalist Theory." *Linguistic Inquiry* 31:319-355.

Collins, Chris. 1994. *The Factive Construction in Kwa*. Ms., Cornell University.

De Vries, Mark. 2002. *The Syntax of Relativization.* LOT 53. Utrecht, The Netherlands.

Enç, Mürvet. 1991. "The Semantics of Specificity." *Linguistic Inquiry* 22:1-26.

Giusti, Giuliana. 1996. "Is there a FocusP and a TopicP in the Noun Phrase structure?" *University of Venice Working Papers in Linguistics* 6:105-128.

Giusti, Giuliana. 1997. "The Categorial Status of Determiners." In Liliane Haegeman (ed.) *The New Comparative Syntax*. London, Longman Linguistics Library.

Ihsane, Tabea, and Genoveva Puskás. 2001. "Specific is not Definite". In *Generative Grammar in Geneva* 2:39-54

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge MA, MIT Press.

Koopman, Hilda. 1999. "Unifying Predicate Cleft Constructions". In *The Syntax of Specifiers and Heads.* London and New York, Routledge.

Longobardi, Giuseppe. 1994. "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form." *Linguistic Inquiry* 25:609-665.

Longobardi, Giuseppe. 2001. "The Structure of DPs: Some Principles, Parameters and Problems." In M. Baltin and C. Collins (eds.) *Handbook of Syntactic Theory*. Oxford, Blackwell.

Newmeyer, Frederick J. 2003. *Peripheral Grammatical Relations in English : Theoretical Implications of Non-DP Subjects and Objects.* Paper presented at the Syntactic Functions–Focus on the Periphery. Helsinki, 2003.

Panagiotidis, Phoevos. 2000. "Demonstrative Determiners and Operators: the Case of Greek". *Lingua* 110:717-742.

Pesetsky, David. 1987. "Wh-in-situ: Movement and Unselective Binding." In Eric Reuland, and Alice ter Meulen (eds.) *The Representation of* (*In*)*definiteness*. Cambridge MA, MIT Press.

Pesetsky, David. 1996. *Zero syntax.* Cambridge MA, MIT Press.

Ritter, Elizabeth. 1995. "On syntactic Category of Pronouns and Agreement." *Natural Language and Linguistic Theory* 13:405-443.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery." In Liliane Haegeman, (ed.) *Elements of Grammar*. Dordrecht, Kluwer.

Siloni, Tal. 1997. "Event Nominals and the Construct State." In L. Haegeman (ed.) *The New Comparative Syntax*. London, Longman Linguistics Library.

Szabolcsi, Anna. 1994. "The Noun Phrase." In Ferenc Kiefer and Katalin E. Kiss (eds.) *Syntax and Semantics. The Syntactic Structure of Hungarian,* 27:179-274.

Toribio, Almeida J.. 1992. "Proper Government in Spanish Subject Relativization." *Probus* 4:291-304.

Chinese relative clauses: restrictive, descriptive or appositive?

Francesca Del Gobbo
University of Venice

1. Introduction

This paper addresses the issue of the distinction between restrictive and 'descriptive' relative clauses for Mandarin Chinese. While it is well-known what a restrictive relative clause is, the term 'descriptive' is more elusive, and has only been used to illustrate a certain kind of relative clause in Mandarin Chinese. In particular, it appears that with this term various authors seek to identify a variety of Chinese relative clauses that share some properties with the so-called appositive relative clauses of languages like English and Italian.

In order to assess what a Chinese 'descriptive' relative clause is, in this paper I first show that this particular type of relative clause is not equivalent to an appositive relative clause. More specifically, I will use some traditional diagnostics to distinguish restrictive from appositive relative clauses to show that Chinese relative clauses are never appositive.

In section 3 I provide evidence, following work by Larson and Takahashi (2002), to show that Chinese 'descriptive' relative clauses can be better analyzed as generic or i-level modifiers, while Chinese restrictive relative clauses can be classified as deictic or s-level modifiers.

In section 4 I will prove that even when they modify proper names and/or pronouns, Chinese relative clauses behave as restrictives. This is not so surprising once we acknowledge that appositive adjectives and restrictive relative clauses are more similar to one another than they are to appositive relative clauses (section 5). In Chinese, we can either have relative clauses that behave as appositive adjectives (when they modify proper names and/or pronouns) or relative clauses that behave as restrictive ones (in all other

cases). I claim that the impossibility for Chinese relative clauses to be truly appositive is ultimately due to the nature of appositive relative clauses, which, according to the hypothesis put forth in Del Gobbo (2003), are independent sentences (propositions of type t) and instances of E-type anaphora.

2. Chinese relatives do not behave as appositive relatives

According to Chao (1968) and Hashimoto (1971), among others,[1] a relative clause in Chinese is 'descriptive' if it follows a demonstrative, but it is restrictive if it precedes it:

(1) na-ge [dai yanjing de] nanhai
that-CL wear glasses DE boy
'that boy, who wears glasses'

(2) [dai yanjing de] na-ge nanhai
wear glasses DE that-CL boy
'the boy that wears glasses'

Notice that the translations I used in (1) and (2) are indicative of the assumption made when the above mentioned authors discuss the so-called 'descriptive' relative clauses. Even though they do not specifically mention it, it is clear by the translations they use that by 'descriptive' they mean 'appositive'.[2]

[1] Huang's (1982) account of the facts in (1)-(2) is in terms of the scope of modification: if the relative clause is in the scope of the demonstrative as in (1), the demonstrative is deictic, and it fixes the reference of the 'head' of the relative clause. The relative clause is then used non-restrictively. But when the demonstrative is in the scope of the relative clause as in (2), it is used 'anaphorically' onto the relative clause. And it is the relative clause, now restrictive, which contributes in determining the reference of the 'head' noun. A terminological clarification is due at this point. In the literature on Chinese, it is not always clear if a non-restrictive relative clause is an appositive relative clause. Huang (1982) explicitly says that his use of the term 'non-restrictive' is different from the use of the same term to describe English relative clauses. He uses the term 'non-restrictive' to mean that the relative does not specify the reference of a preceding demonstrative.

[2] The actual examples used by Chao (1968: 286) are:

(i) Na-wei dai yanjing de xiansheng shi shei?
that-CL wear glasses DE gentleman is who
'Who is that gentleman (who incidentally is) wearing glasses?'

But, as I already show elsewhere (Del Gobbo 2001, 2003), Chinese relatives consistently have the behaviour of restrictive relative clauses. I provide some of the most compelling evidence below.

Safir (1986) observes that no quantifier in the matrix clause can have scope over a pronoun in the appositive clause. In the following pair of examples, *every Christian* can bind into the restrictive in (3a), but it cannot bind into the appositive in (3b):

(3) a. [Every Christian]$_i$ forgives a man who harms him$_i$.
b.*[Every Christian]$_i$ forgives John, who harms him$_i$.

In Chinese, surprisingly, this generalization does not hold. It is in fact possible to bind inside the relative, regardless of the order between the relative and the demonstrative:

(4) a. [Mei-ge xuesheng]$_i$ dou yuanliang na-xie [$_{RC}$ cengjing shanghai-guo tamen$_i$ de] ren.
every-CL student all forgives those formerly insult GUO them DE people
b. [Mei-ge xuesheng]$_i$ dou yuanliang [$_{RC}$ cengjing shanghai-guo tamen$_i$ de] na-xie ren.
every-CL student all forgives formerly insult GUO them DE those people
'Every student forgives those who insulted him.'

Emonds (1979), citing Ogle (1974), notices that certain adverbs that in general appear only in main clauses also appear in appositive relatives. Thus, *frankly* can appear inside the appositive relative in (5a), but not inside the restrictive relative in (5b):

(ii) Dai yanjing de na-wei xiansheng shi shei?
wear glasses DE that-CL gentleman is who
'Who is the gentleman who is wearing glasses (not the one who is not wearing glasses)?

Chao (1968: 286) also adds that "... if a contrasting stress is placed on a modifier, it is used restrictively, so that if *dai yanjing* in (i) is constrastively stressed, the sentence will have the same restrictive sense as in (ii)". I will get back to the restrictive reading of modifiers within the scope of the demonstrative in section 3. As for the contrastive stress, in this paper I will only consider sentences that do not contain it, leaving this interesting issue for future research.

(5) a. The boys, who have frankly lost their case, should give up.
b.*The boys that have frankly lost their case should give up.

In Chinese, the insertion of a sentential adverb of modification yields ungrammaticality, in both cases of possible orders:

(6) a.*Na-xie [$_{RC}$ shunbianshuo kaoshi shibai de] nanhai
those incidentally exam fail DE boys
zhuandao lingwai yi-suo daxue qu-le.
transfer other one-CL university go-LE
b.*[$_{RC}$ Shunbianshuo kaoshi shibai de] na-xie nanhai
incidentally exam fail DE those boys
zhuandao lingwai yi-suo daxue qu-le.
transfer other one-CL university go-LE
*'The boys that incidentally failed the exam transferred to another university.'

Ross (1967) maintains that quantified noun phrases cannot usually serve as antecedents of an appositive relative clause:

(7) a. Every student that wears socks is a swinger.
b.*Every student, who wears socks, is a swinger.

In Chinese no difference arises if the quantifier precedes or follows the relative clause, as pointed out also by Lin (1997):

(8) a. Mei-ge [$_{RC}$ chuan wazi de] xuesheng dou shi tiaowude.
every-CL wear socks DE student all be dancers
b. [$_{RC}$ Chuan wazi de] mei-ge xuesheng dou shi tiaowude.
wear socks DE every-CL student all are dancers
'Every student that wears socks is a dancer.'

Chierchia and McConnell-Ginet (1990) argue convincingly that appositives are not presuppositions, but backgrounded assertions. More precisely, they maintain that the content of an appositive relative clause is a backgrounded component of what is being asserted, not of what is being presupposed. They show this by reasoning that the truth of an appositive relative clause is not taken for granted. In the case of a presupposition, instead, its assumed truth is a precondition for the felicitous utterance of the sentence and places a kind of constraint on discourse contexts that admit the sentence for interpretation.

Consider the following examples (slightly modified from Chierchia and McConnell-Ginet 1990):

(9) a. That woman, who lost her luggage on the flight from Ithaca to New York, likes to travel by train.
b. That woman, who lost her luggage on the flight from Ithaca to New York, doesn't like to travel by train.
c. Does that woman, who lost her luggage on the flight from Ithaca to New York, like to travel by train?
d. If that woman, who lost her luggage on the flight from Ithaca to New York, likes to travel by train, she probably flies infrequently.
e. That woman lost her luggage on the flight from Ithaca to New York.

As Chierchia and McConnell-Ginet (1990) point out, a sentence S presupposes p if S implies p and further implies that p is somehow already part of the background against which S is considered. In order for S to imply p, p needs to stay attached to S not only when it is asserted, but also when it is denied, questioned, or offered as a hypothetical assumption. This is what the P-family test does:

(10) a. S
b. It is not the case that S.
c. Is it the case that S?
d. If S, then S'.

We can see from (9) that the appositive relative passes the implication tests: each of (9a-d) implies (9e), the content of the relative clause. But the appositive is not presupposed. In (9a) our assertion is articulated in two parts: a main assertion in the foreground and a secondary one in the background. To show that this is the case, Chierchia and McConnell-Ginet (1990) embed (9a) in a context – provided in (11) – that does not take for granted the truth of (9e):

(11) Let me tell you something about Jill Jensen, a woman I met while flying from Ithaca to New York last week.

If we utter (9a) after (11), the result is felicitous. But this is not the case if after (11) we utter a sentence with the restrictive counterpart of the appositive in (9a):

(12) The woman that lost her luggage on the flight from Ithaca to New York was pretty upset.

The infelicitous result of uttering (12) after (11) is due to the fact that, because of the definite determiner in (12), the content of the restrictive relative clause is presupposed, hence taken for granted. But nothing in the context provided in (11) tells us that the woman lost her luggage. Thus, a restrictive relative clause with a definite determiner cannot be uttered felicitously after a sentence that does not take its truth for granted.

In Chinese, we find that the utterance of a relative clause modifying a definite nominal in a context that does not take the truth of the relative for granted does not yield a felicitous result. Consider the following group of sentences:

(13) a. Na-ge [$_{RC}$ zai cong Beijing dao NiuYue lutu zhong diu-le
that-CL in from Beijing to New York travel middle lost-LE
xingli de] nuren xihuan zuo huoche.
luggage DE woman like sit train
'That woman who lost her luggage on the flight from Beijing to New York likes to travel by train.'

b. Na-ge [$_{RC}$ zai cong Beijing dao NiuYue lutu zhong diu-le
that-CL in from Beijing to New York travel middle lost-LE
xingli de] nuren bu xihuan zuo huoche.
luggage DE woman not like sit train
'That woman who lost her luggage on the flight from Beijing to New York doesn't likes to travel by train.'

c. Na-ge [$_{RC}$ zai cong Beijing dao NiuYue lutu zhong diu-le
that-CL in from Beijing to New York travel middle lost-LE
xingli de] nuren xi bu xihuan zuo huoche?
luggage DE woman like not like sit train
'Does that woman who lost her luggage on the flight from Beijing to New York like to travel by train?'

d. Ruguo na-ge [$_{RC}$ zai cong Beijing dao NiuYue lutu zhong diu-le
if that-CL in from Beijing to New York travel middle lost-LE
xingli] de nuren xihuan zuo huoche, ta dagai jiu bu chang
luggage DE woman like sit train she probably just not often
zuo feiji.
sit plane
'If that woman who lost her luggage on the flight from Beijing to New York likes to travel by train, she probably flies infrequently.'

e. Na-ge nuren zai cong Beijing dao NiuYue lutu zhong
that-CL woman in from Beijing to New York travel middle
diu-le ta de xingli.
lost-LE her DE luggage
'That woman lost her luggage on the flight from Beijing to New York.'

Not surprisingly, the relative clause passes the P-family test. But this does not prove that the relative is appositive, since both appositives and restrictives in the scope of a definite determiner pass the P-family test. What we need to do is embed (13a) in a context that does not take the truth of the relative for granted. I provide such a context in (14):

(14) Rang wo gaosu ni guanyu Xiao Yu de yi-xie shi.
let me tell you about Xiao Yu DE one-CL thing
Wo shangzhou zai cong Beijing dao NiuYue de lutu zhong
I last-week on from Beijing to New York DE travel during
pengjian ta.
meet her
'Let me tell you something about Xiao Yu, a woman I met while flying from Beijing to New York last week.'

Interestingly, if (14) is followed by the sentence in (13a), the result is awkward. I take this to show that the relative clause in the sentences in (13) is not appositive: if it were such, it would only provide backgrounded information and consequently should be able to be felicitously uttered after the context in (14).

Summarizing, all the tests above indicate that, regardless of the order of the relative clause with respect to the determiner, relative clauses in Chinese have the syntactic and semantic properties of restrictive relative clauses.

3. Deictic and generic reading of Chinese relative clauses

In the previous section, I showed that regardless of the order of the determiner and the relative clause, Chinese relative clauses are consistently restrictive. In this section I propose to analyze the so-called 'descriptive' relative clauses in Chinese as generic or i-level modifiers, in the sense of Larson and Takahashi (2002).

3.1. *Deictic and generic modifiers*

Bolinger (1967) observes that the pairs in (15) show a meaning difference:

(15) a. the stars visible (include Capella, Betelguese, and Sirius)
the visible stars
b. the rivers navigable (include the Nile, the Amazon and the Ganges)
the navigable rivers
c. the individual responsible (were contacted)
the responsible individuals
d. the jewels stolen (were on the table)
the stolen jewels

Postnominal adjectives attribute a property temporally or episodically, while prenominal adjectives can attribute a property characteristically or intrinsically. According to Svenonius (1994), this is the s-level/i-level distinction (respectively):

(16) a. The stars visible include Capella. False, when the sky is overcast.
b. The visible stars include Capella. True, even when the sky is overcast.

The interesting observation made by Larson and Takahashi (2002) is that the i-level/s-level distinction does not have a one-to-one correspondance with the prenominal or postnominal position of the adjective:

(17) a. The visible stars visible include Capella.
b. The visible visible stars include Capella.

(18) a. The nonvisible visible stars include Capella.
b.#The visible nonvisible stars include Capella.

The postnominal adjective visible in (17a) has an s-level reading, but this is also the case for the first of the prenominal adjectives in (17b). That this slot is dedicated to s-level modification is made clear by the pair in (18): while the example in (18a) is coherent, the one in (18b) is odd, as it is awkward to assign the temporary property of being visible to stars that are intrinsically nonvisible.

Larson and Takahashi (2002) also observe that time modifiers in prenominal positions are ambiguous. They can have both a deictic and a generic reading. When they are doubled in prenominal position, generic modifiers occur closer to N than deictic ones:

(19) a. (Tuesday lectures are usually interesting and Thursday lectures boring but:)
the Thursday Thursday lecture (was interesting this week)
b. the Wednesday Thursday lecture (was packed)
(said of a regular Thursday lecture moved to Wednesday during one week)

The generalization that Larson and Takahashi (2002) draw from these examples is that i-level adjectives occur closer to the noun than s-level ones, and that generic time modifiers occur closer to the noun than deictic ones. In Larson (2000) these facts are drawn together under the idea that the nominal, composed of NP and DP, comprises two distinct domains of modification. The outer modifier vs. inner modifier contrasts reflect DP-modification vs. NP-modification (respectively):

(20) $[_{DP}\ D\ \beta\ [_{NP}\ \alpha\ N\ ]\ \beta]$ (β = DP-modifier; α = NP-modifer)

The fact that the inner domain expresses generic/i-level predicates is captured by Larson and Takahashi (2002) by proposing that NP always contains a generic quantifier Γ whose scope is limited to NP:[3]

(21) $[_{DP}\ D\ \ \beta\ [_{NP}\ \Gamma e\ [\ \alpha\ N]\ ]\ \beta\ ]$

This means that modifiers inside NP (α) will have potential generic/i-level readings, while modifiers outside NP (β) will be outside the scope of the generic quantifier and will not get i-level/generic readings.

Larson and Takahashi (2002) also observe that in Japanese and Korean prenominal relatives show ordering restrictions resembling those found in English prenominal adjectives. Relative clauses expressing i-level properties and relatives expressing s-level properties order freely among themselves; but when the two types of relatives combine, i-level ones must occur closer to the noun than s-level ones. Therefore, they propose that prenominal relative clauses in Japanese and Korean may attach to either NP or DP, like prenominal adjectives and time modifiers in English:

[3] The idea stems from Chierchia's (1995) proposal according to which generic verbs, predicate nominals, and i-level adjectives are all bound by a generic quantifier Γ ranging over eventualities.

(22) $[_{DP}$ RC $[_{NP}$ Ге $[$ RC N]]]
s-level i-level

This accounts for the ordering restrictions they observed and for the fact that i-level relatives have the semantics of generics.

3.2. *Chinese 'descriptive' relatives as generic modifiers*

Since in Chinese nominal phrases are head-final, the structure in (21) translates into the following structure:

(23) $[_{DP}$ β D β $[_{NP}$ Ге [α N]]]

Assuming (23) and a left-adjunction structure for Chinese relatives (see Aoun and Li 2003, Del Gobbo 2003), we make two predictions. The first prediction is that Chinese relatives preceding the demonstrative in D have a deictic or s-level meaning, and Chinese relatives following the demonstrative have either a deictic/s-level or a generic/i-level meaning. The second prediction is that in Chinese we should find the same ordering restrictions found in Korean and Japanese.

The first prediction is confirmed by the semantic difference that the native speakers attribute to the example in (1) and (2), repeated below with different translations as (24) and (25):[4]

(24) na-ge [dai yanjing de] nanhai
that-CL wear glasses DE boy
'that boy who wears glasses' (preferred reading: generic/i-level)

(25) [dai yanjing de] na-ge nanhai
wear glasses DE that-CL boy
'the boy that wears glasses' (only reading: deictic/s-level)

In the example (24), the relative clause has a preferred generic or i-level reading, even though it could also have a deictic/s-level one.[5] In the example

[4] An anonymous reviewer points out that in Italian, as in other Romance and Germanic languages, the demonstrative form *questo*, 'this', has a stronger deictic value than the form *quello*, 'that'. In Chinese, it seems that if we replace *na-ge*, 'that' with the corresponding Chinese form for *this,* no semantic difference is attested, as far as the interpretation of the relative is concerned.

in (25), the relative clause instead only has a deictic reading. This is also reinforced by the fact that the demonstrative in (25) – but not the one in (24) – can lose its deictic meaning and be semantically bleached to the meaning of a definite determiner.[6] Finally, the relative in (24) doesn't have an appositive reading, but a generic/i-level one: this is the 'descriptive' reading advocated in the literature.[7]

As for the second prediction, in Chinese as well we do indeed find the same ordering restrictions noticed in Japanese and Korean. The following examples show how i-level relative clauses and s-level relative clauses order freely among themselves:

i-level RCs

(26) [$_{RC}$ Hui shuo Yidaliyu de] [$_{RC}$ xihuan qu yinyuehui de] ren shi
can speak Italian DE like go concerts DE person is
Zhangsan.
Zhangsan
'The person who speaks Italian who likes to go to concerts is Zhangsan.'

(27) [$_{RC}$ Xihuan qu yinyuehui de] [$_{RC}$ hui shuo Yidaliyu de] ren shi
like go concerts DE can speak Italian DE person is
Zhangsan.
Zhangsan
'The person who likes to go to concerts who speaks Italian is Zhangsan.'

s-level RCs

(28) [$_{RC}$ Cong Yidali huilai de] [$_{RC}$ wo zuotian kanjian de] ren shi Lisi.
from Italy come-back DE I yesterday meet DE person be Lisi
'The person who came back from Italy who I met yesterday is Lisi.'

---

[5] Recall that according to Chao (1968: 286), if a constrasting stress is placed on a modifier, it is used restrictively. See also footnote 2.

[6] Hashimoto (1966: 25) maintains that when the demonstrative follows the relative clause, it doesn't have any deictic value, but it is only anaphoric. On the contrary, when the demonstrative precedes the relative, it has deictic value.

[7] As a matter of fact, Chao (1968: 287) himself had already noticed this fact, as he observes: "A somewhat weaker effect of the order of demonstratives has to do with permanent as against temporary characteristics in the modifier."

(29) [$_{RC}$ Wo zuotian kanjian de] [$_{RC}$ cong Yidali huilai de] ren shi Lisi.
I yesterday meet DE from Italy come-back DE person be Lisi
'The person who I met yesterday who came back from Italy is Lisi.'

Interestingly, when the two types of relatives combine, i-level ones need to occur closer to the 'head' noun than s-level ones:

<u>i-level + s-level RCs</u>

(30) [$_{RC}$ Wo zuotian kanjian de] [$_{RC}$ xihuan qu yinyuehui de] ren shi
I yesterday meet DE like go concerts DE person is
Zhangsan.
Zhangsan
'The person I met yesterday who likes to go to concerts is Zhangsan.'

(31) * [$_{RC}$ Xihuan qu yinyuehui de] [$_{RC}$ wo zuotian kanjian de] ren shi Lisi.
like go concerts DE I yesterday meet DE person is Lisi
'The person who likes to go to concerts who I met yesterday is Lisi.'

Also, if a demonstrative is present, only s-level relatives can precede it, while between the demonstrative and the 'head' noun, the only possible order is again s-level preceding i-level:

<u>s-level + i-level RCs + demonstrative</u>

(32) [$_{RC}$ Zuotian meiyou lai de] na-ge [$_{RC}$ hen xihuan shang ke de]
yesterday not come DE that-CL very like go class DE
xuesheng jiao Zhangsan.
student call Zhangsan
'The student who didn't come yesterday who likes to come to class very much is called Zhangsan.'

(33) * [$_{RC}$ Hen xihuan shang ke de] na-ge [$_{RC}$ zuotian meiyou lai de]
very like go class DE that-CL yesterday not come DE
xuesheng jiao Zhangsan.
student call Zhangsan

(34) Na-ge [$_{RC}$ zuotian meiyou lai de] [$_{RC}$ hen xihuan shang ke de]
that-CL yesterday not come DE very like go class DE
xuesheng jiao Zhangsan.
student call Zhangsan
'The student who didn't come yesterday who likes to come to class very much is called Zhangsan.'

(35) * Na-ge [$_{RC}$ hen xihuan shang ke de] [$_{RC}$ zuotian meiyou lai de]
that-CL very like go class DE yesterday not come DE
xuesheng jiao Zhangsan.
student call Zhangsan

Summarizing, Chinese relative clauses show the same ordering restrictions that we find in English prenominal adjectives and in Japanese and Korean relative clauses. Following Larson and Takahashi (2002), these ordering restrictions can be explained in terms of two layers of modification: the DP-layer, dedicated to s-level or deictic modification, and the NP-layer, for i-level or generic modification. I claim that this exact structure is responsible for the difference in Chinese between the so-called 'descriptive' relative clauses and the restrictive ones. This proposal should finally clarify the confusion in the literature regarding the meaning of term 'descriptive' as referred to Chinese relative clauses: these are generic or i-level modifiers, and not appositive relative clauses.

4. Chinese relatives modifying Proper Names and Pronouns

Even though we have established that the so-called 'descriptive' relative clauses of Chinese are not real appositive relative clauses, the question of the existence of appositives in Chinese remains, as relative clauses in this language are still able to modify pronouns and proper names. In Del Gobbo (2003) and Del Gobbo (2004), I show that in Chinese prenominal relatives modifying proper names and/or pronouns consistently behave like restrictives. I repeat some of the relevant tests here.

We know that these relatives are restrictives because, even though they modify pronouns or proper names, it is possible to bind inside them:

(36) [Mei-ge xuesheng]$_i$ dou xihuan [$_{RC}$ shi ta$_i$ daoshi de] Huang laoshi.
every-CL student all like be he advisor DE Huang professor
lit. 'Every student likes Prof. Huang who is his advisor.'

(37) [Mei-ge xuesheng]$_i$ dou yuanliang (na-ge)[$_{RC}$ cengjing shanghai-guo
every-CL student all forgives that-CL formerly insult-GUO
ta$_i$ de] Lisi.
him DE Lisi
lit. 'Every student forgives Lisi who insulted him.'

(38) [Mei-ge laoshi]$_i$ dou jiaoxun [$_{RC}$ meiyou zunzhong ta$_i$ de] women.
every-CL professor all scold not respect him DE us
lit. 'Every professor scolded us who disrespected him.'

Moreover, the insertion of a sentential adverb of modification yields ungrammaticality:

(39) *[Shunbian-yi-sheng xihuan yinyue de] Zhangsan changchang qu
by-the-way-one-voice like music DE Zhangsan often go
yinyuehui.
concert
lit. 'Zhangsan, who by the way likes music, often goes to concerts.'

Finally, Giorgi (1984), on the basis of the behavior of the long-distance anaphor *proprio*, 'self', in Italian, observes that if the long-distance anaphor is inside an appositive relative clause, it can only be bound by the 'head' of the relative clause. If instead it is inside a restrictive relative clause, it can be bound both by the 'head' of the relative and by the matrix subject:

(40) Gianni$_i$ pensa che Mario$_j$, che t$_j$ ama la propria$_{j/*i}$ moglie,
Gianni thinks that Mario that loves the own wife
sia intelligente.
is smart
'Gianni thinks that Mario, who loves his own wife, is smart.'

In Chinese, the long-distance anaphora ziji, 'self', can be bound both by the 'head' of the relative and by the matrix subject:

(41) Zhangsan$_i$ renwei [$_{RC}$ ai ziji$_{j/i}$ de qizi de] Lisi$_j$ hen congming.
Zhangsan thinks love own DE wife DE Lisi very smart
'Zhangsan thinks that the Lisi who loves his own wife is smart.'

(42) Zhangsan$_i$ renwei [$_{RC}$ ai ziji$_{j/i}$ de qizi de] ta$_j$ hen congming.
Zhangsan think love himself DE wife DE him very smart
'Zhangsan thinks that he who loves his own wife is smart.'

In the examples in (41) and (42), even though the relative clauses modify a proper name and a pronoun, respectively, the relative clauses cannot be appositive, because the anaphor can be bound by the matrix subject.

Summarizing, the tests used in this section provide evidence that even when they modify pronouns or proper names, Chinese relative clauses behave as restrictive relatives and not as appositive ones.

5. Appositive Adjectives vs. Appositive Relative Clauses

The key to solve the puzzle of Chinese relative clauses behaving as restrictive even when modifying proper names and pronouns is linked to Larson and Takahashi's (2002) proposal to consider prenominal relative clauses as more 'adjectival' than postnominal ones. I would like to push this claim further to show that Chinese relatives, being 'adjectival', can be restrictive – i-level/generic or s-level/deictic – and appositive (in the sense that they can modify proper names and pronouns), but cannot be real appositive relative clauses.

I have shown elsewhere (Del Gobbo 2004) that despite their similarities, appositive adjectives are semantically and syntactically different from appositive relative clauses. The evidence I use comes from tests on question/answer pairs, NP-deletion and VP-deletion. I will not repeat those tests here, but I will report the main results: appositive relative clauses always behave as if they are detached from the matrix sentence, in the sense that they do not seem to be part of the main assertion. As a matter of fact, they cannot be repeated in answers, and with both NP-deletion and VP-deletion, their content disappears at LF. In other words, appositive relative clauses seem to be detached from the matrix sentence that contains them, at least at LF, maybe, at the discourse level, where by discourse level I mean a level where not only sentences, but sequences of sentences are represented. Appositive adjectives, instead, pattern with restrictive relative clauses: they can be repeated in answers, and with both NP-deletion and VP-deletion their content is always present at LF.

In order to account for the different behavior between appositive relative clauses on one side, and appositive adjectives and restrictive relative clauses on the other side, in Del Gobbo (2004) I propose the following distinction. Appositive relatives denote propositions (type t) (see Sells 1985, Demirdache 1991, Del Gobbo 2003), hence independent sentences. Syntactically, they are detached from the 'head', at some relevant level. Appositive adjectives denote predicates of individuals (type <e,t>), and they are always adjoined to the noun they modify, through LF. In a sense, they are very similar to restrictive relative clauses, the only difference is really in

the noun they modify: appositive adjectives can modify pronouns and proper names, restrictive relative clauses cannot. Syntactically they are then very similar; semantically, we need to assume a different mechanism of composition. In Del Gobbo (2004), I propose that when a predicate of individuals (type <e,t>), modifies a singular term (type e), it simply intersects with the singleton set formed from the denotation of the latter. The structure is acceptable if the predicate denoted by the relative is true of the individual denoted by the DP.

6. Where do Chinese relative clauses stand?

Once we establish that there is a syntactic and semantic similarity between appositive adjectives and restrictive relative clauses, we can explain the behavior of the Chinese sentences seen in section 4. Those relative clauses are predicates of individuals that modify proper names or pronouns, namely singular terms (type e). This means that, for reasons that remain to be investigated thoroughly, relative clauses in Chinese, differently from relative clauses in English, have the potential to behave like appositive adjectives of English: in other words, they are more 'adjectival' than English relatives. Namely, they can modify nominals that denote individuals (type e), without detaching themselves from the same nominals like real appositives do. This explains why pronouns and anaphors inside these relatives are always in the scope of elements within the matrix clause. It also explains why sentential adverbs of modification yield ungrammaticality if added to Chinese relative clauses: this is so because these adverbs need to be part of a main assertion, and not of a predicate.[8]

In sum, Chinese relative clauses are either restrictive relatives or appositive adjectives. By this I mean that they always denote predicates of individuals (type <e,t>), and that they are always adjoined to the noun they modify, at all levels. What they cannot do is being real appositive relative clauses, namely propositions (type t) that detach from the matrix at some relevant level, certainly at LF, most probably also at the discourse level.

[8] As for the reasons why appositives cannot modify quantified heads, I refer the reader to my dissertation, Del Gobbo (2003).

7. Conclusion

In this paper I address the issue of the distinction between 'descriptive' and restrictive relative clauses in Mandarin Chinese. In relation to this, I investigate the meaning and the structure of Chinese relative clauses.

I first show that the so-called 'descriptive' relative clauses in Chinese are not appositive relative clauses, despite translations found in the literature that seem to indicate so. I propose to account for the restrictive vs. 'descriptive' contrast in terms of the s-level/deictic meaning versus i-level/generic meaning. According to Larson and Takahashi (2002), these two meanings have dedicated slots within the nominal phrase, namely the DP-layer is dedicated to s-level/deictic meaning and the NP-layer is dedicated to i-level/generic meaning. The ordering restrictions found for Chinese relative clauses show that such structure is to be assumed for Chinese as well and that it is ultimately responsible for the restrictive versus 'descriptive' contrast.

Even though I establish that Chinese so called 'descriptive' relative clauses are not appositive relative clauses, but restrictive modifiers, of the i-level/generic type, the question remains as to why relative clauses in this language are able to modify proper names and pronouns. My proposal is to distinguish appositive adjectives from appositive relative clauses, both semantically and syntactically, and to claim that when they modify proper names or pronouns, Chinese relative clauses behave like appositive adjectives. I propose that appositive relative clauses are propositions (type t), while appositive adjectives are predicates of individuals (type <e,t>). Syntactically, the former need to detach from the noun they modify at LF (and in discourse), the latter stay adjoined to it throughout the derivation. In Del Gobbo (2003) I propose to treat appositive relative clauses as an instance of E-type anaphora (following for example work by Sells 1985 and Demirdache 1991). In order for the appositive relative pronoun to be correctly interpreted as E-type, it needs to temporally follow the 'head' it modifies. This is ultimately the reason why Chinese relative clauses cannot be appositive: they are prenominal, hence they always precede the 'head' they modify. But this doesn't prevent them to act like appositive adjectives, namely like predicate of individuals (type <e,t>) that modify singular terms. By assuming that Chinese relative clauses have the property of behaving like appositive adjectives we explain why they can still modify proper names and pronouns while retaining the syntactic and semantic behavior of restrictive relative clauses. As a matter of fact, tests on question/answer pairs, NP-deletion and VP-deletion show that appositive adjectives are strikingly more similar to restrictive relative clauses than they are to appositive relative clauses. Appositive adjectives and restrictive relative clauses are predicates

of individuals (type <e,t>) and they are always adjoined to the noun they modify, while appositive relative clauses are propositions (type t) and they detach from the noun they modify after Spell-Out. The only feature that appositive adjectives and appositive relative clauses have in common and that differentiates them from restrictive relative clauses is their ability to modify singular terms.

There are still two important questions to be answered: if we assume, as I do, that appositive adjectives and restrictive relative clauses are both predicates of individuals, what is ultimately responsible for their main difference, namely the ability versus inability to modify singular terms? And, maybe more interestingly, what allows Chinese relative clauses to act like appositive adjectives and prevents English relative clauses to do so? My suspicion at this point is that answering the second question will lead us to an answer for the first one, but I leave this for future research.

References

Aoun, Joseph and Y.-H. Audrey Li. 2003. *Essays on the Representational and Derivational Nature of Grammar: the Diversity of Wh-constructions*. Cambridge MA, MIT Press.

Bolinger, Dwight. 1967. "Adjectives in English: Attribution and Predication", *Lingua* 18:1-34.

Cinque, Guglielmo. 2003. "The Prenominal Origin of Relative Clauses". *Workshop on Antisymmetry and Remnant Movement*, NYU.

Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles, University of California Press.

Chierchia, Gennaro. 1995. "Individual-level Predicates as Inherent Generics". In G. Carlson and J. Pelletier (eds.) *The Generic Book*. Chicago, The University of Chicago Press.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and Grammar*. Cambridge MA, MIT Press.

Del Gobbo, Francesca. 2001. "Appositives Schmappositives in Chinese". In M. Irie and H. Ono (eds.) *UCI Working Papers in Linguistics*, 7:1-25.

Del Gobbo, Francesca. 2003. *Appositives at the Interface*. Ph.D. dissertation, UC Irvine.

Del Gobbo, Francesca. 2004. "On Prenominal Relative Clauses and Appositive Adjectives". In B. Schmeiser, V. Chand, A. Kelleher and A. Rodriguez (eds.) *WCCFL 23 Proceedings*. 182-194. Somerville, MA, Cascadilla Press.

Demirdache, Hamida. 1991. *Resumptive Chains in Restrictive Relatives, Appositives and Dislocation structures*. Ph.D. dissertation, MIT.

Emonds, Joseph. 1979. "Appositive Relatives have no Properties". *Linguistics Inquiry* 10:211-243.

Giorgi, Alessandra. 1984. "Towards a Theory of Long-Distance Anaphors: a GB Approach". *The Linguistic Review* 3:307-362.

Hashimoto, Anne Yue. 1966. *Embedding Structures in Mandarin*. Ph.D. dissertation, The Ohio State University.

Hashimoto, Anne Yue. 1971. "Mandarin Syntactic Structures". *Unicorn (Chi-Lin)*, 8:1-149.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Ph.D. dissertation, MIT.

Larson, Richard. 2000. "Temporal Modification in Nominals". *International Roundtable on the Syntax of Tense*. Paris.

Larson, Richard and Naoko Takahashi. 2002. "Order and Interpretation in Prenominal Relative Clauses". *LSA* hand-out.

Lin, Jowang. 1997. "Some remarks on the semantics of prenominal relative clauses in Chinese". *The International Conference on Chinese Linguistics VII, Leiden University*. Leiden, The Netherlands.

Ogle, Richard. 1974. *Natural order and dislocated syntax*. Ph.D. dissertation, UCLA.

Ross, John. R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, MIT.

Safir, Ken. 1986. "Relative Clauses in a Theory of Binding and Levels". *Linguistic Inquiry* 17:663-689.

Sells, Peter. 1985. "Restrictive and Non-Restrictive Modification". *CSLI Report* No. CSLI-85-28. Stanford, California.

Svenonius, Peter. 1994. *Dependent nexus: Subordinate predication structures in English and the Scandinavian languages*. Ph.D. dissertation, UCSC.

A comparative analysis of restrictive and appositive relative clauses in Cushitic languages*

Mara Frascarelli and Annarita Puglielli
Università degli Studi Roma Tre

1. Introduction

Relative clauses represent an extremely intriguing empirical domain, both because of the complexity of the data and of the theoretical relevance of the construction. In this respect, a particularly interesting area of research is the distinction between restrictive and appositive clauses, for which different analyses have been proposed in terms of adjunction sites (for a general survey cf. the *Introduction* in Alexiadou et al., eds. 2000, Bianchi 2002).

All these analyses, however, have been challenged by Kayne's (1994) antisymmetry hypothesis, which excludes right-hand adjunction. Kayne takes Chomsky's (1977) approach to connectivity effects (in terms of Operator-movement of the head) to support a revised version of Vergnaud's (1974) "Raising Analysis", according to which the D° head selects the relative CP as its unique complement and the lexical NP is generated within the relative clause:

(1) $[_{DP}$ the $[_{CP}$ $[_{NP}$ **book**$]_k$ $[_{C'}$ that $[_{IP}$ I bought t_k]]]]

* We wish to thank Axmed Cabdullaahi Axmed, Cabdalla Omar Mansur (Somali) and Mohammed Ali Mahmoud (Afar) for their help and patience in providing, testing and discussing our data. We are also grateful to the members of the XXX IGG (Venice, February 26th-28th 2004) for their helpful suggestions and comments. General disclaimers apply.

The relativization chain is thus completely assimilated to the standard A'-movement chain: the head leaves a Variable within the relative clause which is an identical copy of the head itself. Crucially, Kayne assumes this explanation for *both* restrictive and appositive clauses, so that the two types of relative clauses are only distinguished by the following covert operation:

(2) $[_{DP}\,IP\,[_{D^\circ}\,[_{CP}\,NP\,[_{C'}\,C^\circ\;\;t_{IP}\,]]]]$

As is shown in (2), the appositive IP raises to Spec,DP in Logical Form (LF), thus escaping from the scope of the article in D°. Kayne thus concludes that the "intonational break" which typically distinguishes appositives from restrictives is only a "PF property".

Though a unified analysis of relative clauses is advantageous for many respects (cf. Bianchi 2000, 2002), it also triggers both theoretical and empirical problems. In the former instance, it implies that the PF interface can have access to LF operations (contrary to standard tenets), since the intonational break which characterizes appositive clauses crucially depends on LF movement. Regarding empirical problems, restrictive and appositive clauses show important asymmetries cross-linguistically, that challenge the feasibility of a unified derivation (cf. Alexiadou et al., eds., 2000).

The aim of this paper is to illustrate and discuss some major discrepancies between restrictive and appositive clauses, in order to propose a structural distinction between these constructions. The relevant data are taken from Somali and Afar, two Cushitic languages, whose major morphosyntactic properties will be illustrated in the next two sections.

2. Somali[1]

Somali is an SOV, non *pro*-drop, polysynthetic language (in the sense of Baker 1996). The latter condition entails that argument roles (also, "θ-roles") are only assigned through *incorporation onto the verbal head* (the so-called "Morphological Visibility Condition"), so that the argument structure of a verb is only realised by means of *clitic pronouns*, which are disposed in the Verbal Complex (VC).[2] The SOV order is thus strictly realised within the VC (cf. Puglielli 1981), while *full DPs are merged in non-argument position* and connected to the sentence by means of resumptive pronouns, which bind constituents to their thematic roles.

[1] When not otherwise indicated, the Somali data exposed in this paper are original sentences.

[2] A number of facts support polysynthesis in Somali like, for instance, the absence of non-finite clauses and multiple wh-questions (cf. Svolacchia and Puglielli 1999 for details).

Somali is also a *Focus-prominent* language, so that in a main declarative sentence one constituent must be overtly realised as the Focus of the sentence. Nominal Focus occurs in preverbal position and is immediately followed by the Focus Marker (FM) *baa* (cf. (3)). In the absence of nominal Focus, *waa* marks the VC on its right side as new information (cf. (4)). Non-focused DPs are realised extrasententially as Topics (Focus is capitalized, as in standard use):[3]

(3) Shalay jamacadda CALI **baan** (baa+ aan) ku arkay.
yesterday university.ART Cali FM.SCL1SG at saw.1SG
'Yesterday, I saw CALI, at the university.'

(4) Shalay jamacadda Cali **waan** (waa+ aan) ku ARKAY.
yesterday university.ART Cali FM.SCL1SG at saw.1SG
'As for Cali, yesterday, I SAW him, at the university.'

When the subject is focused (or wh-questioned), Somali shows *Antiagreement effects* (cf. Ouhalla 1993, Frascarelli and Puglielli 2004):

(5) Hilib NIMANKAAS baa cunayá.
meat men.those.ABS FM eat.PRG.RD
'THOSE MEN are eating meat.'

(6) a. *Hilib NIMANKAAS**U** baa cunayá.
meat men.those.**NOM** FM eat.PRG.RD
b. *Hilib NIMANKAAS b**ay** cunayá.
meat men.those.ABS FM.**SCL3PL** eat.PRG.RD
c. *Hilib NIMANKAAS baa cunay**aan**.
meat men.those.ABS FM eat.PRG.**3PL**

As the ungrammaticality of (6a-c) shows, a focused subject does not show NOM Case (but the unmarked Absolutive (ABS) Case), it cannot be resumed

[3] The list of the abbreviations used in the glosses is the following:

ABS	=	absolutive case	NOM	=	nominative case
ART	=	definite article	PL	=	plural
AN	=	anaphoric article	POSS	=	possessive pronoun
CONJ	=	conjunctive head	PRG	=	(present) progressive
DEP	=	dependent paradigm	RD	=	reduced paradigm
F	=	feminine	RPR	=	relative pronoun
FM	=	Focus Marker	SCL	=	subject clitic
M	=	masculine	SG	=	singular

by a clitic (SCL) (which is obligatory in all other cases) and the verb appears in a reduced (but not invariable[4]) form of agreement, the so-called "reduced paradigm" (RD).

As far as relative clauses are concerned, both restrictives and appositives take the form of *postnominal modifiers* (though Somali is head-final in the VC). They are not introduced by any Complementizer-like element and belong to the head deletion type, so that neither a clitic nor a relative pronoun resume the head within the modifying clause (cf. Antinucci 1981, Gebert 1981). When the head-noun has a subject role in the relative clause, the verb shows Antiagreement (henceforth, AA) effects:

(7) a. Wiilka [Maryam la hadlay**á** /*hadlay**aa**]waa walaalkay.
boy.ART M. with speak.PRG.**RD** /***3SGM** FM brother.POSS.1SG
'The boy that is talking to Maryam is my brother.'

b. Wiilka [toosay**á** /*toosay**aa**] baa warshad ka shaqaysaa.
boy.ART wake up.PRG.**RD** /***3SGM** FM farm in work.PRG.3SGM
'The boy that is waking up works in a farm.'

c. Wiilaasha [baabuurka ku dhoof**áy** /*dhoof**een**] baan
boys.ART car.ART in leave.PRG.**RD** /***3PL** FM.SCL1SG
macasalaameyay.
greeted.1SG
'I greeted the boys that are leaving by car.'

As we can see, AA effects in the restrictive clause are independent of the type of Focus construction (*waa* in (7a), *baa* in (7b-c)) and of the syntactic role of the head in the main clause (a subject in (7a-b), an object in (7c)). This is also the case for appositive clauses:

(8) a. Cali [**oo** Maryam la hadlay**á** /*hadlay**aa**] waa walaalkay.
Cali Maryam with speak.PRG.**RD/*3SGM** FM brother.POSS.1SG
'Cali, who is talking to Maryam, is my brother.'

b. Cali [**oo** toosay**á** /*toosay**aa**] baa in yarsexday.
Cali wake up.PRG.**RD**/ ***3SGM** FM quantity little slept.3SGM
'Cali, who is waking up, didn't sleep much.'

c. Cali iyo Maryam [**oo** baabuurka ku dhoof**áy** /*dhoof**een**] baan
C. and M. car.ART in leave.PRG.**RD/*3PL** FM.SCL1SG
macasalaameyay.
greeted.1SG
'I greeted Cali and Maryam, who are leaving by car.

[4] The RD shows three forms: one for 3SGF, one for 1PL and one for all other persons.

According to our informants, no particular intonational break is produced between the head and the appositive clause, which is introduced by a specific element, namely *oo* (to be discussed in section 4).

3. Afar[5]

Afar is also a Cushitic SOV language. However, its morphosyntactic properties are quite different from Somali, so that a comparison of the two languages is very effective for the issues of the present study. Afar is, in fact, an inflectional *pro*-drop language, so that DPs carry argument role and pronouns are "strong" elements, realised as object of either verbs or postpositions. Consider the following examples (from Bliese 1981):

(9) amoy't-i 'sara daa'me.
the-chief.NOM clothes.ABS bought.3SGM
'The chief bought clothes.'

(10) 'kimal moo'tar-at 'bilu-k yemee'te.
yesterday car.ABS-by Bilu.ABS-from came.3SGM
'Yesterday he came by car from Bilu.'

(11) a'nu 'kaa 'ko-h ruu'be.
I him you-to sent.1SG
'I sent him to you.'

As is shown, the Afar Case system distinguishes NOM Case (for subjects) from ABS Case (marking non-subject constituents, nominal predicates and Foci). As with Somali, ABS Case is considered the unmarked (i.e., "citation") form of nouns.

As for relative clauses, restrictives precede the antecedent, independent of the syntactic role of the head and consistent with the head-final character of the language. Like in Somali, no pronominal form resumes the head within the relative clause. However, differently from Somali, AA effects are not present. Consider the following (from Bliese 1981):

(12) [a'nu ub'le] aw'ki 'daa cammi'se.
I saw.1SG boy.NOM stone.ABS threw.3SGF
'The boy that I saw threw a stone.'

[5] Afar data are either the result of our original research or taken from Bliese's (1981) grammar. The latter case is indicated in the text.

(13) [yemee'te] ab'ba tuble.
came.3SGM father.ABS saw.2SG
'You saw the father who came.'

(14) 'is ['yo-h yex'e] kabel'la bey'te.
she me-to gave.3SGM shoe.ABS took.3SGF
'She took the shoe that he gave me.'

Appositive clauses, on the other hand, are *head-initial*, they show the presence of a relative pronoun (*iyya*) that is marked for ABS Case (independent of its grammatical role)[6] and, finally, the verb always shows 3SGM agreement. Consider the following examples, in which the appositive head is, in turn, a subject, a direct object and an indirect object within the appositive clause (from Bliese 1981):

(15) 'awka ['abba fan ge'da-h-iyya] wee'ce.
boy.ABS father towards go.3SGM-*h*-RPR.ABS cried.3SGM
'The boy, who is going to his father, cried.'

(16) 'woo gi'ra [aw'ki yabbi'de-h-iyya] tiddigi'le.
that gun.ABS boy.NOM hold.3SGM-*h*-RPR.ABS broke.3SGM
'That gun, which the boy holds, is broken.'

(17) 'siiniy [yeddee'reh su'ge-h-iyya-h] 'daagu war'se.
you.ABS be-far.3SGM-*h*-RPR.ABS-to news told.3SGM
'He told the news to you (PL) who were [lit.: was] far.'

Some of these properties seem to support Kayne's (1994) analysis. Indeed, the absence of a Complementizer in prenominal restrictive clauses is immediately explained in terms of IP-movement to Spec,DP, which leaves a (null) C° head stranded in postnominal position (cf. (2)). However, prenominal restrictives show full agreement in Afar, so that splitting I° from C° cannot be taken as a cross-linguistic explanation for AA effects in these kinds of constructions (cf. Kayne 1994: 95). Moreover it is not clear why AA should appear in postnominal appositive clauses, where a(n invariable) relative pronoun is present.

[6] The pronoun *iyya* is also used as a wh-constituent. In this case it shows the NOM form *iyyi*:

(i) **'iyya** tab'le?
who.ABS saw.2SG
'Who did you see?'

(ii) **'iyyi** yemee'te?
who.NOM came.3SGM
'Who came?'

4. Restrictive vs. Appositive Relative Clauses

The semantic distinction between restrictive and non-restrictive modification is formally realised to different extents cross-linguistically, so that it is very often difficult to propose a structural distinction which can be taken as universal. In Somali and Afar, on the other hand, restrictive and appositive clauses show crucial formal and interpretative asymmetries (in terms of word order, binding, and scope properties). A comparative analysis can thus provide a crucial insight into the understanding of these constructions.

4.1. *Extraposition Effects*

While extraposition is totally blocked for restrictive clauses, some languages seem to allow this operation in appositives (cf. Alexiadou et al., eds. 2000). In the two Cushitic languages examined, restrictive clauses can never be extraposed, as is shown in (18b) for Somali and in (19b) for Afar:

(18) a. WIILKA [TOOSAYÁ] baa warshad ka shaqaysaa.
boy.ART wake up.PRG.RD FM factory in work.PRG.3SG
'THE BOY THAT IS WAKING UP NOW works in a factory.'
b. *WIILKA baa warshad ka shaqaysaa [TOOSAYÁ].

(19) a. [a'nu a'mo 'kaa-k oogo're] 'too 'num yer'de
I head him-off hit.1SG that man ran.3SGM
'That man [that I hit on the head] ran.'
b.*'too 'num yer'de [a'nu a'mo 'kaa-k oogo're]

On the other hand, appositive clauses can be extraposed in both languages, but only in *nominal Focus constructions*. So, in Somali extraposition is allowed in *baa* constructions (with a slight marginal effect), but not in the presence of *waa*:

(20) a. CAASHA [oo soor sameysáy] **baa** soo gashay.
Caasha food prepare.PRG.RD FM in entered.3SGF
'CASHA, who prepared the meal, entered.'
b. ?CAASHA baa soo gashay [oo soor sameysáy].

(21) a. CALI [oo arday ah] **baan** la kulmay.
Cali student be.PRG.RD FM.SCL1SG with met.1SG
'I met CALI, who is a student.'
b. ?CALI baan la kulmay [oo arday ah].

(22) a. Cali [oo wareersan] **waan** ARKAY.
Cali be-confused.RD FM.1SCL saw.1SG
'I SAW Cali, who was confused.'
b. *Cali ***waan*** *ARKAY* [*oo wareersan*].

In Afar, narrow Focus is generally realised in situ and is not lexically marked.[7] Nonetheless, like in Somali, extraposition is only possible when the head-noun is focused. Therefore, sentences like (23a) and (24a) are only allowed as answers to the questions given in the examples, while they are excluded in a broad Focus context (that is to say, as an answer to a question like "what happened?"). Of course, the non-extraposed version is also possible (given in (b)):

(23) **Q:** 'Iyyi wee'ce?
'Who cried?'
a. AHMED wee'ce ['abba fan ge'da-h-iyya]
Ahmed cried.3SGM father towards go.3SGM-*h*-RPR.ABS
'AHMED cried, who is going to his father.'
b. AHMED ['abba fan ge'da-h-iyya] wee'ce

(24) **Q:** maxa tiddigi'le?
'What is broken?'
a. GI'RA tiddigi'le [aw'ki yabbi'de-h-iyya]
gun.ABS broke.3SGM boy.NOM hold.3SGM-*h*-RPR.ABS
'THE GUN is broken, which the boy holds.'
b. GI'RA [aw'ki yabbi'de-h-iyya] tiddigi'le

The block imposed on restrictives (shown in (18)-(19)) is consistent with Kayne's raising analysis. Since right-hand movement is excluded by antisymmetry, extraposition does not imply movement of the relative clause, but (leftward) movement of the head. Hence, this movement is ungrammatical in restrictive clauses, since the string [D°+ HEAD-NOUN] *does not form a constituent* (cf. (1)). However, what prevents bare nouns from moving in appositives? And – what is more intriguing – how can Focus make this operation legitimate? It is clear that a raising approach to appositives cannot explain these facts.

[7] Bliese (1981) claims that the fronted option is also available, but apparently only to realise contrastive Focus: "any element of the sentence (except postposition and conjunctions) may be moved to or toward the front of the sentence to give emphasis or focus" (p. 102).

4.2. *Case Marking of the head*

In Afar, the head-noun of a restrictive clause is marked for Case according to its syntactic role in the matrix clause (cf. (12)-(14)). In appositive clauses, on the other hand, relative heads always show the unmarked ABS Case, as is shown by *'awka, gi'ra* and *'siiniy* in (15)-(17).

The realization of Case in Somali relative clauses also provides an interesting point for discussion. Case marking in Somali appears on the rightmost element within the DP. So, if a subject is modified by one or more adjectives, NOM Case only appears on the last constituent within the relevant DP:

(25) a. [gabart**a** yaree qurxoon**i**] guriga
girl.ART.**ABS** small and nice.**NOM** home.ART
bay aadeysaa.
FM.SCL3SGF go.PRG.3SGF
'The small and nice girl is going home.'
b. *[gabart**i** (**NOM**) yar ee qurxoon (**ABS**)] guriga bay aadeysaa

Consistently, when the head of a restrictive clause has a subject role, NOM Case marking is found at the end of the relative clause:[8]

(26) a. [wiilk**a** [aan af Talyaaniga ku hadlin**i**]]
boy.ART.**ABS** NEG language Italian.ART in speak.NEG.**NOM**
waa walaalkay.
FM brother.POSS.1SG
'The boy that cannot speak Italian is my brother.'
b. *[wiilk**u** (**NOM**) [aan af Talyaaniga ku hadlin]] waa walaalkay.

On the other hand, in appositive clauses *NOM Case marking is never realised*, whether it be on the head or at the end of the appositive clause:

(27) a. Adig**a** [oo aadan af Talyaaniga ku
you.**ABS** NEG.2SG language Italian.ART in
hadleynin] waa walaalkiis.
speak.PRG.NEG FM brother.POSS.3SG
'You, who cannot speak Italian, are his brother.'

[8] Since in affirmative clauses NOM marking on the verb is only realised through prosody (by means of a low tone), to show Case marking we have used a negative verb in (26), in which NOM Case is morphologically realised.

b. *Adig**u** (**NOM**) [oo aadan af Talyaaniga ku hadleynin] waa….
c. *Adig**a** [oo aadan af Talyaaniga ku hadleynin**i** (**NOM**)] waa…

The comparison between sentences (26) and (27) shows that the string [HEAD+RESTR] counts as *one constituent* for the purposes of Case marking and, specifically, that these two elements form *one and the same* DP. However, what about the string [HEAD+APPOS]? Not only do sentences like (27) show that they *do not form a (unique) DP,* they also prove that neither the head nor the appositive are in a position to receive NOM Case. This proves an additional challenge for a uniform analysis of relative clauses.

4.3. *Stacking effects*

In many languages, it is possible that both types of relative clause refer to the same head. In this case, however, the linear order must be one in which the restrictive clause is adjacent to the head and the appositive follows. This is also the case in Somali:

(28) a. Wiilka [hadlayá], [oo aan ku baray],
boy.ART speak.PRG.RD SCL1SG you introduced.1SG
baa Landan ka yimid
FM London from came.3SGM
'The boy that is talking, whom I introduced you before, comes from London.'
b. *Wiilka [oo aan ku baray], [hadlayá], baa Landan ka yimid.

It is clear that in a uniform analysis the order of relative and appositive clauses should not be affected by such a restriction.

In Afar, where restrictive and appositive clauses are located on different sides, linear order is not an issue. However, it is interesting to point out that in this type of "complex modification" an introducing head appears between the head and the appositive clause, namely *–ay* (to be discussed later):

(29) [yinniki'se] 'awka [**ay** 'abba fan ge'da-h-iyya] wee'ce
fell.3SGM boy.ABS father towards went.3SGM-*h*-RPR.ABS cried.3SG
'The boy that fell, who is going to his father, cried.'

4.4. *Binding by external Operators and Scope of Negation*

As is generally agreed in the literature, appositives are impervious to syntactic binding by external Operators (cf., among others, Fox 2002). So, for instance, a QP can be the antecedent for an element within a restrictive relative clause, but it cannot be coreferent to a DP located within an appositive clause. This is precisely the case in Somali:

(30) a. Arday walba$_k$ gabarta [Ø$_k$ jecel] buu
student every girl.ART (him) love.RD FM.SCL3SGM
la cayaar aaday.[9]
with danced.3SGM
'Every student$_k$ danced with the girl [that loves him$_k$]'
b. *Arday walba$_k$ Maryam [oo Ø$_k$ jecel] buu la cayaar aaday.
'*Every student$_k$ danced with Mary, [that loves him$_k$]'

Similarly, appositive clauses cannot have a negative polarity item (NPI) as an antecedent, as is shown in (31b):

(31) a. Ciidna wiil [calaacalay] ma maqlo.
no boy complained.RD NEG heard.NEG
'I noticed no boy that complained.'
b. *Ciidna wiilkii [oo calaacalay] ma maqlo.
no boy.AN complained.RD NEG heard.NEG
'*I noticed no boy, who complained.'

These facts also represent a major cross-linguistic difference between the two types of relative clause and provide additional evidence against a mirror analysis.

4.5. *Extraction*

In some languages, it is possible to extrapose elements from within a restrictive relative clause, while appositive clauses behave as syntactic islands, so that no constituent can be extracted from them (cf. Engdahl 1997 for Swedish). This is also the case in Somali, as is shown in the following sentences (in which the "Complex NP" is taken into account):[10]

[9] Notice that 3rd person object clitics in Somali are realised as Ø forms.
[10] In Afar, on the other hand, extraction is not allowed in either case (as in many other languages, like English and Italian).

(32) a. qofka [buugga qoráy] baan bartay.
person.ART book.ART wrote.RD FM.SCL1SG knew.1SG
'I knew the person who wrote the book.'
b. buugga$_i$ qofka [t_i qoráy] baan bartay.

(33) a. Cali [oo buugga qoráy] baan bartay.
Cali book.ART wrote.RD FM.SCL1SG knew.1SG
'I knew Cali, who wrote the book.'
b. *buugga$_i$ Cali [oo t_i qoráy] baan bartay.

5. The syntactic proposal

Our investigation has shown that, despite their different core grammars, Somali and Afar show a number of morphosyntactic and interpretative phenomena (summarised in the table 1) that invoke a distinction between restrictive and appositive clauses.

Table 1: morphosyntactic and interpretative phenomena in Somali and Afar

	SOMALI		AFAR	
PHENOMENON	RESTR	APPOS	RESTR	APPOS
unmarked position	yes	yes	yes	no
extraposition	no	yes	no	yes
NOM Case on the head	yes	no	yes	no
NOM Case on RPR	no RPR	no RPR	no RPR	no
stacking	yes		not relevant	
binding by external QP	yes	no	no available data	
extraction	yes	no	no	no

5.1. *A structural distinction*

In the light of the data examined, we consider Kayne's (1994) analysis to relativization as appropriate to account for *restrictive* clauses. We thus maintain a promotion analysis for restrictives, according to which the relative clause is the complement of a D° head and the NP-head is an Operator sitting in Spec,CP and connected with a variable (i.e., a *deleted NP*, according to the "copy and delete" theory of movement; cf. Chomsky 1995,

2004). Therefore, the NP-head is part of the relative CP and forms a (movement) chain with a Variable in IP-internal position:[11]

(34) $[_{DP}\ [_{D'}\ [_{\mathbf{CP_{RESTR}}}\ [\text{NP-head}]_k\ [_{C'}\ [_{IP}\ \ldots \text{VAR}_{k\ldots}]]]]]$

On the other hand, we propose an analysis of appositive clauses in terms of an *assertive sentence "conjoined" to its antecedent* (in the spirit of Rebuschi 2002, 2005[12]):

(35) $[_{\text{ConjP (=DP)}}\ [\text{DP-head}]_k [_{\text{Conj'}}\ \text{CONJ}^\circ\ [_{\mathbf{CP_{APPOS}}}\ \sout{\mathbf{DP}}_k/\text{RELP}_k\ [_{C'}\ [_{IP}\ldots \text{VAR}_{k\ldots}]]]]]$

According to this analysis, the *antecedent* is merged as an *independent* and *fully referential DP*, while the *appositive clause is a CP* providing some additional information about the antecedent (e.g., a property, a definite description, background/new information, etc.; cf. Chierchia and McConnel-Ginet 1990, Doron 1994, Del Gobbo 2003). The appositive clause is *conjoined with its antecedent* through an "asymmetric conjunctive structure", whose head is overtly realised in some languages (like Somali and Afar) and left covert in others (for reasons to be made clear later).

The antecedent DP is therefore *not* the head of an A'-chain, but is coindexed with either a deleted copy (a ∅ form) or a relative pronoun (an "e-type" pronoun, cf. Heim 1990, Del Gobbo 2003), sitting in Operator position in the appositive CP. Given Spec-head agreement, the whole ConjP assumes the properties of its Spec (Johannesen 1998) hence, it is a "big DP".[13]

Finally, as is shown in (34) and (35), both restrictive and appositive clauses include the presence of a Variable within the relative IP. Specifically, in restrictive clauses the Variable is connected to the NP-head through a movement chain, while in appositive clauses this connection is mediated by the relative (or zero) pronoun sitting in Spec,CP. This is a crucial point, that provides an explanation for a number of structural and interpretative issues, as the analysis of AA effects will show (section 6.1).

---

[11] Given the structure in (34), the IP-initial order in Afar (an N-final language) is obtained through IP-raising to Spec,DP, while in polysynthetic Somali the head-noun further raises to D°, so as to incorporate into the Determiner (as assumed in Kayne 1994 for Rumanian).

[12] Rebuschi (2002, 2005), however, assumes a "linking morpheme" both in restrictive and appositive clauses, while our proposal claims for a structural distinction and supports a promotion analysis for restrictives.

[13] This analysis clearly explains why the head of appositive clauses can be a proper name or a pronoun across languages, while this is excluded for restrictive clauses. The independent generation of the antecedent as a fully referential DP also explains why, crosslinguistically, the head of an appositive clause *cannot be an Operator* (i.e., a QP or a WH-constituent).

This analysis of appositive clauses can be considered an extension of Kayne's (1994) theory on coordination. It is in fact reminiscent of the Kaynian approach to possessive constructions and APs (Kayne, 1994:85*ff.*)[14] and, as such, it is fully in the spirit of antisymmetry. This proposal is also in line with Chomsky's (2004) "late merge" analysis of non-argument information:

(36) We saw[$_{NP}$a painting], (that is) [~~a painting~~ [$_{ADJP}$ from the museum]]

Like adjuncts (and unlike restrictives), appositives may be left out without loss of grammaticality. In this sense, they can be considered an "afterthought structure" in which the head of the apposition undergoes ellipsis.

5.2. *Evidence for a ConjP Structure*

In this section we will consider additional data supporting a conjoined structure to account for the syntax of appositive clauses. In particular, we will show that the antecedent DP is merged as an independent constituent but, at the same time, shows a clear *morphosyntactic connection with the appositive CP*.

In this respect, the first element that is worth mentioning is the so-called "anaphoric article" in Somali, that is to say, a specific type of Determiner that modifies the antecedent DP in appositive clauses (when it is not a proper name). Consider the following:

(37) Xasan gabar**tii** /*gabar**ta** [oo guriga u socota].
Xasan girl.**AN** /***ART** CONJ house.ART to go.PRG.RD
buu arkay
FM.SCL3SGM saw.3SG
'Xasan met [that (specific) girl]$_{k}$, [who$_{k}$ is now going to home].'

As we can see, the definite article cannot be used in this context, while it is present in restrictive clauses (cf. (7)). This means that the selection of the Determiner is connected with the presence of a specific modifier and, in particular, that the anaphoric article *and* the appositive CP refer *conjointly* to some previous information. This connection is well explained syntactically through a conjoined structure, in which the denotation of the antecedent is calculated on the basis of the entire "big DP" (cf. also Doron 1994).

[14] Rebuschi (2005) discusses many similarities between relative clauses, APs and locative PPs in Chinese and Turkish. Crucial similarities are also present in Somali (cf. note 16).

Another important consequence of the structure proposed in (35) is that the appositive CP must be considered a *dependent clause* (as also argued in Demirdarche 1991). This is clearly shown in Somali by the *absence of focalization*, which is instead present in every main, declarative sentence (as it is a Focus-prominent language). So, no element can be focused in an appositive clause and this is evidence that the appositive CP is *structurally dependent on its antecedent*:

(38) a. Cali [oo MARYAM (***baa**) la hadlayá] waa macallin.
Cali CONJ Maryam FM with speak.PRG.**RD** FM teacher
'Cali, who is talking to Maryam, is a teacher.'
b. Cali [oo Maryam (***waa**) la HADLAYAA] waa macallin.
Cali CONJ Maryam FM with speak.PRG.3SG FM teacher
'Cali, who is talking to Maryam, is a teacher.'

In particular, we claim that this structural dependency is triggered by an "Identification Condition": the DP-head must *c-command* and *identify* the anaphoric pronoun in the Spec,CP of the appositive clause. Strong support for this claim comes from the analysis of sentences like (39) and (40) below, that reproduce the same meaning in Somali and Afar respectively:

(39) a. [Isaga$_k$ [**oo** Ø$_k$ isbitaalka ku jira]] buu dhintay.
he CONJ hospital.ART in stayed.RD FM.SCL3SGM died.3SGM
'He died while he was in the hospital.'
b. *[**oo** Ø$_k$ isbitaalka ku jira] isaga$_k$ baa dhintay.

(40) a. [Ahmed$_k$ [isbital suge-iyya$_k$]] rabe.
Ahmed hospital.ART be.3SGM.RPR.ABS died.3SGM
'Ahmed died while he was in the hospital.'
b. *[isbital suge-iyya$_k$] Ahmed$_k$ rabe.

As we can see, the appositive clause in the relevant sentences is interpreted as an adverbial clause. Indeed, from a semantic point of view, it provides a piece of information concerning the time of the event expressed in the matrix clause. However, it is structurally an appositive and, as such, it is *strictly dependent* on the DP-head (*isaga*/*Ahmed*). Hence, it cannot be extraposed

and precede the antecedent-DP (or the anaphoric pronoun does not meet the identification requirement).[15]

Bianchi (2000) proposes that appositives raise and attach to the matrix sentence at LF ("Long raising analysis"). The availability of an adverbial meaning for appositive clauses seems to support this analysis. In this respect, the following sentences should also be considered:

(41) a. [gabartii [oo markaas guriga **gashay**]] baan arkay
girl.AN CONJ time.ART house.ART entered.RD FM.1SCL saw.1SG
'I saw the girl (soon) after she got home.'
b. [gabartii [oo markaas guriga **galaysá**]] baan arkay
girl.AN CONJ time.ART house.ART enter.PRG.RD FM.1SCL saw.1SG
'I saw the girl when she was coming home.'

(42) [gabartii [oo ku taqan]] baa ag taadii markay
girl.AN CONJ you know.PRG.RDFM.1SCL near you.POSS passed.3SGF
'That girl, though she knows you, she did not greet you.'

As we can see, the temporal meaning associated with these sentences depends on the tense in the relative clause. So, with a past tense, the adverbial apposition expresses a sequence of two events (41a), while with a present tense the adverbial clause expresses contemporaneous events (41b). Finally, adverbial apposition can also encode concessive/adversative information (as is shown in (42)).

Let us finally consider the CONJ head (*oo* in Somali, *ay* in Afar). Though it *connects* the DP-head with the appositive CP, it cannot be considered a coordinative head in a proper sense. First of all, it connects two constituents of a different kind, thus forming what we have defined as an "asymmetric conjunction". Secondly, its behaviour shows crucial asymmetries with respect to "real" coordinative (CRD) heads, like *–na* in Somali. Consider the following:

(43) a. Cali wuu bukay adigu**na** dawaysay.
Cali FM.SCL3SGM be-sick.3SGM you.NOM-CRD cured.3SGM
'Cali was sick and you have cured him.'
b. *Cali wuu bukay **oo** adigu dawaysay.

[15] Also note that in (39) the presence of the subject clitic on the FM *baa* (yielding ***buu***) shows that the DP *isaga* is *not* the Focus of the main clause (since subject focusing excludes clitic resumption, cf. (6b)). This implies that the Focus in (39) is the *entire ConjP*.

(44) a. Cali hilib buu cunay caano buu**na** cabbay.
C. meat FM.SCL3SGM ate 3SGM milk FM.SCL3SGM-CRD drank.3SGM
'Cali ate meat and drank milk.'
b. *Cali hilib buu cunay **oo** caano buu cabbay.

As we can see, the CONJ *oo* cannot be used to connect two independent sentences (regardless of subject identity). Moreover, as is clearly shown, sentences coordinated by *–na* can both contain a Focus marker, while *oo*-clauses exclude focusing (cf. (38)). We thus conclude, with Thompson (1971), that "an appositive cannot be simply the counterpart of a coordinated structure" and that the CONJ head is *not a genuine coordinative head*: it is a partially *unspecified element* whose value is defined by distributional properties (cf. Rebuschi 2003).

As for its overt/covert realization, we suggest that an overt CONJ head in appositive constructions is required to differentiate meanings and avoid ambiguities in those languages in which categorial distinction crucially relies on syntactic structure. This is a very complex issue, that is far beyond the scope of this work. However, to exemplify our suggestion, consider "minimal pairs" like the following, in Somali:

(45) a. [Cali [keli ah]] baa yimid
Cali alone be.RD FM came.1SG
'**Only** CALI came.'
b. [Cali [**oo** keli ah]] baa yimid
Cali CONJ alone be.RD FM came.1SG
'CALI came, **alone**.'

The contrast offered in (45a-b) shows that the presence of the CONJ head in the relative clause distinguishes an adverbial interpretation from an appositive reading of the noun *keli*. Since in Somali "adverbs" and "adjectives" are not morphologically characterized (i.e., they do not form "categories" in the traditional sense), adverbial and adjectival interpretation only depends on the specific type of relative clause that is headed by the relevant head-noun. An overt functional head like *oo* is therefore needed to make this distinction clear.[16]

[16] Specifically, adjectives in Somali are realised as a particular class of verbs and they modify nouns in the form of restrictive clauses (cf. (ia)). It is thus interesting to note that, in the presence of more than one "adjectival" modification, the second modifier must take the form of an *appositive* clause (cf. (ib)) and their word order cannot be changed (cf. (ic)):

Avoidance of ambiguities can also be invoked to explain the presence of the CONJ head *ay* in Afar when the DP-head is modified both by a restrictive and an appositive (cf. (29) above). Since sentence coordination is generally realised through simple juxtaposition in Afar, it is plausible to assume that the CONJ head *ay* appears in this kind of complex structure to signal that the following predicate is an *apposition* to the DP-head (*awka* in (29)) and avoid an interpretation in which the Noun *abba* is interpreted as the subject of the following verb (thus obtaining the (wrong) reading: "the boy that fell *and* his father who is going cried")[17].

6. Back to data: a full account

Let us now resume the different phenomena examined in section 4. and see how the structural distinction proposed in section 5.1. can provide a comprehensive explanation for the asymmetries shown by restrictive and appositive clauses.

6.1. *Antiagreement effects*

We have seen in section 2. that both restrictive and appositive clauses show AA effects in Somali, when the head-noun has a subject role, and that such effects typically arise when the subject is a Focus (or wh-questioned).

In recent works, Frascarelli and Puglielli (2004, forthcoming) have argued for a cleft-like structure in Focus marking languages. According to this analysis, Focus is realised by means of a copular construction in which a matrix Focus Marker (an original copular form) selects a Small Clause (SC) as its complement. The subject of the relevant SC is a (restrictive) relative clause headed by a generic (overt/covert) NP ("person", "time", "place", etc.), that is the piece of information we lack and that we are going to provide in the predication (i.e., the rhematic part of the sentence). The predicate is therefore the focused DP that Operator-moves to Spec,FocP in order to identify the NP-head in the relative clause.

(i) a. Cali [$_{DP}$ baaburka$_k$ [fiican ah]] buu soo gatay
Cali car.AN nice be.RD FM.SCL.3SGM bought.RD
'As for Cali, he bought A NICE CAR.' [lit: Cali, he bought a car that is nice]

b. *Cali* [$_{DP}$ *baaburka*$_k$ [*fiican ah*]] [***oo*** [*duug ah*]] *buu soo gatay*
'Cali bought a NICE, OLD CAR.' [lit: Cali, he bought a car that is old and is good]

c. *Cali [$_{DP}$ baaburka$_k$ [**oo** [duug ah]] [fiican ah]] buu soo gatay

[17] The relevant ambiguity could not be solved by Case marking on the DP '*abba* because, as shown in section 3., the DP-head of an appositive clause always shows ABS Case.

This amounts to saying that the Focus is *not* the head of the relative clause, but is merged as an independent constituent in the SC. Hence, in the case of subject focusing, the relevant DP is in fact only *reinterpreted* as the subject of the verb (after the identification of the Variable), as is shown in (46b):

(46) a. CALI baa Soomali ah.
Cali FM Somali be.RD
'CALI is (the person that) is Somali.'
b. [$_{FocP}$ CALI$_k$ [$_{Foc'}$baa [$_{IP}$ t$_{baa}$ [$_{SC}$[$_{DP}$[$_{CP}$ ∅$_k$ [$_{IP}$ VAR$_k$ soomaali ah]]] t$_k$]]]]

This provides an immediate explanation for AA effects: the reduced ("participial") form of the verb is triggered by the presence of an empty subject (within the relative clause) in a non *pro*-drop language[18] (for discussion and details, cf. Frascarelli and Puglielli 2004, forthcoming).

This analysis plays a crucial role in the present proposal since, as we have seen in section 5.1., both restrictive and appositive clauses include the presence of a Variable within the relative IP. Thus, if the Variable plays a subject role, this creates the structural condition for AA effects. This is shown for sentences (7b)-(8b), repeated below as (47)-(48):

(47) a. Wiilka [toosayá] baa warshad ka shaqaysaa.
boy.ART wake up.PRG.RD FM farm in work.PRG.3SGM
'The boy that is waking up works in a farm.'
b. [$_{DP}$ **[wiil$_k$**ka [$_{CP}$ **t$_k$** [$_{IP}$ **VAR$_k$** toosayá]]] …

(48) a. Cali [oo toosayá] baa in yar sexday.
Cali CONJ wake up.PRG.RD FM quantity little slept.3SGM
'Cali, who is waking up, didn't sleep much.'
b. [$_{ConjP}$ [$_{DP}$ **Cali$_k$**] [$_{ConJ'}$ oo [$_{CP}$ **∅$_k$** [$_{IP}$ **VAR$_k$** toosayá]]]] …

[18] The connection between AA and the *pro*-drop parameter is also evident in non-Afroasiatic languages. Consider, for instance, cleft sentences in English as compared to *pro*-drop Italian (for further discussion, cf. Frascarelli 2000b).

(i) a. It is ME (1SG) that goes (3SG) to America.
b. Sono IO (1SG) che vado (1SG) in America.

The reduced paradigm is thus dependent on the presence of a Variable sitting in subject position in the relative clause, and *not* on the particular FM in the matrix sentence. Consider, for instance, (49) below (an appropriate answer to "where is Cali, with the present?"), where *waa* is used:

(49) [Cali$_k$ [oo ∅$_k$ hadiyadda keenaya /*keenay**aa**]] waa IMANAYAA
Cali CONJ present.ART bring.PRG.**RD**/***3SGM** FM come.PRG.3SGM
'Cali is ARRIVING, with the present.'
(lit.: Cali, who is bringing the present, is arriving)

As we can see, the verb in the matrix clause (*imanayaa*) shows full agreement, given the unfocused status of the subject. Nonetheless the verb in the appositive clause is necessarily in the reduced form, due to the presence of an empty element in subject position.

6.2. *Extraposition Effects*

As shown in section 4.1., restrictive clauses in Somali and Afar can never be separated from their head, while extraposition is possible for appositives when the relevant head is a narrow Focus. The present analysis can explain this asymmetry. In restrictive clauses the NP-head does not form a maximal projection with the Determiner and, as such, these two elements cannot be subject to movement. On the contrary, the antecedent DP in appositive clauses is merged independently. Movement is thus possible, but *only if triggered by interface requirements*, in the spirit of Minimalism: this condition makes Focus relevant.

In nominal Focus constructions the DP-head must move to SpecFocP to check the [+F] feature (cf. Frascarelli 2000a). This interface requirement does not extend to the appositive CP, which is *not* included within the DP-head. Therefore, given a Merge structure as in (50a), either the entire ConjP moves to SpecFocP (as in (50b)) or *Cali* only moves, leaving the appositive CP stranded (as in (50c)). Hence, extraposition is in fact CP-stranding:

(50) a. [$_{FocP}$ baa [$_{SC}$ [$_{DP}$[$_{CP}$∅$_k$ [$_{IP}$ aan arkay]] [$_{ConjP}$ Cali$_k$ [oo wareersan]]]]]
FM SCL1SG saw.1SG C. CONJ be-confused.RD
lit.: (what) I saw (is) [Cali [who was confused]]
b. [$_{FocP}$**[$_{ConjP}$ CALI$_k$ [oo wareersan]]$_j$** [$_{Foc'}$ baan [$_{SC}$[$_{DP}$[$_{CP}$∅$_k$ arkay]]t$_j$]]]
c. [$_{FocP}$**[CALI]$_k$**[$_{Foc'}$ baan [$_{SC}$[$_{DP}$[$_{CP}$∅$_k$ arkay]] [$_{ConjP}$t$_k$ [oo wareersan]]]]]
'I saw CALI, who was confused.'

On the other hand, extraposition is excluded for non-focused DP-heads because no Operator movement is required for them. As argued in Frascarelli and Puglielli (2004), *waa* constructions are existential statements used to give salience to an *in situ* predicate. Moreover, in a verb Focus construction the "subject" DP is in fact a Topic and, as such, is merged in extrasentential position. It is thus evident that, given the basic order (51a), extraposition (in (51b)) is simply *impossible to derive*: the ConjP sits in extrasentential position and a lower position of the appositive clause could only be obtained through rightward movement (that is excluded for independent reasons):

(51) a. [$_{TopP}$ [$_{ConjP}$ Cali [oo [wareersan]]] [$_{FocP}$ waan [$_{TopP}$ ARKAY]]]
Cali CONJ be-confused.RD FM.1SCL saw.1SG
'I SAW Cali, who was confused.'
b. *Cali waan ARKAY [oo wareersan]

This analysis is further supported by the observation that extraposition is particularly frequent when the DP-head is indefinite. As is well known, indefinite nouns convey new information, hence their movement to SpecFocP with the stranding of additional (non-focused) information is considered the most natural option by speakers (cf. Gebert 1981:89):

(52) a. Cali GABAR buu arkay [oo talyaani ah]
Cali girl FM.3SGM.SCL saw.RD CONJ Italian be.RD
'Cali met a girl who is Italian.'
b. MAGAALO baan tagnay [oo la yirahdo Galkacyo]
town FM.1SG.SCL went.RD CONJ IMP call.DEP Galkacyo
'We went to a city called Galkacyo.'

Also in this case, the presence of *waa* makes the relevant sentences completely unacceptable, as expected:

(53) a. *Cali gabar wuu ARKAY [oo talyaani ah]
'Cali MET a girl, who is Italian.'
b. *Magaalo waan TAGNAY [oo la yirahdo Galkacyo]
'We WENT to a city, called Galkacyo.'

6.3. *Case marking and Stacking effects*

Data in section 4.2. have shown that, both in Somali and Afar, the head of an appositive clause always shows the unmarked ABS Case, independent of its syntactic role. This is a clear consequence of the ConjP that we propose.

We have seen that Case marking in these languages only takes place on the *rightmost* constituent within the DP. So, in coordinated structures it only shows on the second conjunct, as is illustrated in (54)-(55) for Somali and Afar, respectively:

(54) [$_{DP}$[$_{DP}$ qalin**ka**] iyo [$_{DP}$ buugag**gu / *-ga**]] miiska way saran yihiin
pencil.ART.**ABS** and books.ART.**NOM/*ABS** table.ART FM.SCL3PL lay.3PL
'The pencil and the books are on the table.'

(55) [$_{DP}$[$_{DP}$mu'sac ca'd**a**] 'kee [$_{DP}$nabuw'wab bar's**i /*-**'s**a**]] 'tu 'maduu'da
Moses law.**ABS** and prophets teaching.**NOM**/***ABS** thing NEG.able.IMPF
'Moses' law and the prophets' teaching achieve nothing.'

Consistently, in appositive constructions the first conjoined element (the DP-head) *cannot* be marked for Case, while the second conjunct is an independent CP and, as such, it is not subject to Case marking.[19]

Stacking effects (shown in (28)) also represent an immediate consequence of the present analysis since, in that kind of complex modification, the NP-head of the restrictive clause is part of the DP-head which is the antecedent of the (conjoined) appositive CP. Hence, an appositive clause must necessarily follow a restrictive one (and any other internal modifier referring the same head-noun), as is shown in (56) below:[20]

---

[19] It is worth mentioning that predicative DPs also show ABS Case in Cushitic languages. Consider, for instance, the following copular sentences in Somali (i) and Afar (ii):

(i)	Tani	waa	sonkort**a**	/ *sonkort**i**	'This is sugar.'
	this.NOM	FM	sugar.ART.**ABS**	sugar-ART.**NOM**	
(ii)	a'l**i**	ra'ku**ubu**	/* ra'kub		'The animal is a camel.'
	animal.NOM	camel.**ABS**	camel.**NOM**		

[20] This word order constraint is reminiscent of Kayne's (1994) analysis for sentences like (ib):

(i) a. The student of chemistry from New Jersey.
b. *The student from New Jersey of chemistry.

Kayne suggests that *of chemistry* is a complement of *the students* while *from New Jersey* is a *predicate*. This is very much in the spirit of the distinction that we are proposing.

(56) [$_{IP}$ [$_{ConjP}$ [$_{DP}$[$_{NP}$ wiilka$_k$ [$_{IP}$ VAR$_k$ hadlayá]]] [$_{Conj'}$oo [$_{CP}$ ~~wiilka~~$_k$ [$_{IP}$ aan ku Ø$_k$ baray]]]] baa Landan ka yimid]]] (= (28))
'The boy that is talking, whom I introduced you before, comes from L.'

As for Afar, this analysis can account for word order without positing an *ad hoc* exception to the otherwise very consistent (S)OV order of the language. Indeed, since the appositive clause is a predicative CP (in a conjoined structure), it must follow the head, as expected:

(57) [$_{IP}$ [$_{ConjP}$ [$_{DP}$[$_{IP}$ VAR$_k$ yinniki'se] [$_{NP}$ 'awka$_k$ t$_{IP}$]] [$_{Conj'}$ ay [$_{CP}$ [$_{IP}$ Ø$_k$ 'abba fan ge'da-h-] iyya$_k$ t$_{IP}$] wee'ce]]] (= (29))
'The boy that fell, who is going to his father, cried.'

As indicated in (57), the prenominal position of the restrictive clause is derived through IP-movement to Spec,DP, consistent with Kayne's assumptions. The same kind of operation applies in the appositive clause, thus deriving the final position of the relative pronoun *iyya* (located in C°).[21]

6.4. *Binding by external Operators, scope of Negation and Extraction*

In sentences (30) and (31) we have seen that QPs/NPIs can be the antecedent for an element within a restrictive clause, while this is excluded for appositive clauses. Given the structural distinction proposed, this asymmetry is also easily explained.

Restrictive CPs are complements of a D°-head whose Spec,DP provides the target for QR (*Quantifier Raising*) at LF. Hence, after (covert) Operator-movement, both QPs and NPIs scope over the sentence and bind a DP therein contained.[22] Appositive clauses, on the other hand, are impervious to syntactic binding by external Operators because the DP-head is merged in the Spec, ConjP position, as the "subject" of a Conjoined structure. Hence, a QP referring to this head cannot reach "the highest Operator position" and scope over the entire "big DP". Therefore, appositive clauses form an independent binding domain with respect to their head.

[21] The head status of the relative pronoun in Afar is shown by its incorporation into the leftward verb, so that it cannot be separated from the rest of the sentence by any kind of linguistic material. This "grammaticalized" status can also account for the obligatory 3SG agreement in appositive relative clauses.

[22] Remember that, according to Kayne's analysis (1994:26-27), constituents located in the highest Specifier position within a subject DP can c-command out of it.

As far as extraction is concerned, we can say that restrictive clauses allow for this option because the relative clause is the complement of a D° head and, as is known, complements are not islands and extraction is possible. On the other hand, appositive clauses are included in a Conjoined structure, that is to say, in a *syntactic island*. Extraction from an appositive clause is thus ungrammatical as it is from any second term of a coordinated structure.

7. Conclusions

The data discussed from Somali and Afar have provided evidence that the morphosyntactic properties of restrictive and appositive clauses cannot be fully explained within a uniform approach to relative clauses. In particular, we have shown that restrictive and appositive clauses have contrasting behaviour redarding extraposition, extraction and binding from external Operators. Moreover, appositive clauses are characterised by some specific properties concerning Case assignment, that never arises for restrictives.

We have therefore argued for a promotion analysis of restrictive clauses and refuted this approach for appositives. We have thus proposed an analysis of appositive clauses in terms of a "conjoined" structure, in which an independent DP-head is conjoined to the appositive CP and is connected to a relative (or zero) pronoun (an E-type anaphora) sitting in Spec,CP.

This analysis provides a clear explanation for the data discussed in the paper and sheds new light on the cross-linguistic understanding of some interpretative asymmetries between restrictive and appositive clauses. From a theoretical point of view, the structure proposed is in line with recent cross-linguistic studies of appositive clauses (cf. Del Gobbo 2003, Rebuschi 2002, 2003) and is fully consistent with antisymmetric requirements.

References

Alexiadou, Artemis, Paul Law, André Meinunger and Chris Wilder (eds.). 2000. *The Syntax of Relative Clauses*. Amsterdam/Philadelphia, John Benjamins.

Antinucci, Francesco. 1981. "Tipi di frase". In A. Puglielli (ed.) *Studi Somali 2. Sintassi della lingua somala*. 217-300. Roma, Ministero AA EE.

Baker, Mark. 1996. *The Polysynthesis Parameter*. New York/Oxford, Oxford University Press.

Bianchi, Valentina. 2000. *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin, De Gruyter.

Bianchi, Valentina. 2002. *Headed relative clauses in generative syntax. I-II. Glot*, Vol. 6:7-8.

Bliese, Loren F. 1981. *A generative grammar of Afar*. The Summer Institute of Linguistics and the University of Texas at Arlington.

Chierchia, Gennaro and S. McConnell-Ginet. 1990. *Meaning and grammar*. Cambridge MA, MIT Press.

Chomsky, Noam. 1977. "On wh-movement". In P. W. Culicover, T. Wasow and A. Akmajian (eds.) *Formal Syntax*. 71-132. New York, Academic Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA, MIT Press.

Chomsky, Noam. 2004. "Beyond Explanatory Adequacy". In Belletti A. (ed.) *Structures and Beyond. The Cartography of Syntactic Structures, Vol 3*. 104-131. New York/Oxford, Oxford University Press.

Del Gobbo, Francesca. 2003. *Appositives at the interface* Ph.D. dissertation, University of California, Irvine.

Demirdarche, Hamida. 1991. *Resumptive chains in restrictive relatives, appositives and dislocation structures*. Ph.D. dissertation, MIT.

Doron, Edit. 1994. "The discourse function of appositives". *Proceedings of IATL 1* 53-62. IATL, Hebrew University, Jerusalem.

Engdahl, Elisabet. 1997. "Relative Clause Extraction in Context". *Working Papers in Scandinavian Linguistics* 60:50-79.

Fox, Danny. 2002. "Antecedent-contained deletion and the copy theory". *Linguistic Inquiry* 33:63-96.

Frascarelli, Mara. 2000a. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht, Kluwer.

Frascarelli, Mara. 2000b. "Frasi Scisse e 'Small Clauses': Un'Analisi dell'Inglese". *Lingua e Stile* XXXV.3:417-446.

Frascarelli, Mara and Annarita Puglielli. 2004. "The Focus System in Cushitic Languages". In P. Fronzaroli e P. Marrassini (eds.) *Proceedings of the 10th Hamito-Semitic Linguistics Meeting* (Quaderni di Semitistica 23), Florence.

Frascarelli, Mara and Annarita Puglielli. Forthcoming. "Focus Markers and Universal Grammar". In Azeb Amha, Maarten Mous and Graziano Savà (eds.) *Cushitic and Omotic languages: Proceedings of the fourth International Conference on Cushitic and Omotic Languages (Leiden, 10-12 April 2003)*. Cologne, Rüdiger Köppe Verlag.

Gebert, Luciana. 1981. "Il sintagma nominale", "La coordinazione". In A. Puglielli (ed.) *Studi Somali 2. Sintassi della lingua somala*. 45-214. Roma, Ministero AA EE.

Heim, Irene. 1990. "E-type pronouns and donkey anaphora". *Linguistics and Philosophy* 13:137-177.

Johannesen, Janne. 1998. *Coordination*. New York/Oxford, Oxford University Press.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge MA, MIT Press.

Ouhalla, Jamal. 1993. "Subject-Extraction, Negation and the Anti-agreement Effect". *Natural Language and Linguistic Theory* 11:477-518.

Puglielli, Annarita. 1981. "Frase dichiarativa". In A. Puglielli (ed.) *Studi Somali 2*. *Sintassi della lingua somala.* 3-44. Roma, Ministero AA EE.

Rebuschi, Georges. 2002. "Coordination et subordination. Deuxième partie: vers la co-jonction généralisée". *Bulletin de la Société de Linguistique de Paris*, Tome XCVII. 1:37-94.

Rebuschi, Georges. 2005. "Generalizing the antisymmetric analysis of coordination to nominal modification". *Lingua* 115.4:445-459.

Svolacchia, Marco and Annarita Puglielli. 1999. "Somali as a polysynthetic language". In L. Mereu (ed.) *Boundaries of Morphology and Syntax*. 97-120. Amsterdam/Philadelphia, John Benjamins.

Thompson, Sandra. 1971. "Deep structure of relative clauses". In C. Fillmore and D.T. Langedoen (eds.) *Studies in Linguistic Semantics.* 79-96. New York, Holt, Rinehart and Winston.

Vergnaud, Jean Roger. 1974. *French Relative Clauses*. Ph.D. dissertation, MIT.

The Com(p-)position of DP-internal infinitival clauses*

Petra Sleeman
University of Amsterdam

1. Introduction

Since the publication of Pollock's seminal paper (1989) on the splitting of Infl into two distinct functional projections, TP and AgrP, generative linguists have tried to find evidence for the splitting of the left periphery or the middle field of the clause or the DP in several functional projections. Rizzi (1997) claims that the C-domain consists of four functional projections: ForceP, TopicP, FocusP and FinP. Cinque (1999) argues for the existence of a fixed universal hierarchy of a myriad of clausal functional projections in the middle field of the clause, based on the distribution of adverbs among others. On the basis of the distribution of adjectives, Cinque (1994) splits the middle field of the DP in discrete functional projections. Along the lines of the cartographic approach, Giusti (2002) and Aboh (2004) assign a split structure to the left periphery of the DP.

In this paper, I defend Rizzi's (1997) split structure of the left periphery of the clause, and more specifically the existence of the functional projection FinP. In Rizzi's system, the presence of FinP in the C-system is motivated by the fact that the choice of the complementizer reflects certain properties of the verbal system of the clause, e.g. in English *that* co-occurs with a finite verb and *for* co-occurs with an infinitive. Rizzi shows that while the finite complementizer *che* in Italian can be followed by a left-dislocated phrase, which he claims to be in TopP, the infinitival complementizer *di* can only be preceded by a left-dislocated phrase, which suggests that whereas *che*

* Thanks to an anonymous reviewer for the useful comments. Any remaining errors are my responsibility.

manifests the force position, *di* manifests the finiteness position. In this paper, I will claim that DP-internal subordinate clauses can be introduced by an infinitival complementizer in FinP. Whereas Rizzi bases the evidence for the location of *che* in ForceP and *di* in FinP mainly on word order data, I will adduce evidence for the position of complementizers from extraction facts. I will base my argumentation on French, although the construction that will be discussed also exists in e.g. other Romance languages and English.

The infinitival complementizer that I will claim to be located in FinP is the French prepositional complementizer *à* occurring in an infinitival construction that seems to be an equivalent of the relative clause:

(1) a. Il a été le seul Français à avoir atteint les sommets.
'He has been the only Frenchman to have reached the tops.'
b. Il a été le seul Français qui ait atteint les sommets.
'He has been the only Frenchman who has reached the tops.'

However, although *à* + infinitive seems to be an equivalent of the relative clause, Siloni (1995) shows that they behave differently with respect to extraction:

(2) a. les sommets$_i$ qu'il a été le seul Français à atteindre t_i
'the tops that he was the only Frenchman to reach'
b.*les sommets$_i$ qu'il a été le seul Français qui ait atteints t_i
the tops that he was the only Frenchman who has reached

Although *à* + infinitive behaves differently from the full relative clause, I will analyze both as a complement, adopting for both Kayne's (1994) raising analysis of relative clauses. The difference in syntactic behavior between *à* + infinitive and the full relative clause will be attributed to a difference in the left periphery of the clause. One of the arguments in favor of the idea that both *à* + infinitive and relative clauses are complements will be the dependency of both on antecedents that favor the subjunctive mood, although it will be shown that there can be differences between the two types of clauses with respect to the position of the antecedents in syntax. These differences will also be attributed to a difference in the left periphery of the clause.

The paper is organized as follows. I will first show that although in a framework such as Kayne's Antisymmetry Theory both relative clauses and *à* + infinitive are analyzed as complements (section 2), they behave differently with respect to extraction (section 3). In sections 4 and 5, I will show that both types of clauses depend on the same set of antecedents,

which trigger the subjunctive mood in the finite relative clause. I will furthermore show in section 5 that the antecedent of *à* + infinitive can be located within the infinitival clause in syntax. In section 6, I will relate the difference with respect to extraction and the possible difference with respect to the position of the antecedent in syntax to a difference in the left periphery of the subordinate clause within DP: in the relative clause the complementizer occupies Force° whereas as a prepositional complementizer it occupies Fin° in the infinitival clause. Finally, in section 7, I will summarize the results.

2. The equivalents of the relative clause

Kayne (1994) shows that the demonstrative pronoun *celui* 'the one' (fem.sg. *celle*, masc.pl. *ceux*, fem.pl. *celles*) necessarily has to be followed by a relative clause (3) or an equivalent. Kayne mentions three equivalents of the relative clause: a past participial phrase (4), a possessive PP (5) and the particles *ci* 'here' and *là* 'there' (6):

(3) celui que j'ai envoyé à Jean
'the one that I have sent to Jean'

(4) celui envoyé à Jean
'the one sent to Jean'

(5) celui de Jean
'John's'

(6) celui-ci ou celui-là
'this one or that one'

In Kayne's raising analysis of relative clauses, *celui* has to be interpreted in the Spec of CP or another clausal structure such as D/PP for the possessive phrase:

(7) $[_{DP}$ D° $[_{CP}$ celui$_i$ $[_{C'}$ que $[_{IP}$ j'ai lu t$_i$]]]]

(8) $[_{DP}$ D° $[_{D/PP}$ celui$_i$ $[_{D/P'}$ de $_{IP}$[Jean I° t$_i$]]]]

As Sandfeld (1965) shows, *celui* can furthermore be combined with present participles (9), *à* + infinitive with a "passive" meaning (10), and adjectives followed by a complement (11):

(9) ceux parlant quatre langues
'those speaking four languages'

(10) toutes celles à commettre
'all those to commit'

(11) celui capable de sacrifier sa vie
'the one able to sacrifice his life'

Sandfeld shows that *celui* can even be followed by adjectives ending in one of the suffixes *-able*, *-ible*, or *–uble*:

(12) les bouteilles en plastique non recyclables et aussi celles recyclables
'the plastic bottles that cannot be recycled and also those than can be recycled'

(13) Nous préférons ceux réutilisables.
'We prefer those that can be used again.'

Sleeman and Verheugd (1998) analyze all equivalents of the relative clause as reduced relative clauses. This means that in their analysis all these equivalents have a clausal structure.

In the next section, I show that, although in Kayne's analysis both the full relative clause and all its equivalents are analyzed as complements, *à* + infinitive behaves differently with respect to extraction than the full relative clause or the other equivalents.

3. Differences with respect to extraction

I have shown, in the previous section, that *à* + infinitive with a "passive" meaning can follow the demonstrative pronoun *celui*. *À* + infinitive can also follow certain adjectives such as *seul* 'only', *premier* 'first' or *dernier* 'last' and also superlatives. In this case *à* + infinitive can also have an "active" meaning:

(14) Elle est la seule à avoir participé à six éditions des Jeux Olympiques d'hiver.
'She is the only one to have participated in six editions of the Olympic Wintergames.'

(15) Il fut le premier à atteindre le pôle Nord.
'He was the first one to reach the North pole.'

(16) Ce sont les derniers à avoir rendu visite à la Lune.
'They are the last ones to have visited the moon.'
(17) Elle est la plus jeune à avoir publié un roman.
'She is the youngest one to have published a novel.'

Siloni (1995) observes that whereas extraction from the infinitival constituent in the "active" reading is possible, extraction from a relative clause or its other equivalents is not:

(18) Qu'est-ce qu'il est le seul à avoir fait?
'What is he the only one to have done?'

(19) * Qu'est-ce qu'il est le seul qui ait fait?
what is he the only one who has done

(20) Quel rallye a-t-il été le premier à couvrir?
'Which rally was he the first one to finish?'

(21) * Quel rallye a-t-il été le premier qui ait couvert?
which rally was he the only one who has finished

(22) * A qui_i Jean est-il le seul parlant régulièrement?
to whom is Jean the only one speaking regularly

The following examples involve a relative pronoun instead of an interrogative pronoun:

(23) ces paroles épouvantables que je fus le seul à entendre
'these horrible words that I was the only one to hear'

(24) le sommet qu'il fut le premier à atteindre
'the top that he was the first one to reach'

(25) * le $prix_i$ qu'elle est la seule personne fière d'avoir gagné t_i
the prize that she is the only person proud to have won

Siloni adopts a traditional analysis of relative clauses. She claims that the ungrammaticality of (19) and (21) results from a Subjacency violation, since Spec,CP of the relative clause is filled by an empty operator. A similar analysis would apply to (22) and (25):

(26) *Qu'$_i$ est-ce qu'il est le seul [$_{CP}$ *OP* qui ait fait t_i]?

In the infinitival construction there would be no empty operator in Spec,CP but PRO as the subject of the infinitive:

(27) Qu'$_i$ est-ce qu'il est le seul à [$_{CP}$ PRO avoir fait t_i]?

Siloni is not explicit about the structure of the sentence containing the infinitival clause. Probably the infinitival clause is not an adjunct, i.e. a kind of relative clause, but rather the complement of a preposition. The PP itself is probably the complement of the adjective *seul*:

(28) le seul [$_{PP}$ à [$_{CP}$ [$_{IP}$ PRO avoir fait]]]

However, the adoption of such a structure becomes problematic if we consider the next sentences:

(29) Il est le seul homme à avoir fait cela.
'He is the only man to have done that.'

(30) Il a été le premier journaliste à couvrir le rallye Paris-Dakar.
'He was the first reporter to finish the Paris-Dakar rally.'

(31) Il est le plus jeune joueur à avoir remporté déjà sept titres majeurs.
'He is the youngest player to have already won seven major titles.'

In (29-31), the infinitival constituent cannot be the complement of the adjective, but rather must be a relative clause, which is adjoined to the DP in Siloni's framework, see (26). However, (32) shows that extraction from the infinitival constituent in (29-31) is possible just as in (24):

(32) le sommet qu'il a été le seul Français à atteindre
'the top that he has been the only Frenchman to reach.'

If *à* + infinitive is the complement of *le seul* in this sentence, it becomes difficult to explain that they can be separated by a noun. If one admits that *à* + infinitive is a relative clause rather than a complement, however, it has to be explained why extraction is allowed in (32).

In Kayne's analysis of relative clauses and reduced relative clauses, the presence of the noun between the adjective *seul* and *à* + infinitive can easily

be accounted for. The noun moves from a position within the relative clause, the complement of *seul* in (33), to Spec,CP:[1]

(33) [$_{DP}$ le seul [$_{CP}$ Français$_i$ à t$_i$ atteindre le sommet]]

In line with Kayne's analysis, I analyze *à* + infinitive as a complement, which is also in line with Siloni's analysis. However, contrary to Siloni, I also analyze relative clauses as complements, in line with Kayne's analysis.

An argument in favor of an analysis as a complement in all cases, is that at least *le seul* necessarily has to be followed by a relative clause (or one of its equivalents). If the relative clause is an adjunct instead of a complement, this is unexpected:[2]

(34) Les seuls livres *(qui me plaisent) sont là.
'The only books (that please me) are there.'

Another argument in favor of the idea that the relative clause is a complement and not an adjunct, is the possibility to use the subjunctive mood in the relative clause after an antecedent containing e.g. a superlative, *seul*, *premier* or *dernier* (Carlsson 1969, Kampers-Manhe 1991):

(35) Le roi était fier d'avoir pour épouse la plus belle femme qui soit$_{subj}$ au monde.
'The king was proud of having as his wife the most beautiful woman in the world.'

(36) Elle est la seule (femme) qui ait$_{subj}$ vraiment compté pour lui.
'She is the only (woman) that has really been important to him.'

[1] In fact, the noun must even move to a position outside the (reduced) relative clause, probably to NumP, raising over "postnominal" adjectives (Cinque 1994):

(i) Elle est la femme$_i$ la plus âgée [$_{CP}$ t$_i$ à t$_i$ avoir eu un enfant].
'She is the oldest woman to have had a child.'

[2] In Kayne's (1994) analysis, the relative clause is the complement of a determiner. The dependency relation between *seul* and the relative clause in (34), however, shows that the relative clause can also be the complement of an adjective, cf. (28-29). I assume that some other adjectives, such as ordinals and superlatives (30-31) can also take a clausal complement.

(37) La première chose que l'on doive$_{subj}$ posséder, c'est une maison et c'est aussi la dernière chose que l'on doive$_{subj}$ vendre.
'The first thing that one has to own is a house and it is also the last thing that one has to sell.'

Besides its use in relatives clauses, the subjunctive is only used in sentential (substantive or adverbial) complements introduced by the complementizer *que*. If the relative clause is analyzed as a complement, it is possible to give a more uniform explanation of the use of the subjunctive in subordinate clauses. Furthermore, the fact that the relative clause containing a subjunctive can be replaced by *à* + infinitive might also plead in favor of an analysis of the relative clause as a complement:

(38) Elle est la femme la plus âgée qui ait$_{subj}$ gagné un prix.
'She is the oldest woman who has won a prize.'

(39) Elle est la femme la plus âgée à avoir gagné un prix.
'She is the oldest woman to have won a prize.'

In order to support the analysis of both *à* + infinitive and the relative clause as complements, I will show in the next two sections that there is a relation between the two constructions. What remains to be done then, in the last section, is to account for the differences between the two types of complements, e.g. the difference with respect to extraction mentioned above.

4. Subjunctive relative clauses

The adjectives and determiners that combine with a relative clause in which the subjunctive mood can be used are the superlatives, *premier* 'first' and *dernier* 'last', *principal* 'principal', *seul* 'only' and *unique* 'unique', and *(un des) rares* '(one of the) rare' or *peu de* 'few':

(40) C'est la meilleure chose qui puisse$_{subj}$ arriver.
'It is the best thing that can happen.'

(41) Le premier homme qui ait$_{subj}$ volé dans un avion à vapeur fut Ader.
'The first man who flew in a steam aeroplane was Ader.'

(42) La dernière chose qu'elle ait$_{subj}$ vue c'est un pare-brise lui broyant la boîte crânienne.
'The last thing that she saw was a windshield shattering her skull.'

(43) Le principal péril que nous courions$_{subj}$ aujourd'hui, c'est de ne pas écrire assez clair.
'The main risk that we run nowadays is not to write clearly.'

(44) C'est la seule possibilité que nous ayons$_{subj}$.
'It is the only possibility that we have.'

(45) Mais elle avait cette excuse, c'est qu'il était le premier, l'unique des jeunes hommes à qui elle eût$_{subj}$ jamais fait attention dans sa vie.
'But she had this excuse that he was the first, the only one of the young men she had ever paid attention to in her life.'

(46) Philip K. Dick est une des rares personnes qui aient$_{subj}$ compris que la bonne science-fiction est en fait la science-fiction sociale. La technologie est un reflet ou un écho de ce qui se passe dans la société.
'Philip K. Dick is one of the rare persons who have understood that good science fiction is indeed social science fiction. Technology is a reflection or an echo of what is going on in society.'

(47) Il y a bien peu de personnes qui sachent$_{subj}$ aimer.
'There are really few persons who know how to love.'

What are the semantic properties of the adjectives or determiners permitting the use of the subjunctive in the relative clause, and possibly also the use of *à* + infinitive?

Fauconnier (1980: 134) observes that subjunctive relative clauses depending on a superlative, *seul* etc. are ideal environments for the use of negative polarity items. They are in the domain of a monotone decreasing operator, the superlative element, which can reverse polarity (Zwarts 1981):

(48) Ce cadeau est le plus beau qu'on m'ait$_{subj}$ *jamais* fait.
'This gift is the most beautiful ever given to me.'

(49) Gustave est l'homme le plus compétent qui soit$_{subj}$ *le moindrement* intéressé par ce travail.
'Gustave is the most competent man who is somewhat interested in this work.'

(50) Ce sont les derniers récitals qu'elle ait$_{subj}$ donnés *de sa vie*.
'Those were the last recitals that she has given in her life.'

Kampers-Manhe (1991) analyzes subjunctive clauses as clausal complements that are in the domain of a negation provoking the use of the subjunctive (May 1985). The reason for this is that apart from superlatives, *seul* etc., a negative antecedent also provokes the use of the subjunctive and of negative polarity items:

(51) Je ne vois pas de voiture qui ait$_{subj}$ *la moindre* tache de rouille.
'I see no car that has any rust spots.'

According to Carlsson (1969), the superlative has this negative meaning if it is followed by a subjunctive clause. The meaning of (52) is: there is no other work that you can do than this one. In (53), in which the indicative mood is used, *seul* has a more positive meaning:

(52) C'est le seul travail que vous puissiez$_{subj}$ (jamais) faire.
'It is the only work that you can (ever) do.'

(53) C'est le seul travail que vous pouvez$_{ind}$ faire.
'It is the only work that you can do.'

I suggest that the more negative meaning of *seul* in (52) is due to focalization, implying the exclusion of a complement set (cf. Kiss' 1998 definition of identificational focus).

5. À + infinitive

Kampers-Manhe (1991) observes that relative clauses containing the subjunctive mood can be replaced by *à* + infinitive:

(54) Lucie est la seule de mes soeurs qui se soit$_{subj}$ mariée.
'Lucie is the only one of my sisters who has got married.'

(55) Lucie est la seule de mes soeurs à s'être mariée.
'Lucie is the only one of my sisters to have got married.'

Apart from *seul*, all other adjectives and determiners that can be followed by a subjunctive relative clause can be followed by *à* + infinitive. The infinitival clause can also contain a negative polarity item, as in (56):

(56) Il était le seul à avoir jamais osé la contredire.
'He was the only one to have ever dared to contradict her.'

(57) Elle est la première femme à occuper ce poste.
'She is the first woman to occupy this post.'

(58) La dernière femme à avoir reçu la peine de mort en Angleterre était Ruth Ellis.
'The last woman to have been executed in the U.K. was Ruth Ellis.'

(59) Elle était l'unique femme à avoir pris part à cette compétition.
'She was the only woman to have taken part in this competition.'

(60) Il fut l'un des principaux hommes à affrêter le navire.
'He was one of the principal men to prepare the ship.'

(61) Les rares hommes à avoir marché sur la lune étaient revêtus de scaphandres.
'The rare men to have walked on the moon wore diving-suits.'

(62) des conditions dont il est très peu d'hommes à vouloir entendre parler
'conditions of which very few men want to hear'

On the Internet, but not in the literary database Frantext, I also found examples of *à* + infinitive following ordinals other than *premier* and *dernier*, and following cardinal numbers, after which the subjunctive is never used in the normal case:

(63) Elle est la quatrième femme qui a_{ind} été élue à l'Académie française en 2000.
'She is the fourth woman who has been elected at the Académie française in 2000.'

(64) Elle est la quatrième femme à avoir été élue à l'Académie française en 2000.
'She is the fourth woman to have been elected at the Académie française in 2000.'

(65) Elle est l'une des trois femmes à jamais avoir accompli cet exploit au Canada.
'She is one of the three women who have ever succeeded to do this in Canada.'

It might be that *à* + infinitive is used here because the numerals have a focalized interpretation excluding others: there was almost no one before these women who had done this. This explanation is supported by the use of the negative polarity item *jamais* 'ever' in (65).

A more intriguing case is the use of *à* + infinitive (with an "active" meaning) after *celui*, of which I found many examples on the Internet but not in the literary database Frantext:

(66) C'est celui à avoir posé la dernière carte.
'It is he who has played the last card.'

(67) Celle à avoir survécu le plus longtemps fut Lilith.
'The one who survived the longest was Lilith.'

(68) Le vainqueur est celui à avoir le plus de cartes.
'The winner is the one who has the most cards.'

(69) Celui à avoir posé la première brique de cet assemblage sonore se nomme Third Eye Foundation.
'The one who has laid the first stone of this sound assembly is called Third Eye Foundation.'

Instead of *celui* I also found examples of determiner + noun followed by *à* + infinitive:

(70) Il est le joueur à avoir gagné le plus de matches cette année .
'He is the player who has won most matches this year.'

(71) Charles de L'Ecluse reste le personnage du 17ème siècle à avoir le plus contribué à l'introduction de la tulipe en Europe.
'Charles de L'Ecluse remains the person from the 17th century who has most contributed to the introduction of the tulipe in Europe.'

After *celui* or determiner + noun, a relative clause containing a subjunctive cannot be used:

(72) * C'est celui qui ait$_{subj}$ posé la dernière carte.
'It is him who has played the last card.'

(73) * Il est le joueur qui ait$_{subj}$ gagné le plus de matches.
'It is the player who has won most matches.'

What is striking in the examples of the use of *à* + infinitive after *celui* or determiner + noun, is that most of them contain a superlative, *premier* or *dernier* within the infinitival clause. There seems thus to be a difference between subjunctive relative clauses and *à* + infinitive clauses. Whereas the first group is only licensed by a real antecedent with a focalized interpretation, the second group is not only licensed by a real antecedent with a focalized interpretation but also by an adjective or an adverb with a focalized interpretation inside the *à* + infinitive clause.

I have shown that there is a close resemblance between subjunctive relatives and *à* + infinitive, both being used after antecedents with a "negative" meaning due to focalization, which motivates their analysis as complements, the noun being raised within the clausal complement. I have also shown that there is a difference between the two types. Whereas a focal element can also license an infinitival clause if it is inside the infinitival clause, this is not possible in the case of subjunctive relatives. In the next section, I investigate whether this difference can be related to the difference in extraction from the two types of clauses.

6. Analysis of the differences

I have shown in section 3 that whereas extraction from *à* + infinitive is possible, extraction from a full relative clause or an equivalent is not:

(74) Qu'$_i$ est-ce qu'il est le seul à avoir fait t_i ?
'What is he the only one to have done?'

(75) le sommet$_i$ qu'il a été le seul Français à atteindre t_i
'the top that he has been the only Frenchman to reach'

(76) * Quel rallye$_i$ a-t-il été le premier qui ait couvert t_i ?
'What rally was he the only one who has finished?'

(77) * A qui$_i$ Jean est-il le seul parlant régulièrement t_i ?
to whom Jean is the only one speaking regularly

(78) * le prix$_i$ qu'elle est la seule personne fière d'avoir gagné t_i
the prize that she is the only person proud to have won

Whereas Siloni (1995) analyzes *à* + infinitive as a complement, but the relative clause and its equivalents as adjuncts, I have adopted Kayne's (1995) analysis of relative clauses and reduced relative clauses and have analyzed all cases as complements. As an argument for this common

analysis, I have advanced the close resemblance between *à* + infinitive and subjunctive relative clauses. I have shown that they both depend on antecedents with a "negative" meaning due to focalization, such as superlatives, *seul*, *premier* and *dernier*.

If the relative clause and all its equivalents are complements, another account for the difference in extraction than the one proposed by Siloni (1995) has to be given. I claim that the prepositional complementizer *à* is located, within the C-domain, in Fin°, which reflects certain properties of the verbal system of the clause, such as mood. The relation between the subordinate clause and the DP is established by ForceP. The noun or empty pronoun moves to the Spec of FinP (and from there possibly out of the subordinate clause, cf. fn. 1), but, crucially, leaving Spec,ForceP empty so that another constituent can move through this position:[3]

(79) $[_{\text{ForceP}}$ Qu'$_j$ est-ce qu'il est le seul $[_{\text{ForceP}}$ t_j $[_{\text{FinP}}$ *pro*$_i$ à t_i avoir fait $t_j]]]$?

(80) les $[_{\text{ForceP}}$ sommets$_j$ qu'il a été le seul $[_{\text{ForceP}}$ t_j $[_{\text{FinP}}$ Français$_i$ à t_i atteindre $t_j]]]$

I assume that participles or adjectives followed by a complement are not dominated by functional projections of the Force-Finiteness system. They are simply IPs to the Spec of which the noun or empty pronoun moves. Since there is no ForceP, there is no position through which a constituent can move out of the reduced clause:

(81) * A qui$_i$ Jean est-il le seul $[_{\text{IP}}$ *pro* parlant régulièrement $t_i]$?

(82) * le prix$_i$ qu'elle est la seule $[_{\text{IP}}$ personne fière d'avoir gagné $t_i]$

The full relative clause is also dominated by ForceP, but this time the (empty) noun has to move to Spec,ForceP which contains the complementizer *que* in its head position (after agreement with *pro* in

[3] Extraction is only possible from *à* + infinitive with an "active" meaning, which suggests that *à* + infinitive with a "passive" meaning has a different structure. I suggest that (ii) is ungrammatical, because the "antecedent" noun moves to Spec,ForceP. This blocks extraction of another constituent out of the clause:

(i) C'est la seule chose à demander aux parents.
'It is the only thing to ask the parents.'

(ii) *A qui$_j$ est-ce la seule $[_{\text{ForceP}}$ chose$_i$ à PRO demander t_i $t_j]$?
who is it the only thing to ask

Spec,ForceP in (83) *que* is spelled out as *qui*). This means that there is no empty Spec left that could be used for the extraction of a constituent out of the relative clause:

(83)*Quel rallye$_j$ a-t-il été le premier [$_{ForceP}$ *pro*$_i$ [$_{Force'}$ qui [$_{FinP}$ t$_i$ ait couvert t$_j$]]]?

In the previous section, I showed that whereas subjunctive relative clauses always depend on a real antecedent with a focalized, "negative", interpretation, *à* + infinitive can also be licensed by a focal element within the infinitival clause:

(84) Charles de L'Ecluse reste le personnage du 17ème siècle à avoir le plus contribué à l'introduction de la tulipe en Europe. 'Charles de L'Ecluse remains the person from the 17th century who has most contributed to the introduction of the tulipe in Europe.'

(85) * Charles de L'Ecluse reste le personnage du 17ème siècle qui ait$_{subj}$ le plus contribué à l'introduction de la tulipe en Europe. 'Charles de L'Ecluse remains the person from the 17th century who has most contributed to the introduction of the tulipe in Europe.'

The adoption of a different structure for both types of clauses makes an account of this difference possible in terms of scope differences at LF. Normally, the subjunctive clause and the infinitival clause are the complement of focalized elements such as the superlative, *seul*, etc. and are therefore in their domain. I propose that for speakers who accept (84), an element with a "negative" meaning can also move out of the infinitival clause to a scope position, possibly Spec,FocP dominated by DP (cf. Aboh 2004), at LF and take scope over the clause.

(86) [$_{DP}$ le [$_{FocP}$ le plus$_i$ [$_{ForceP}$ t$_i$ [$_{FinP}$ personnage du 17ème siècle à avoir t$_i$ contribué à l'introduction de la tulipe en Europe]]]].

Since only the infinitival clause contains an extraction site, viz. Spec,ForceP, but full relatives do not, focalized elements can only take scope over and license the infinitival clause, but cannot take scope over the subjunctive clause and hence cannot license it.

7. Conclusion

In this paper I have claimed that DP-internal subordinate clauses, such as French *à* + infinitive clauses occurring with a superlative antecedent, can be

introduced by an infinitival complementizer in FinP. I have argued that both *à* + infinitive clauses and their finite counterparts, subjunctive relative clauses, are complements of an adjective or determiner. Both types of clauses being complements, I have attributed the difference with respect to extraction to a difference in the left periphery of the clause. I have claimed that the prepositional complementizer *à* is located in Fin° attracting the "antecedent" NP to its specifier position and leaving ForceP available for extraction, whereas in full relative clauses the "antecedent" noun raises to Spec,ForceP, blocking extraction. I have proposed that this difference can also account for a difference between the two types of clauses with respect to fronting of the superlative element at LF.

References

Aboh, Enoch. 2004. *The Morphosyntax of Complement-head Sequences: Clause Structure and Word Order Patterns in Kwa*. New York/Oxford, Oxford University Press.

Carlsson, Lennart. 1969. *Le Type 'C'est le Meilleur Livre qu'il ait Jamais Écrit' en Espagnol, en Italien et en Français.* Uppsala, Linquist and Wiksells.

Cinque, Guglielmo. 1994. "On the Evidence for Partial N-Movement in the Romance DP". In G. Cinque, J. Koster, J.-Y. Pollock, and R. Zanuttini (eds.) *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne.* 85-110. Washington D.C., Georgetown University Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. NewYork/Oxford, Oxford University Press.

Fauconnier, Gilles. 1980. *Etude de Certains Aspects Logiques et Grammaticaux de la Quantification et de l'Anaphore en Français et en Anglais*. Paris, Honoré Champion.

Giusti, Giuliana. 2002. "The Functional Structure of Noun Phrases: a Bare Phrase Structure Approach". In G. Cinque (ed.) *Functional Structure in DP and IP*. 54-90. New York/Oxford, Oxford University Press.

Kampers-Manhe, Brigitte. 1991. *L'opposition Subjonctif/Indicatif dans les Relatives*. Amsterdam, Rodopi.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge MA, MIT Press.

Kiss, Katalin. 1998. "Identificational Focus versus Information Focus". *Language*, 74.2:245-273.

May, Robert. 1985. *Logical Form: its Structure and Derivation*. Cambridge MA, MIT Press.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery." In L. Haegeman (ed.) *Elements of Grammar: Handbook of Generative Grammar*. 281-339. Dordrecht, Kluwer.

Sandfeld, Kristian. 1965 (reedition of 1928). *Syntaxe du Français Contemporain I: les Pronoms*. Paris, Honoré Champion.

Siloni, Tal. 1995. "On Participial Relatives and Complementizer D°: a Case Study in Hebrew and French". *Natural Language and Linguistic Theory* 13.3:445-487.

Sleeman, Petra and Els Verheugd. 1998. "Licensing DP-internal Predication". In A. Schwegler, B. Tranel, and M. Uribe-Etxebarria (eds.) *Romance Linguistics. Theoretical Perspectives*. 271-283. Amsterdam/ Philadelphia, John Benjamins.

Zwarts, Frans. 1981. "Negatief Polaire Uitdrukkingen". *Glot* 4:35-132.