Current Issues in Linguistic Theory

Romance Languages and Linguistic Theory 2002

EDITED BY Reineke Bok-Bennema Bart Hollebrandse Brigitte Kampers-Manhe Petra Sleeman

ROMANCE LANGUAGES AND LINGUISTIC THEORY 2002

AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE

General Editor

E.F. KONRAD KOERNER

(Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Berlin)

Series IV - CURRENT ISSUES IN LINGUISTIC THEORY

Advisory Editorial Board

Lyle Campbell (Salt Lake City); Sheila Embleton (Toronto) Brian D. Joseph (Columbus, Ohio); John E. Joseph (Edinburgh) Manfred Krifka (Berlin); E. Wyn Roberts (Vancouver, B.C.) Joseph C. Salmons (Madison, Wis.); Hans-Jürgen Sasse (Köln)

Volume 256

Reineke Bok-Bennema, Bart Hollebrandse, Brigitte Kampers-Manhe and Petra Sleeman (eds)

Romance Languages and Linguistic Theory 2002. Selected papers from 'Going Romance', Groningen, 28-30 November 2002.

ROMANCE LANGUAGES AND LINGUISTIC THEORY 2002

SELECTED PAPERS FROM 'GOING ROMANCE', GRONINGEN, 28–30 NOVEMBER 2002

Edited by

REINEKE BOK-BENNEMA BART HOLLEBRANDSE BRIGITTE KAMPERS-MANHE

University of Groningen

PETRA SLEEMAN
University of Amsterdam

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Romance languages and linguistic theory 2002 : selected papers from "Going Romance," Groningen, 28–30 November 2002 / edited by Relneke Bok-Bennema ... [et al.].

p. cm. -- (Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, ISSN 0304-0763; v. 256)

Includes bibliographical references and index.

1. Romance languages--Congresses. I. Bok-Bennema, Relneke, 1946- II. Series.

PC11.R833 2004

440--dc22

2004046449

ISBN 90 272 4770 6 (Eur.) / 1 58811 585 2 (US) (Hb; alk. paper)

© 2004 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 36224 • 1020 ME Amsterdam • The Netherlands John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

INTRODUCTION

Going Romance is a major European annual discussion forum for theoretically relevant research on Romance languages; it is an international initiative of the Dutch university community involved in research on Romance languages. The Romance Languages and Linguistic Theory series consists of a selection of papers of the Going Romance conferences, which have been organized by and held at the various universities of the country. The first volume contained the selected papers of the thirteenth conference, held in 1999.

This is the fourth volume of *Romance Languages and Linguistic Theory*. The articles form a selection of the papers that have been presented at the occasion of *Going Romance 2002 (XVI)* - which was held at the State University of Groningen on November 28 through November 30. The three day program included a workshop on Acquisition. The volume contains articles on specifics of one or more Romance languages or varieties: clausal structure, verb-movement, topic, focus and reinforcement constructions, nominal ellipsis, (absence of) pronouns in child language, and other interesting phenomena.

The editors would like to thank everyone who contributed to the success of *Going Romance XVI*. Next to the editors, the organization committee consisted of Frank Drijkoningen (Utrecht, UiL OTS), Angeliek van Hout (Groningen, CLCG), Aafke Hulk (Amsterdam, HIL), Haike Jacobs (Nijmegen, CLS) and Marianne Starren (Nijmegen, CLS).

The selection committee for the more than fifty abstracts for the main session and the workshop consisted of editors and organizers and was assisted by the invited speakers Heles Contreras (Washington), Celia Jakubowicz (Paris), Marie-Thérèrse Vinet (Sherbrooke) and Alessandro Zucchi (Salerno).

The organizers and the editors gratefully acknowledge financial support from the the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO), the Groningen University Fund (GUF), the Centre of Behavioral and Cognitive Neurosciences (BCN, Groningen), the Centre for Language and Cognition, Groningen

(CLCG) and the Department for Romance Languages and Cultures (RTC) of the State University of Groningen.

Reineke Bok-Bennema Bart Hollebrandse Brigitte Kampers-Manhe Petra Sleeman

CONTENTS

Tense, Quantification and Clause Structure in EP and BP. Evidence from a Comparative Study on Sempre Manuela Ambar, Manuela Gonzaga & Esmeralda Negrão			
Early 'Pragmatic' Competence and its Implications regarding the Null Subject Phenomenon Cécile de Cat	17		
On the Impact of French Subject Clitics on the Information Structure of the Sentence <i>Cécile de Cat</i>	33		
A Restricted View of Head Movement Heles Contreras	47		
The Effects of Phonological Cues on the Syntax of Focus Constructions in Spanish Laura Dominguez	69		
Optional Infinitives or Silent Auxes? New Evidence from Romance Cristina Dye	83		
Dislocation, Clitic Resumption and Minmality. A Comparative Analysis of Left and Right Topic Constructions in Italian <i>Mara Frascarelli</i>	99		
Focus on Negative Concord Daniela Isac	119		
Romance and 'Something Else' Daniela Isac & Charles Reiss	141		
Learnability Order in the French Pronominal System Jacqueline van Kampen	163		

viii CONTENTS

Nominal Ellipsis and Morphological Structure in Spanish Laura Kornfeld & Andrés Saab		
The Development of Inalienable Possession in English and Spanish Ana T. Pérez-Leroux, Cristina Schmitt & Alan Munn	199	
Segmental and Featural Strategies to Avoid Adjacent Sibilant Segments in Balearic Catalan. An Optimality Account Clàudia Pons Moll	217	
-Tu in Quebec French as a (Super)positive Marker Marie-Thérèse Vinet	235	
Feature Checking and Object Clitic Omission in Child Catalan and Spanish Ken Wexler, Anna Gavarró & Vincenç Torrens	253	
Index of languages and dialects	269	
Subject Index	270	

TENSE, QUANTIFICATION AND CLAUSE STRUCTURE IN EP AND BP EVIDENCE FROM A COMPARATIVE STUDY ON SEMPRE*

MANUELA AMBAR, MANUELA GONZAGA¹ & ESMERALDA NEGRÃO U. Lisboa U. Lisboa-CLUL U. São Paulo

1. Main Goal

Various authors have pointed out crucial structural differences between European Portuguese and Brazilian Portuguese (henceforth EP and BP) regarding several aspects of their syntax (Ambar & Veloso 1999, Barbosa, Oliveira & Müller 2001, Brito 1999, Costa & Figueiredo Silva 2001, E. Duarte 1995, I. Duarte 1997, 2000, Galves 1993, 1998, Kato & Raposo 1996, Matos & Cyrino 2001, Peres & Negrão 2001, and Nunes & Martins 2001, a.o.).

The main goal of this paper is to derive the distribution of the adverb *sempre* ("always", "really/indeed"), two crucial different interpretations associated with it and the variation between EP and BP from core structural properties of the clause in the two languages.²

*

^{*} The results presented in this paper are part of a research project designed to compare the Syntax and the Semantics of European and Brazilian Portuguese, coordinated by Mary A. Kato and João Andrade Peres, funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brasil) and Instituto de Cooperação Científica e Tecnológica Internacional (ICCTI-Portugal).

¹ The presentation of this paper and the work developed for it were supported by FCT (Ph.D. research - Ref. BD/5252/2001).

² As will be clearer in what follows, in this paper we concentrate on two readings – temporal and confirmative/assertive. Other interpretations exist however – either related to the confirmative reading or to the temporal one or to different scope effects. Limitations of space preclude the treatment of the latter in this paper as well as the study of other adverbial-like elements that, depending on their distribution, may provide different readings (point of view, focus, inclusiveness), e.g. *mesmo*. For example, this last adverb in its confirmative reading coexists with confirmative *sempre* (in EP, though not in BP) and even mirrors the position of *sempre* w.r.t. the verb (cf. orders *sempre*-V, *V-*sempre vs.* **mesmo*-V, V-*mesmo*, for the confirmative reading). We also limit our analysis to finite and non-periphrastic contexts. Thus, the question raised by a reviewer on the behavior of *sempre* in infinitives cannot be fully answered here. Infinitives have however been considered and support our proposal where the status of Tense and Agr play an important role. We just observe that *sempre* behaves

In explaining the puzzling behavior of *sempre*, which is tied to a set of other phenomena, the paper sheds light on some of the parameters responsible for the variation between the two languages.

2. The problem

The first striking difference between the two languages in regard to the behavior of *sempre* involves its interpretation. In EP, two interpretations are available for *sempre*: (i) a confirmative interpretation, in which case *sempre* means something like "really" or "indeed", and (ii) a temporal interpretation similar to "always". In BP, on the other hand, the confirmative reading of *sempre* does not exist, the temporal reading being the only one available. The glosses of the sentence in example (1) show the contrast between the two languages:

The two main readings available in EP are dependent on two factors: 1) the position occupied by *sempre* in relation to the verb; 2) the morphological tense of the clause. The confirmative reading, only available in EP, appears when *sempre* occurs in pre-verbal position either with Present or Past (Gonzaga 1997):

As already pointed out, sentence (2) in BP does not have a confirmative reading. The only interpretation associated with it is the temporal interpretation, as shown in (3):

(3) O João <u>sempre</u> vai/foi para Paris de trem. the John 'always' / * 'really/indeed' goes/went to Paris by train (BP)

differently according to the different status of Tense and Agr in infinitival structures. Actually, confirmative *sempre* is available in inflected infinitives complement of epistemic and declarative verbs, but unavailable in obligatory control or exceptional Case marking contexts – but see Negrão, Gonzaga & Ambar (forthcoming).

The two languages also differ in regard to the distribution of temporal *sempre*. In BP, the pre-verbal position of *sempre* is the preferred order, regardless of the morphological tense of the clause (Present or Past), as shown in (4):

(4) a. O João <u>sempre</u> vai para Paris de trem. the John always goes to Paris by train (BP)

b. O João <u>sempre</u> foi para Paris de trem. the John always went to Paris by train (BP)

In EP, on the other hand, the preferred order for the temporal *sempre* adverb is the post-verbal position:

(5) a. *O João vai <u>sempre</u> a Paris de comboio*. the John always goes to Paris by train (EP)

b. *O João foi <u>sempre</u> a Paris de comboio*. the John goes always to Paris by train (EP)

Sentences in the Present tense with temporal *sempre* in pre-verbal position are rejected by most EP speakers:

Comparing the temporal reading available in both positions, some subtle contrasts show up in both languages. The pre-verbal temporal *sempre* in Past clauses conveys a reading expressing a universal quantification over the event argument of the predicate both in BP and EP. The interpretation associated with (7b) can be expressed as something like: "In all the events of drinking that took place in his life, John drank in them". This universally quantified reading is also available for *Present* clauses when *sempre* is in the pre-verbal position in BP, though not in EP (cf. (4) above for BP and (7a) below):

(7) a. *O João <u>sempre</u> bebe.* (*EP, BP) the John always drinks

b. *O João <u>sempre</u> bebeu*. (EP, BP) the John always drank

The post-verbal temporal *sempre* of EP Present and Past sentences activates a correlation of events interpretation. In order to establish this correlation, more information is needed:

Sentence (8) has the interpretation that correlates the events of drinking wine to the events of having a meal: "For all the events of John having a meal there is an event of him drinking wine". The absence of lexical material offering another event, through which *sempre* could establish a correlation with the event of the main predicate, causes oddness to the sentence, as observed in (9), unless *sempre* is given a different intonation, as in (10):

Sentence (10) has a reading closer to the universally quantified reading, denoting the whole set of temporal events associated with John, that is, John drank his whole life, John is a drinker.

In BP, the correlation of events reading is obtained when *sempre* occupies the pre-verbal position in both Present and Past sentences. When in post-verbal position, *sempre* elicitates a reading expressing a pattern of behavior. Compare the following sentences:

Sentence (11a) can be paraphrased by the following sentence: "Whenever John buys/bought books, he does/did it at FNAC". Sentence (11b), on the other hand, can be paraphrased as: "John buys/bought books regularly at FNAC".

Evidence in favor of the correlation between this *pattern* special meaning and the position of the adverb in BP is the strangeness of sentence (12b) when compared to sentences (12a) and (12c):

In sentence (12a), in which *sempre* appears in post-verbal position, the interpretation is that John exercised himself with regularity during this year. It can even be inferred that he exercised himself a lot. Sentence (12c), on the other hand, has the interpretation that whenever John ran this year, he ran two miles, which does not necessarily mean that he ran very often. Quite the opposite, it can even be true that he did that only twice this year.

Another fact to be pointed out is that the determination of the object with event verbs interacts with the distribution and interpretation of *sempre*. Arguments introduced by a definite determiner cause ungrammaticality because they force a single punctual event reading, incompatible with the [+distributive-universal] feature correlating events, which characterizes the adverb *sempre*:

d. *O João <u>sempre</u> come/comeu bolos*. (EP*Present, BP) the John always eats/ate cakes

3. Our analysis 4

In this section we argue that the differences observed in the distribution and interpretation of the adverb *sempre* in EP and BP correlate with some well-known facts of the grammar of the two languages.

Linguists working with the syntax of BP (Galves 1993, Figueiredo Silva 1996, Duarte 1995, a.o.) have shown that this language is going through a parametric change. A drastic restructuring of the pronominal paradigm and a

³ Note that sentence (13c) is grammatical, in EP, in the confirmative reading, where such restrictions do not operate.

⁴ For a different analysis of *sempre* in both languages see Brito (1999).

weakening of the verbal inflectional paradigm of BP have been attested: the second person subject pronouns, both singular and plural (tu and $v\delta s$), and the first person plural subject pronoun ($n\delta s$) have been substituted by forms that trigger third person agreement ($voc\hat{e}/voc\hat{e}$ and a gente). Negrão (1999) and Negrão & Viotti (2000) argue that a change in the way the basic predicative relation is established in the sentences of this language has weakened the role played by the inflectional markers. For our purposes here in this paper, this change amounts to saying that the AgrSP functional projection is weak in BP as opposed to the AgrSP strong projection of EP.

In order to develop our analysis, besides the functional and lexical projections traditionally considered in the literature (*NegP*, *AgrSP*, *TsP*, *VP*), we will assume the following projections to be necessary:

AssertiveP (Ambar 1999) – it is located above WhP and projects whenever confirmative features associated to the speaker attitude must be checked.

Searle (1979) defines "assertive" as "the illocutory act where we tell our hearers (truly or falsely) how things are".

Distributive-Universal Phrase (DistP) – according to Beghelli & Stowell (1997), it is a functional projection headed by Dist^o, where the distributive operator resides, to whose specifier position (Spec) operators and quantified phrases functioning as distributive universals like every man move in order to check their [+distributive-universal] feature (Beghelli & Stowell 1997, Negrão 1999, 2001). Following Choe (1987), Beghelli (1995) takes distributivity to be a binary relation that requires a distributor and a distributee (in Choe's terms, sorting key and distributed share, respectively). In our analysis, this projection will be responsible for deriving the correlation of events reading.

Tense Object Phrase (TobjP) – according to Ambar (1996, 1998), it is a projection intended to account for the restrictions on clauses imposed by the Aktionsart of predicates and by the quantificational properties of the complements.

The following structure shows the hierarchical distribution of those functional projections:

As already pointed out, *sempre* is a frequency or pattern adverb (Vlach 1993:237): "The notion of frequency or pattern is in the nature of processes; they consist of some pattern of distribution of one or more eventuality types over an interval".

We assume that *sempre* is generated in the head position of TobjP in as much as it is related to the characteristics of the Aktionsart of the predicate of the sentence. It has a [+distributive-universal] feature and it may have a [+assertive] feature.⁵

_

⁵ A reviewer notes that this possibility of having an assertive feature conveys to the analysis the assumption that we have in fact two different lexical entries for /sempre/, an option that we would not like to assume. The same reviewer also proposes an interesting alternative analysis which would consist in considering that sempre can be either X° or XP and is always a universal quantifier. As X° it would be generated in Spec of a functional projection (FP) c-commanding TP, therefore quantifying over Tense. This would remove the 'double lexical listing' of sempre. Evidence for the X°/XP opposition is drawn by the reviewer from modification with nem (not) and quase (almost), which is possible for the temporal reading but excluded for the confirmative one. As observed in section 3.1, we agree with the reviewer w.r.t. the fact that universal quantification is involved in the confirmative reading – let us take it as quantification over 'point of view' located in AssertiveP ("indeed", "by all means", "in all aspects") as the reviewer also suggests. However, we are not sure about the X°/XP opposition. Notice that either in its confirmative reading (corresponding to the reviewer's X°) or in its temporal reading (reviewer's XP), nothing but clitics can intervene between sempre and V, suggesting that they form a complex head. The behaviors of nem and quase are complex – on

3.1 *Confirmative* sempre

In EP, when tense is strong, as is the case of Past, the verb raises to T° to check the strong V feature; then it raises to AgrS° to check the strong D feature of INFL. If tense is weak, as is the case of Present (tense *zero* – Giorgi & Pianesi 1991, Kayne 1993), the verb raises to AgrS° directly. That is, in EP, the verb always raises to AgrS° and, in the case of Present, it does not pass through T° (cf. Nash & Rouveret 1998). In BP, in contrast, due to the parametric change taking place in its syntax, as already pointed out, the verb raises as high as TsP.

In turn, the frequency/pattern adverb *sempre* raises from the head position of TobjP to the head position of DistP, a projection which attracts operators and quantified phrases carrying the [+distributive-universal] feature. In checking this feature, it elicitates the universal quantified reading associated to *sempre*. In as much as *sempre* raises to Dist^o in both languages, it is the difference in the movement of the verb that explains why the preferred order is *sempre* in the pre-verbal position in BP and *sempre* in post-verbal position in EP.

The quantified universal reading is involved in the confirmative interpretation. First, *sempre* raises to DistP to check [+distributive-universal]. The confirmative *sempre* also has a [+assertive] feature. AssertiveP projects and its feature has to be checked by *sempre*. When *sempre* raises, the verb is in AgrS°, *sempre* adjoins to V°. On its way to AssertiveP, it 'pied-pies' the verb. *Sempre* lands in Assertive°, carrying the adjoined verb with it. The subject DP, in turn, moves to the specifier position of TopP⁶. The structure in (15) shows the derivation of the confirmative reading of *sempre*:

the one hand nem involves values that seem to be also related to 'point of view' (evaluative in some sense), and on the other hand while nem seems to have scope over XPs (over an NP, for example: $Nem\ o\ João\ saiu$ "Not even John left"), quase is submitted to other types of restrictions (* $Quase\ o\ João\ saiu$ "Almost John left"), impossible with scope over the NP $o\ João\$). Clearly there are restrictions on the assertive reading imposed by negation, e.g. $nem\$ cannot combine with $sempre\$ in this case – but different hypotheses may be considered, e.g. intervention effects, c-commanding relations, minimal link condition. Anyway, in work in progress, where negation is being studied in its relation with $sempre\$, the reviewer's suggestion will be taken into consideration. Moreover, coming back to the X°/XP opposition we still would have to explain what would attract X° to the higher domain – if it is just quantification over the 'point of view', our analysis can then also dispense with the [assertive] feature on $sempre\$ and the problem of 'double lexical listing' would be solved.

⁶ As a reviewer notes, it is not clear whether the subject should be in Spec,AssertiveP or in TopicP. For the time being we are assuming it ends up in TopicP given the bad results provided by topicalization of quantified DPs in pre-verbal *sempre* contexts, either in the confirmative or in the temporal reading (but see the discussion about quantifier floating, section 4.2).

(15) Confirmative

 $\begin{bmatrix} T_{ODP} O \ Jo\~{a}o_1 [A_{SSertiveP} \ sempre_2 - foi_v \ [W_{hP} \ [F_{OCUSP} \ [A_{grSP} \ t_1 \ t_{2v} \ [D_{IStP} \ t_2 \ [T_{SP} \ t_v \ [T_{ObjP} \ t_2 \ [v_P \ t_1 \ t_v \ a \ Paris]]]]]]]]] \end{bmatrix}$

The order *sempre*+verb is the only one possible with the confirmative reading, as shown in (16). An adverb such as *ontem* "yesterday", breaking the adjacency between the adverb *sempre* and the verb, causes the ungrammaticality of sentence (16a):

- (16) a. *Ele <u>sempre</u> ontem foi a Paris.
 he really/indeed vesterday went to Paris
 - b. *Ontem* ele <u>sempre</u> foi a Paris yesterday he really/indeed went to Paris
 - c. *Ele <u>sempre</u>* foi ontem a Paris. He really/indeed went yesterday to Paris
 - d. *Ele <u>sempre</u>* foi a Paris ontem. He really/indeed went to Paris yesterday

Only clitics, or other adjoined heads, can appear between *sempre* and the verb:

- (17) a. *Ele <u>sempre</u> lhe disse isso*. he really/indeed him said that
 - b. ?*O João <u>sempre</u> já chegou*. the John really/indeed already arrived

In AssertiveP, *sempre* has scope over the whole proposition; this gives to the confirmative interpretation the value of singularity of the situation that is confirmed by the speaker ("the single of all").

In BP, the confirmative interpretation of *sempre* is impossible because, due to the typological changing of this language in the recent past, the verb stops in T^0 . It does not have access to AgrSP, the locus of the adjunction.

3.2 Universal Temporal <u>sempre</u>

As claimed above, the two main differences between EP and BP are: the absence of the confirmative reading in BP, already explained in 3.1, and the different preferred orders of temporal *sempre*: post-verbal in EP and pre-verbal in BP, as shown in (18) and (19), (presented in (11) for BP and repeated here for convenience):

- (18) a. O João compra <u>sempre</u> livros na FNAC. (EP preferred), the John buys always books at FNAC (BP non-preferred) b. O João comprou <u>sempre</u> livros na FNAC. (EP preferred), the John bought always books at FNAC) (BP non-preferred)
- (19) a. *O João <u>sempre</u> compra livros na FNAC*. (BP preferred ,*/??EP) the John always buys books at FNAC
 - b. *O João <u>sempre</u> comprou livros na FNAC*. (BP preferred), the John always buys books at FNAC (EP *sempre* has focus)

In EP, when tense is strong, as in the case of Past clauses, the verb raises to T^{o} to check its strong V feature; then it raises to $AgrS^{0}$ to check the strong D feature of INFL. If tense is weak, as in the case of Present clauses, the verb raises to $AgrS^{0}$ directly. In BP, on the other hand, the fact that agreement has become weak in BP blocks the movement of the verb to $AgrS^{0}$.

The frequency/pattern adverb *sempre*, in turn, raises from the head position of TobjP to the head position of DistP to check its [+distributive-universal] feature, elicitating the universal quantified reading associated to *sempre*. The verb in AgrS° and *sempre* in Dist° generate the post-verbal position of *sempre* in EP, as shown in (20):

(20) EP post-verbal present (universal temporal reading)

In BP, the verb is always in T^o and *sempre* is in Dist^o, resulting in the preferred pre-verbal position of *sempre* (universal temporal reading).

(21) BP pre-verbal present/past (universal temporal reading)

 $[\begin{smallmatrix} T_{OPP} \ O \ Jo\~{a}o_1 \end{smallmatrix} [\begin{smallmatrix} AssertiveP \ [WhP \ [FocusP \ [AgrSP \ [DistP \ sempre_2 \ [TSP \ t_1 \ compra/-ouv \ [TobjP \ t_2 \ [vP \ t_1 \ t_v \ livros \ na \ FNAC]]]]]]]]]]$

3.3 Past and correlation of events

The universal temporal reading associated with *sempre* is not the only temporal reading available. A correlation of events interpretation is also possible in sentences containing the adverb *sempre*. In BP, a sentence like (22):

(22) O João <u>sempre</u> bebeu vinho às refeições. the John always drank wine during meals may have two interpretations. The first can be paraphrased as something like "All his life, John drank wine during meals". This interpretation corresponds to the universal temporal interpretation. The second interpretation, as already pointed out, can be paraphrased as something like: "For all the events of John having a meal there is an event of him drinking wine". This interpretation corresponds to the correlation of events interpretation.

In EP, when the verb passes through T^0 and goes to $AgrS^0$, and sempre moves to $Dist^0$ generating the post-verbal position of the temporal operator, the quantificational universal reading is lost. Only the correlation of events interpretation remains available. Subsequent movement of sempre to Focus reestablishes a configuration in which sempre has scope over tense, necessary for the availability of the universal reading. At the same time, the prosody effects observed with sempre in pre-verbal position are derived. Sentence (23), whose structure is presented in (24), exemplifies this focus movement in EP:

(23) O João SEMPRE comprou livros na FNAC. the John always bought books at FNAC

(24) EP pre-verbal past (temporal universal reading)⁷

 $\begin{array}{l} [T_{OPP} \ O \ Jo\~{a}o_1 \ [A_{SSertiveP} \ [W_{hP} \ [F_{OcusP} \ sempre_2\text{-}comprou_v \ [A_{grSP} \ t_1 \ t_{2v} \ [D_{istP} \ t_2 \ [T_{SSP} \ t_v \ [T_{ObjP} \ t_2 \ [v_P \ t_1 \ t_v \ livros \ na \ FNAC]]]]]]]]] \end{array}$

A final point is needed to explain the post-verbal non-preferred order in BP, as exemplified in sentence (12), repeated here as (25) for convenience:

(25) *O João compra/comprou sempre livros na FNAC*. the John buys/bought always books at FNAC

As already shown, this sentence has the preferred *pattern* interpretation⁸: "John buys/bought regularly books at FNAC". This interpretation emerges when *sempre* stays in T^oobj position. In other words, having only VP under its scope, the BP post-verbal *sempre* acquires a *pattern* interpretation.

⁷ For an analysis of *sempre* as an adverb that can be focalized in EP see Martins (1994).

⁸ Another interpretation, in which *sempre* takes *livros* "books" as the element in its scope, has an interpretation that can be paraphrased as: "When John buys at FNAC, he buys books". We will not take this interpretation into consideration.

4. Other predictions

In this section we analyze some data involving extraction and quantifier floating in order to present some further evidence for the analysis developed in this paper, which accounts for the difference in distribution and interpretation of *sempre* in EP and BP.

4.1 Wh-extraction

The first set of data to be discussed are sentences containing *sempre* with the confirmative reading in EP. Sentence (26a) is a declarative sentence in the present tense with *sempre* in the pre-verbal position, elicitating a confirmative reading. If, from sentence (26a), we build an interrogative sentence by extracting the locative argument, now a wh-word, we generate an ungrammatical sentence as (26b), whose structure is given in (26b'):

- (26) a. *O João <u>sempre</u> vai a Paris*. (EP/*BP) the John really/indeed goes to Paris
 - b. *Aonde vai o João <u>sempre</u>? where goes John really/indeed?
 - b.' [TopP [AssertiveP [WhP Aonde₃ [Who vaiv] [FocusP [AgrsP O João₁ sempre₂-tv [DistP t₂ [Tsp tv [TobiP t₂ [vP t₁ tv t₃]]]]]]]]]

The reason for this ungrammaticality is that the movement of a wh-word to the specifier position of the Wh-projection requires raising of the verb to the head position of this projection. Since the confirmative reading results from movement of the complex *sempre*+verb to the Assertive projection, if the verb raises by itself to Wh^o, this movement blocks the movement of the complex. Against the explanation given for the ungrammaticality of sentence (26b), it could be said that the confirmative reading clashes with the interrogative meaning of interrogative sentences. Although the clash in meaning does in fact take place, it cannot be the sole explanation in as much as interrogative sentences with an echo flavor, which do not break the complex *sempre*+verb, are possible EP sentences:

- (27) a. *Aonde é que o João <u>sempre</u> vai*? (EP echo confirmative, *BP) where is that the John really/indeed goes
 - b. Aonde é que <u>sempre</u> vai o João? where is that really/indeed goes the John
 - c. *Aonde é que <u>sempre</u> o João va? where is that really/indeed the John goes?

4.2 Quantifier Floating

Data involving quantifier floating can also be taken as evidence for the analysis proposed in this paper. The paradigm in (28) exhibits the pattern of grammaticality of EP sentences containing confirmative *sempre*, formed from sentence (28a).

- (28) a. Os meninos <u>sempre</u> telefonaram aos amigos.

 The kids really/indeed called their friends
 - b. *Todos os meninos <u>sempre</u> telefonaram aos amigos. all the kids really/indeed called their friends
 - c. Os meninos <u>sempre</u> telefonaram todos aos amigos.

 The kids really/indeed called all their friends
 - d. *Os meninos <u>sempre</u> todos telefonaram aos amigos.

 The kids really/indeed all called their friends
 - e. *Os meninos todos <u>sempre</u> telefonaram aos amigos.

 The kids all really/indeed called their friends
 - b.' $[T_{opP} Todos os meninos_1[AssertiveP] sempre_2$ -telefonaram_v $[W_{hP} [F_{ocusP} [AgrSP t_1 t_{2v} [DistP t_2 [TSP t_v [T_{objP} t_2 [v_P t_1 t_v aos amigos]]]]]]]]]]$
 - c.' [TopP Os meninos₁[AssertiveP sempre₂-telefonaram_v [WhP [FocusP [AgrSP t₁ t₂ [DistP t₂ [TSP t_v [TobjP t₂ [vP todos t₁ t_v aos amigos]]]]]]]]]
 - d.' $[T_{opP} Os meninos_1 [A_{ssertiveP} sempre_2 [W_{hP} [F_{ocusP} [A_{grSP} todos t_1 t_2 telefonaram_v [D_{istP} t_2 [T_{SP} t_v [T_{objP} t_2 [v_P t_1 t_v aos amigos]]]]]]]]]]$

To understand the (un)grammaticality of the sentences in (28), it is necessary to remember that the confirmative reading is generated when the complex *sempre*+verb, formed by head adjunction in AgrSP, raises to the head position of the Assertive Phrase. In (28b), in order to obtain the right word order, we would have to say that the Quantifier Phrase *todos os meninos* has moved to the specifier position of TopP, given the raising of the complex already mentioned above. Quantifier Phrases do not land in Topic projections, due to semantic incompatibilities. Therefore, sentence (28b) is ungrammatical (cf. structure 28b'). On the other hand, this incompatibility does not arise, if only the restriction of the quantifier phrase, a definite expression, moves to TopP. This explains the grammaticality of (28c). Sentence (28d) is ungrammatical because the quantifier *todos* intervenes between the complex *sempre*+verb.

5. Concluding remarks

In this paper we offered an explanation for the differences in distribution and interpretation of the adverb *sempre* attested in European Portuguese and

Brazilian Portuguese. We claimed that those differences are due to a parametric change taking place in the latter language.

References

- Ambar, Manuela. 1996. "The Syntax of Focus a Unified Approach". Paper presented at the *International Workshop on Focus*, University of Lisbon.

- & Rita Veloso. 1999. "On the Nature of Wh-phrases Word Order and Wh-in-Situ: Evidence from Portuguese, French, Hungarian and Tetum". Romance Languages and Linguistic Theory 1999: Selected Papers from 'Going Romance' 1999 ed. by Yves D'Hulst, Johan Rooryck & Jan Schroten, 1-37. Amsterdam & Philadelphia: John Benjamins.
- Manuela Gonzaga & Esmeralda Negrão. (forthcoming). "A Comparative View of European and Brazilian Portuguese Phrase Structure: Evidence from Extraction, Quantification and Negation".
- Barbosa, Pilar, Fátima Oliveira, & Ana Müller. 2001. "Nomes 'Simples' em Posição de Sujeito: Questões Sintáticas e Semânticas". Paper presented at the *EP and BP Colloqium*, Fortaleza.
- Beghelli, Filippo. 1995. *The Phrase Structure of Quantifier Scope*. Ph.D. dissertation, University of California at Los Angeles.
- & Tim Stowell. 1997 "Distributivity and Negation: The Syntax of *Each* and *Every*". *Ways of Scope Taking* ed. by Anna Szabolcsi, 71-109. Dordrecht: Kluwer.
- Brito, Ana-Maria. 1999. "Clause Structure, Subject Positions and Verb Movement. About the Position of *Sempre* in European Portuguese and Brazilian Portuguese". *Romance Languages and Linguistic Theory 1999: Selected Papers from 'Going Romance' 1999* ed. by Yves D'Hulst, Johan Rooryck & Jan Schroten. 63-85. Amsterdam & Philadelphia: John Benjamins.
- Choe, Jae–Woong. 1987. *Anti-Quantifiers and a Theory of Distributivity*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Costa, João & Cristina Figueiredo Silva. 2001. "Ordem VS e Sujeito Nulo em Português Europeu e em Português Brasileiro". Paper presented at the *EP and BP Colloqium*, Fortaleza.

- Duarte, Eugênia. 1995. *A Perda do Princípio 'Evite Pronome' no Português Brasileiro*. Ph.D. dissertation, Campinas, SP, UNICAMP.
- Duarte, Inês. 1997. "Ordem de Palavras. Sintaxe e Estrutura Discursiva". Sentido que a Vida Faz. Estudos para Óscar Lopes ed. by Ana-Maria Brito, Fátima Oliveira, Isabel Pires de Lima & Rosa Maria Martelo, 581-592. Porto: Campo das Letras.
- —— 2000. "Sobre Interrogativas-Q em Português Europeu e Português Brasileiro". Ms., University of Lisbon.
- Figueiredo Silva, Cristina. 1996. *A Posição Sujeito no Português Brasileiro: Frases Finitas e Infinitivas*. Campinas: Editora da UNICAMP.
- Galves, Charlotte. 1993. "O Enfraquecimento da Concordância no Português Brasileiro". *Português Brasileiro: uma Viagem Diacrônica* ed. by Ian Roberts & Mary Kato, 387-425. Campinas: Editora da UNICAMP.
- Giorgi, Alessandra. & Fabio Pianesi. 1992. "For a Syntax of Tense". Ms., University of Catania and IRST, Povo (Trento).
- Gonzaga, Manuela. 1997. Aspectos da Sintaxe dos Advérbios em Português. M.A. dissertation, University of Lisbon.
- Kato, Mary & Eduardo Raposo. 1996. "European and Brazilian Portuguese Word Order: Questions, Focus and Topic Constructions". *Aspects of Romance Linguistics* ed. by Claudia Parodi, Carlos Quicoli, Mario Saltarelli & María Luisa Zubizarreta, 267-277. Washington: Georgetown University Press.
- Kayne, Richard. 1993. "Toward a Modular Theory of Auxiliary Selection". *Studia Linguistica* 47.3-31.
- Martins, Ana Maria. 1994. *Clíticos na História do Português*. Ph.D. dissertation, University of Lisbon.
- Matos, Gabriela & Sônia Cyrino. 2001. "Elipse de SV no Português Brasileiro e no Português Europeu". Paper presented at the *EP and BP Colloqium*, Fortaleza.
- Nash, Lea & Alain Rouveret. 1997. "Proxy Categories in Phrase Structure Theory". *Proceedings of NELS* 27 ed. by Kiyomi Kusumoto, 287-304. University of Massachusetts, Amherst: GLSA.
- Negrão, Esmeralda. 1999. *O Português Brasileiro: Uma Língua Voltada para o Discurso*. Ph.D. dissertation, University of São Paulo.
- —— 2001. "The Scopal Properties of Distributive Quantifiers in Brazilian Portuguese". *The Proceedings of SULA. The Semantics of Under-Represented Languages in America* ed. by Ji-Yung Kim & Adam Werle, 81-85. Amherst, Massachusetts.

- & Evani Viotti. 2000. "Brazilian Portuguese as a Discourse-Oriented Language". *Brazilian Portuguese and the Null Subject Parameter* ed. by Mary Kato & Esmeralda Negrão, 97-116. Frankfurt: Vervuert.
- Nunes, Jairo & Ana Maria Martins. 2001. "Infinitivos Obrigatoriamente Flexionados em Ambientes de Opcionalidade". Paper presented at the *EP and BP Colloqium*, Fortaleza.
- Peres, João & Esmeralda Negrão. 2001. "Algumas Propriedades das Construções Negativas do Português Brasileiro e do Português Europeu". Paper presented at the *EP and BP Colloqium*, Fortaleza.
- Searle, John R. 1979. Expression and Meaning: Studies in the Theory of Speech Acts. Cambridge: Cambridge University Press.
- Vlach, Frank. 1993. "Temporal Adverbials, Tenses and the Perfect". Linguistics and Philosophy 16.231-283.

EARLY 'PRAGMATIC' COMPETENCE AND ITS IMPLICATIONS REGARDING THE NULL SUBJECT PHENOMENON*

CECILE DE CAT University of York

1. In a nutshell

It is commonly assumed in language acquisition research that discourse competence belongs to a pragmatic module which does not form an integral part of Universal Grammar (UG) (e.g. Chien & Wexler 1991) and that the relevant competence matures alongside cognitive development. In this paper, I present clear evidence that from the onset of word combinations, children master the basic discourse notion of topic. This evidence goes against the assumption that children at the null subject stage lack the 'pragmatic' competence necessary to identify and encode topics in a target-like fashion (cf. e.g. Wexler 1998; Schaeffer et al. 2002)— an assumption based on the study of languages such as English, where the structuring of information has little impact on the syntactic form of sentences.

2. Background

2.1. *On topics*

In the vast literature on topics, a confusing number of definitions are proposed. The one that will be adopted for the purpose of this article is as general and wide-ranging as possible, to reflect the fact that the phenomenon under study is most probably universal, although different languages resort to

^{*} This research was supported by the Economic and Social Research Council (grants R00429834373 and 041R00433), which is gratefully acknowledged. It was presented in various forms of completion to the acquisition groups of Groningen and Utrecht, the 2002 Boston University Conference on Language Development and the Going Romance 2002 workshop on language acquisition. I wish to thank the audiences for comments and discussion (with special thanks to Aafke Hulk, Bernadette Plunkett, Kamil Ud Deen and an anonymous reviewer).

different means of encoding it. Topics will be taken to express what the sentence is about (following Reinhart 1981), hence the label aboutness topic.¹

Crosslinguistically, aboutness topics strongly tend to be encoded as dislocated elements (Gundel 1975; Reinhart 1981; Rizzi 1997). In all the examples below, dislocated elements are separated from the rest of the sentence by a comma.

(1) Les mures_i, elles_i sont mures.² the blackberries they are ripe "The blackberries are ripe."

Aboutness topics correspond in most cases to the subject of the sentence/clause (Gundel 1975; Li & Thompson 1976; Givón 1976; Reinhart 1981; Lambrecht 1994). A native speaker of English, if presented with a sentence like (2), will by default interpret the subject (here *your picture*) as the topic. In other words, the context in which a sentence like (2) is felicitous is by default one in which the picture in question is salient (and is being talked about).

(2) Your picture scared me.

Thetic sentences are the only sentence type that is incompatible with an aboutness topic. Such sentences express a state of affairs and are typically uttered out-of-the-blue or after a question like *What happened?*. In thetic sentences, a topic interpretation of e.g. the subject or the object is impossible — hence the dislocation of these elements is not allowed in spoken French. This will be illustrated in the following section.

2.2. Topic encoding in (adult) spoken French

The association between dislocated elements and topic interpretation is

¹ This is intended to avoid confusion with another kind of topic which needs not concern us here. The reader is referred to De Cat (2002) for a definition of *stage topics* (based on Erteschik- Shir 1997) and its manifestations in spoken French. In De Cat (2002), I argue in favour of an overarching definition of topichood in terms of *frame within which the predication is evaluated* (following Erteschik-Shir 1997). The point of the present paper, i.e. that children appear to have the required competence to encode topics from the onset of the multiword stage, holds for topics in general, though it is only explicitly made here with respect to aboutness topics.

² I adopt the new French spelling, proposed by the *Conseil supérieur de la langue française*. Details can be found at http://www.fltr.ucl.ac.be/FLTR/ROM/vdm.html.

extremely robust in spoken French: (heavy) topics are obligatorily dislocated, and dislocated elements are obligatorily interpreted as topics.

- 2.2.1. Heavy topics are obligatorily dislocated In spoken French, topics are obligatorily (left- or right-)dislocated when they are not expressed by a weak pronoun only (De Cat 2002). Weak pronouns include clitics and ca 'it' when unstressed. The presence of a resumptive element such as *elles* 'elles' in (1) can be used as a (sufficient, though not necessary) diagnostic for dislocation in that language.³ The DP *ta photo* 'your picture' is interpreted as a topic in (3) but not in (4).
- (3) Ta photo, elle m'a fait peur. your picture she me has made fear "Your picture scared me."

Heavy elements (i.e. elements other than weak pronouns) in the canonical subject position are obligatorily in focus in spoken French. This is the case either if the sentence is thetic (as in (4)) or if there is a narrow focus on the subject (as in (5), where the main stress falls obligatorily on *aussi* 'too').

- (4) $\int_{F} Ta$ photo m'a fait peur]. (out-of- the-blue only) your picture me has made fear "Your picture scared me."
- (5) [FLES AUTRES AUSSI] me font peur. the others too me make fear "The other ones scare me too."

Evidence to that effect can be found in Individual Level Predicates. Such predicates (ILPs) can be defined by the following three properties: (i) they cannot appear in perception reports, (ii) they do not allow an existential reading of their subject and (iii) they tend to express permanent properties (Milsark

"I love cramique."

³ The resumptive element of a dislocated subject is obligatorily overt in adult French, but not in child French during the null subject stage. The resumptive element of dislocated objects tends to be covert under a generic interpretation, as in (i).

Le cramique, j'adore. (i) the raisinbread I adore

⁴ I take the focus to correspond to the most informative part of the sentence (Cinque 1993; Reinhart 1996).

1974, Jäger 2001).⁵ ILPs cannot appear in thetic sentences (De Cat 2002). Consequently, a DP expressing the subject of an ILP is obligatorily dislocated in spoken French, unless there is a narrow focus on the subject. In (6), both A's statement and B's amendment of it contain an ILP. In B's utterance, there is a narrow (contrastive) focus on the DP expressing the subject. This DP is thus forced to appear in the canonical subject position (where the contrastive focus reading is allowed). Yet the property reading is allowed in both utterances. In A's statement, the topic is *mon voisin* 'my neighbour' and in B's it is a covert aboutness topic corresponding to something like (horrible) people in A's neighbourhood.

(6) A: Mon voisin, c'est une crapule. my neighbour_M it is a ruffian "My neighbour is a ruffian."

B: Non, TA VOISINE est une crapule. no your neighbour is a ruffian "No, your neighbour is a ruffian."

2.2.2. Dislocated elements are obligatorily topics Dislocated elements are obligatorily interpreted as topics in spoken French. As a consequence, existential indefinites are banned from the dislocated position in that language (as illustrated in (7a)). The existential reading is incompatible with a topic interpretation because topic referents must by definition be identifiable in the (discourse) context. That being said, nothing prevents nonexistential indefinites from being dislocated. In (7b), for instance, the dislocated indefinite is interpreted generically, which makes it acceptable as a topic. This sentence is interpreted as transcending particular facts: it is not about a particular child but

(i) *I saw the Muppets know English.

(ILP)

(ii) I saw the Muppets sing in French.

- ILPs do not allow a weak reading of indefinite subjects

(i) Belgians have grey eyes. (ILP. Generic reading of *Belgians*)

(ii) Belgians have invaded Washington. (Specific, existential reading of *B*.)

- ILPs often induce 'lifetime' effects

(i) Claude Deschamps was a wonderful woman. (ILP. Implies C.D. is dead)

(ii) Claude Deschamps was in her garden.

⁵ The following examples illustrate these properties:

⁻ ILPs cannot appear in perception reports

about a typical behaviour of children. In (7a), by contrast, the generic reading is not available (due to the tense specification which forces a specific reading) and the sentence is unacceptable as a result.

- (7) a. #Un enfant, il est arrivé pi il t' a posé une question a child he is arrived then he to-you has asked a question
 - b. *Tsé*, *un enfant*, *il arrive pi il te pose une question* y'know a child he arrives then he to-you asks a question "You know, a child comes and asks you something" (Auger 1994)

3. Signs of early pragmatic competence

It has long been demonstrated that even at the one-word stage, children are sensitive to the information status of discourse referents (cf. e.g. Baker & Greenfield 1988). In this secti on, I provi de evidence from spontaneous production showing that, from the onset of word combinations, French speaking children are able to encode topics adequately.

The data used for this purpose comes from two longitudinal corpora following the acquisition of French by five children from different dialectal areas: Belgium (Léa, Tom, Chloé), Canada (Max) and France (Anne). The data from Léa, Max and Anne are part of the York corpus and the data from Tom and Chloé are part of the Cat corpus. The data was collected fortnightly (York corpus) or monthly (Cat corpus) in an environment familiar to the child, over a period of 18 months in average, starting at the onset of word combination, i.e. around 1;10 (1 year 10 months) or 2 (except for Léa, who was 2;8 at the beginning of the recordings). The period of study corresponds to the null subject stage (which will be explicitly defined in section 3.3). Data from 12 adults interacting with the children were also used for the present analysis.

Transcription and checking were done by native speakers. The transcription conventions used in the examples reported in this paper are as follows: (i) Missing elements are indicated either by 0 (in cases allowing multiple possibilities) or by bracketing (where the missing element was clearly identifiable). (ii) $L\dot{a}$ with a capital is deaccented and is interpreted as a situational deictic rather than as a true locative. (iii) The hash # sign indicates a short pause. (iv) E stands for a verb ending ambiguous between

⁶ The York corpus was collected under the direction of Bernadette Plunkett, who has kindly allowed me to use it. The Cat corpus was collected for my doctoral research, which was funded by ESRC grant R00429834373. Details on transcription and coding procedures can be found in De Cat (2002) and De Cat & Plunkett (2002).

infinitive (er) and participial (é) morphology. (v) False starts with or without correction are indicated by [//] and [/] respectively. (vi) xx correspond to an unintelligible word. (vii) The comment in square brackets introduced by =! describes what the child does to enact the (omitted) object of her sentence. (viii) Interrupted words are marked with an ampersand &. (ix) Double commas indicate a tag. (x) A e in the transcription stands for an embryonic element (e.g. a proto (clitic) pronoun). (xi) Alternative transcriptions are given in square brackets, after a =? sign. (xii) Speakers are designated by a three letter code. E.g. *TAT is Anne's child-minder. (xiii) Proper names are transcribed without capital when dislocated, to distinguish them from vocatives. A first observation of the data reveals that the children's dislocated elements are always compatible with a topic interpretation, even when the utterance is not fully target-like. Even when the dislocated elements fail to be resumed by a weak pronoun (as in (8a)) or occur in verbless utterances (as in (8b)), they still correspond to the topic of the sentence/utterance.

(8) a. O est pas une fille, isabelle.
is not a girl Isabelle
"Isabelle isn't a girl."
b. O dans les briques, le bonhomme.
in the bricks the man
"The mans('s) in the bricks."
(Anne 1;10.26)

In sections 3.1 - 3.3, I point to clear indications that very young children have the competence required for the target-like encoding of topics.

3.1. Children comply by the 'ILP requirement'

Individual Level Predicates (ILPs) provide us with an effective means of testing whether children master the notion of topic: the subject of such predicates obligatorily corresponds to the topic of the sentence, except when a narrow focus reading is appropriate. Showing that children comply by this requirement would indicate that they master the notion of topic and that, like adults, they dislocate any non-weak element interpreted as topic. In order to test this, I will apply the following reasoning. Only heavy elements expressing the subject are targeted by the 'ILP requirement'. So only sentences containing such a heavy element (whether dislocated or in the canonical subject position) need concern us here, not those where the subject is expressed by a weak pronoun only or is missing (on the uncontroversial assumption that missing subjects are missing pronouns).

What needs to be shown is that (i) children show clear signs of compliance with the 'ILP requirement', i.e. they do dislocate heavy elements expressing the subject of ILPs, and (ii) children do not violate the 'ILP requirement', i.e. they do not utter sentences containing an ILP with a heavy subject unless there is a clear narrow focus on that subject.

The sentences in (9) prove the first point. In all cases, a property reading of the predicate was clearly intended and the context (not included here because of space limitation) clearly indicated that the child did not refer to a specific (set of) individual(s).

- (9) a. les sucettes, ça finit pas. (Anne 3;1.15) the lollipops it finishes not "Lollipops don't end."
 - b. parce que c'est blanc, un ours. (Max 2;9.27) because it is white a bear "Because bears are white."
 - c. le coca, ça soigne le hoquet aussi. (Léa 2;9.5) the coke it cures the hiccough too "Coke cures hiccough too."
 - d. et les vaches, # elles mangE # de l'herbe. (Tom 2;1.13) and the cows they eat some the grass "And cows eat grass." (Clearly generic context)

The corpora were then searched for heavy subjects. A total of 186 cases was found in the child data, among 5043 clauses (of which were 4701 finite and 342 non-finite). Only one case (given in (10)) involves an ILP, and it is target-like: the subject is in narrow (contrastive) focus. This sentence was uttered after the child has been asked if her mum is nice. The intended topic in (10) is clearly "the set of nice people in Anne's family". The only new information is conveyed by the subject.

The other 185 cases are clearly not ILPs, and typically contain eventive predicates, as in (11).

- (11) a. ah et tout [/] tout le monde est tombé. (Anne 2;7.1) ah and all all the people is fallen "Ah everybody's fallen over."
 - b. (la) vache # mange # (des) ca(r)ottes. (Max 2;5.29) (the) cow eats (some) carrots "The cow's eating carrots." (Clear ongoing interpretation)
 - c. Parrain, Luc s'est fait mal. (Léa 2;8.22) grandad Luc REFL-is done pain "Grandad. Luc's hurt himself."
 - d. et pourquoi#les lunettes s'étaient xx envolées? (Tom 2;6.12) and why the glasses REFL-were ? flown "And why did the glasses fly away?"

Children's spontaneous production thus clearly suggests that they know that ILPs cannot appear in thetic sentences, which requires them to be able to identify and encode topics in a target-like fashion.

3.2. Children do not dislocate existential indefinites

Another indication that children have the required competence to encode topics adequately comes from the absence from their spontaneous production of dislocated indefinites requiring an existential interpretation. Out of the 19 dislocated indefinites I have found in these five children's spontaneous production, 16 clearly required either a generic interpretation (as in (12), which Anne utters to prevent the interviewer from making a table for the puppets they have been playing with) or a D-linked one (as in (13), uttered after Chloé's mum has asked to draw a man for her). D-linked indefinites refer to a set previously established in the discourse, which makes them possible topic candidates.

- (12) a. parce que c'est pas beau, une table. (Anne 3;9.9) because it is not beautiful a table "Because tables are ugly."
 - b. on fait comment, un monsieur? (Chloé 3;0.21) one makes how a man "How does one draw a man?"

In (13a), the child is talking about a set of cars that has just been fixed by the interviewer; in (13b), the child is referring to a group of animals wanting to enter the zoo enclosure; in (13c), she is talking about the members of a family she has put in bed under a leaf.

- toutes les autos, elles sont réparées. (Max 2:4.18) (13) a. they are fixed all the cars "All the cars are fixed."
 - les copains, tous les copains, i(ls) veulent entrer. (Anne 2;7.1) the friends all the friends they want to-enter "All the friends want to come in."
 - tout le monde, i(l) dort. (Anne 2;8.20) all the people he sleeps "Everybody's sleeping."

The only dislocated indefinites not compatible with a generic or a Dlinked interpretation are given in (14), (16) and (17). However, there are good reasons to believe that the use of indefinites here is due to performance errors or the use of unanalysed chunks by the child. The context immediately preceding (14) is given in (15), where *CAT is the interviewer. Given this context, and that the child is looking at the little baby in question as he utters (14), it is clear that a specific reading is intended, not an existential one.

- (14) un petit bébé, je sais pas c'est quoi. (Max 2;7.25) a little baby I know not it is what

 - "I don't know what the little baby's called."
- (15) *CAT: tu as guéri le petit bébé? "You cured the little baby?"
 - *CAT: comment il s'appelle, le petit bébé '? "What is the little baby called?"

The sentence in (16) is clearly a performance error, as it is immediately followed by a self-correction from the child.

(16) des cochons, ils font 0 =! she swirls. (Léa 3;0.5) they do some pigs "Pigs go..." (She demonstrates what pigs do)

As for the dislocated indefinites in (17), they were prompted by adult speech. In (17a), the child does not seem to be able to identify what de l'herbe 'grass' refers to on the picture she's been looking at. The context in which (17b) is uttered is one in which the child has been asked to place face attributes (eyes, mouth, moustache etc) back onto a sheet, and is highly confused as to what the adults expect from her. In both cases, the intended interpretation is clearly not existential.

```
(17) a.
         c'est
                  où, le l'
                                              l' herbe [//] de l'
                                &her [/] le
                                                                     herbe?
                  where the the gra the the
                                                            PART the grass
                                              grass
         "Where's 'the grass'?"
                                                                 (Anne 2;6.2)
         oui.. un [/] un oeil.
                                              où?
                                                                  (Max 2:4:4)
    b.
                                (c)a va
                  an eve
                                     goes
                                              where
         ves a
                                it
         "Where does (that) eye go?"
```

I conclude that children's spontaneous production gives no indication that they ever dislocate existential indefinites. This does not appear to be due to a sampling artifact, given that dislocated indefinites do appear in their speech (though in accordance with target usage).

3.3. Target-like use of left- vs. right-dislocation

Left- and right-dislocated elements both express the topic of the sentence, but they do so in a slightly different manner (cf. e.g. Lambrecht 1981; Ashby 1988). Essentially, these differences stem from a greater (prosodic and structural) saliency of the left periphery (De Cat 2002). As a result, some nuances are available to lefthand topics but not to their right-hand counterparts. A typical example is that of contrastive topics: these require prosodic saliency (i.e. contrastive stress), which is only possible in the left-periphery. As a consequence, contrastive topics cannot be right-dislocated. 7

Wh-questions are an environment in which the left-vs. right-dislocation contrast is most visible: the majority of dislocated elements coreferential with the subject (henceforth dislocated subjects) appear in the right-periphery in such questions. Significantly, the very same proportion of right-dislocated

⁷ This can be illustrated by the following contrast:

⁽i) Le gros livre, tu peux le mettre ici. 'You can put the big book here.'

a. Les autres, ils iront en haut. the others they will-go upstairs

b. #Ils iront en haut, les autres. they will-go upstairs the others

subjects is observed in adult speech $(78\%)^8$ and in child speech: In the speech of Max, Anne and Tom between the ages of 1;9 and 2;6, dislocated subjects in whquestions appear in the right-periphery in 76% of cases in average. This is shown in Table 1.

	Left	dislocation	Right	dislocation
children	24%	(78/330)	76%	(252/330)
adults	22%	(105/479)	78%	(373/479)

Table 1: Wh-questions involving a subject dislocation

Other indications that children differentiate between left- and right-dislocation in a target-like fashion can be found in their exclusive use of left-dislocation to express contrastive topics (as illustrated in (18)) or whenever there is a need for emphasis on the topic (i.e. whenever structural/prosodic saliency is required — as for instance in the repair strategy illustrated in (19)).

```
(18) maman, e fait xx \# moi, e fais (le) drapeau.
                                                           (Anne 2.0.27)
           e makes?? me e make (the) flag
    "Mum makes the ?? and I make the flag."
(19) *ANN: i(l) est où.
                             mon cass(er)ole?
                                                                 (2:4.2)
                     where my saucepan
            "Where is my saucepan?"
    *TAT:
            qu'
                     est ce que tu cherches?
            "What are you looking for?"
    *ANN: mon\ cass(er)ole, (e)lle\ [=?(i)l]
                                               est où?
            my saucepan
                             it
                                               is where
            "Where is my saucepan?"
```

I conclude that children not only know that heavy topics must be dislocated, they are also aware from an early age of the subtle differences between left- and right-dislocation of topics.

4. Implications for the study of null subjects

It has been proposed in the acquisition literature that subject omission in finite contexts was due to a lack of ability to license topics adequately (e.g. Bromberg & Wexler 1995; Wexler 1998; Schaeffer, Gordishevsky, Hadar, &

⁸ This is the proportion observed in a random sample of 5613 utterances from the adult speakers of the York and the Cat corpora, out of which 479 are root *wh*-questions.

Hacohen 2002). On the assumption that subjects can be dropped in adult English under certain pragmatic conditions, these researchers postulate that children lack the ability to evaluate such pragmatic conditions and that as a result, they omit more subjects than the target grammar would allow. What the relevant pragmatic conditions for subject drop are is however left rather vague. as no precise account of the workings of subject drop in adult English is provided. The closest one comes to a definition is that dropped subjects have to be very strong topics (Wexler 1998:35). I suppose strength here is to be understood in terms of saliency in the context. While I do not dispute that missing subjects are (interpreted as) topics, I believe that I have presented here clear indications that throughout the null subject stage, children have the ability to identify and encode topics adequately. It is therefore highly unlikely that they would omit certain subjects erroneously because they are (salient) topics. Looking at the language production of French speaking children during the null subject stage, I have found no indication that subject omission was influenced by the level of saliency of the subject referent. The relevant figures are given in Table 2 below. The data is from Anne, Max and Tom from the York and Cat corpora, while they were at the null subject stage.

Assuming that very strong topics correspond to topics that are very salient in the context, we could deduce that such topics will not need to be identified explicitly, and can be encoded simply by using a pronoun (without coindexed, dislocated DP). One could think of such cases as involving a covert dislocated element, following e.g. Gundel (1975); Erteschik-Shir (1997). By default, the topic of the sentence corresponds to the grammatical subject (as we have seen in section 2.1), so the assumption that the pronominal subject of a sentence without overt dislocated topic is interpreted as the topic of that sentence seems reasonable. Given this, if children tend to omit only very salient topics in their finite sentences, we could expect the rate of subject omission to be higher in the absence of a dislocated topic. But this is not what we find in either of the sub-periods of the null subject stage, there is no statistically significant difference between the omission of subjects in the presence and in the absence of coreferential dislocated element. Period 1 in Table 2 corresponds to the core of the null subject stage. During Period 2, subject omission dwindles to less than 15% of obligatory contexts.

An alternative interpretation might be that the presence of a dislocated topic coreferential with the subject would facilitate the identification of the subject referent and hence increase the chances of subject omission, but this

⁹ Period 1: γ =2.230, p<0.20. Period 2: γ =1.227, p<0.30.

hypothesis is equally untenable, given the absence of significant difference discussed above. ¹⁰

Period 1	with	without	Total
	overt subject	overt subject	
with dislocated	435	164	599
subject topic	73%	27%	
without dislocated	984	436	1420
subject topic	69%	31%	

Period 2	with	without	Total
	overt subject	overt subject	
with dislocated	548	18	566
subject topic	97%	3%	
without dislocated	2027	89	2116
subject topic	96%	4%	

Table 2: Realization of the subject of finite verbs according to the presence of a coreferential dislocated topic during the null subject stage

Because of the vagueness of Wexler's (1998) proposal, the evidence discussed in this section can only be taken as an *indication* that this proposal rests on shaky grounds. However, it is significant that in a language where topics are encoded in a distinctive way at surface structure, no evidence for a 'pragmatic' delay related to topic has been found. This suggests that the default assumption should be that children *do* have the relevant discourse (or 'pragmatic') competence. It does not entail that no 'pragmatic' cause underlies the null subject phenomenon, but that null subjects are most probably not caused by a *lack* of 'pragmatic' competence.

5. Conclusion

The competence required to encode topics has been shown to be available to children from the earliest attested stages of language production. Dislocated

¹⁰ Left- and right-dislocated topics have been collapsed into one category in Table 2 because the direction of the dislocation had no effect on subject omission. In particular, subjects do not tend to be omitted more when there is a left-dislocated topic. Having explicitly identified the topic *before* uttering the sentence does therefore not increase the likelihood that the element resuming the topic would be omitted.

elements expressing the topic appear at the onset of word combination in child French, and children's use of left- and right-dislocation not only resembles target usage but also complies with its requirements: the present corpora suggest that children only dislocate indefinites when a topic interpretation is allowed, and that they abide by the requirement that the heavy subject of ILPs be dislocated except when it is in focus, which requires them to be able to identify and encode topics in a target-like fashion. These signs of very early competence in handling information structure phenomena are compatible with the findings of Baker & Greenfield (1988), who have shown that in their first year of life, children are already able to distinguish new from established information. By contrast, no support has been found for the hypothesis attributing subject omission in fini te sentences to a lack of 'pragmatic' competence relating to topic encoding. Such early evidence for information structure competence indicates that at least some of what has been considered 'pragmatic' competence in the literature (e.g. Wexler 1998; Schaeffer 2000) is available from the earliest stages of language acquisition.

References

- Ashby, William. 1988. "The Syntax, Pragmatics, and Sociolinguistics of Left-and Right-Dislocations in French". *Lingua* 75(23).203–229.
- Auger, Julie. 1994. Pronominal Clitics in Québec Colloquial French: A Morphological Analysis. PhD thesis. University of Pennsylvania.
- Baker, Nancy & Greenfield, Patricia. 1988. "The Development of New and Old Information in Young Children's Early Language". *Language Sciences* 10(1).3–34.
- Bromberg, Hilary & Wexler, Kenneth. 1995. "Null Subjects in Child Wh-Questions". *MIT Working Papers in Linguistics* 26. 221–247.
- Chien, Yu-Chin & Wexler, Kenneth. 1991. "Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics". *Language Acquisition* 1.225–295.
- Cinque, Guglielmo. 1993. "A Null Theory of Phrase and Compound Stress". Linguistic Inquiry 24.239–298.
- De Cat, Cécile. 2002. French Dislocation. PhD thesis. University of York.
- De Cat, Cécile & Plunkett, Bernadette. 2002. "Qu' est ce qu' i(l) dit, celui+là? Notes Méthodologiques sur la Transcription d'un Corpus Francophone". *Romance Corpus Linguistics: Corpora and Spoken Language*, ed. by C.D. Pusch, & W. Raible, CD-ROM edition. Tübingen: Narr.
- Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.

- Givón, Talmy. 1976. Topic, Pronoun and Grammatical Agreement. *Subject and Topic*. ed. by C. Li, 149–211. New York: Academic Press.
- Gundel, Jeannette. 1975. "Left Dislocation and the Role of Topic-Comment Structures in Linguistic Theory". *OSU WPL 18*.72–131.
- Jäger, Gerhard 2001. "Topic-Comment Structure and the Contrast Between Stage Level and Individual Level Predicates". *Journal of Semantics* 18.83-126.
- Lambrecht, Knud 1981. *Topic, antitopic and verb agreement in non-standard French.* vol. II:6 of *Pragmatics & Beyond.* Amsterdam: John Benjamins.
- Lambrecht, Knud 1994. *Information Structure and Sentence Form. Topic, focus, and the mental representation of discourse referents*. Cambridge: CUP.
- Li, Charles & Thompson, Sandra. 1976. "Subject and topic: A new typology of language". *Subject and Topic* ed by C. Li, NewYork: Academic Press. 457–489.
- Milsark, Gary. 1974. Existential Sentences in English. PhD thesis. MIT.
- Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics". *Philosophica 27(Special issue on Pragmatic Theory)*.53–94.
- Reinhart, Tanya. 1996. "Interface Economy: Focus and Markedness". *The Role of Economy Principles in Linguistic Theory* ed. by Wilder, Gaertner, & Bierwisch, Berlin: Akademic Verlag.
- Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". *Elements of grammar. A handbook of generative syntax* ed. by L. Haegeman, Dordrecht: Kluwer. 281–337.
- Schaeffer, Jeannette. 2000. *The Acquisition of Direct Object Scrambling and Clitic Placement: Syntax and pragmatics*. Amsterdam: John Benjamins.
- Schaeffer, Jeannette, Gordishevsky, Galina, Hadar, Galit, & Hacohen, Aviya. 2002. "Subjects in English-Speaking Children with Specific Language Impairment". *Proceedings of the GALA 2001 Conference* ed. by Joao Costa, & M. J. Freitas, 223–231, Lisboa: Associação Portuguesa de Linguística.
- Wexler, Kenneth. 1998. "Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation of the Optional Infinitive Stage". *Lingua* 106.23–79.

ON THE IMPACT OF FRENCH SUBJECT CLITICS ON THE INFORMATION STRUCTURE OF THE SENTENCE*

CECILE DE CAT *University of York*

1. In a nutshell

In spite of the vast amount of literature dedicated to it, the status of French subject clitics is still an unresolved issue within morpho-syntactic theory. Two main analyses have been proposed and defended over the past three decades: one advocating that French subject clitics are syntactic arguments bearing a θ -role (henceforth 'the syntactic analysis', cf. e.g. Kayne 1975; Rizzi 1986; Laenzlinger 1998; Belletti 1999), the other viewing such clitics as mere inflectional morphemes on the verb (henceforth 'the morphological analysis', cf. e.g. Jaeggli 1982; Roberge 1986; Hulk 1986; Auger 1994; Miller & Monachesi in press).

This paper will not propose arguments for or against either of these analyses. Its aim is to draw the attention to some information structure facts that have been largely ignored in the literature and that any analysis of French subject clitics should account for in a principled way. It demonstrates that in spoken French, a topic interpretation of an XP expressing the subject of the sentence is only possible when this XP cooccurs with a coindexed subject clitic. In the absence of such a clitic, the XP is shown to be obligatorily in focus and is therefore argued to appear in the canonical subject position.

2. Introduction and background

Two factors come into play in the analysis of the string ' XP_i - subject clitic_i', (as exemplified in (2)):

^{*} This research was supported by the Economic and Social Research Council (# 041R00433), which is gratefully acknowledged. For their useful comments, I wish to thank David Adger, Aafke Hulk, Marie Labelle and Bernadette Plunkett as well as two anonymous reviewers.

- (1) a. the locus where the clitic is generated
 - b. the syntactic position occupied by the XP
- (2) Les clitiques_i, ils_i comptent pour du beurre? the clitics they count for some butter "Don't clitics count?"

The subject clitic can be analysed as a syntactic entity (i.e. an element available for syntactic operations) to which a θ -role has been assigned and which has transited via the canonical [spec,TP] subject position, from where it has cliticised phonologically on the verb (e.g. Belletti 1999); alternatively, the subject clitic can be considered as an agreement morpheme generated directly on the finite verb (e.g. Auger 1994).

An XP expressing the subject in spoken French can either appear in the canonical subject position or in a peripheral position allowing a topic interpretation. This is illustrated in (3a) and (3b) respectively:¹

The latter possibility must be acknowledged under either analysis of subject clitics at least in cases where an element intervenes between the XP and the clitic, thereby indicating that the XP cannot be in [spec,TP] in such cases. Examples include cases like (4a), where the XP expressing the subject precedes a fronted *wh*-word and cases like (4b), where the XP expressing the subject of the embedded clause appears in the left-periphery of the matrix clause.

- (4) a. Et la clé_i, où elle_i est? and the key where she is "And where's the key?"
 - b. La clé_i, je pense qu' elle_i est restée dehors. the key I think that she is stayed outside "I think the key's stayed outside."

Given the remarks above, the combination of the two factors in (1) yield three viable possibilities for the analysis of the string ' XP_i - subject clitic_i'

¹ I am abstracting away from the question of whether CP consists of several layers of projections. As explained in section 3, the XP in (3b) will be taken to be adjoined to either TP (possibly as in (i)) or CP (as in (4)).

⁽i) Tu crois que les autres i, ils i aimeraient ça? you think that the others they would-like that

(excluding the possibility of coexistence of both elements in the canonical subject position, which is ruled out by the Subcategorisation Principle, cf. Chomsky 1965). These possibilities are spelled out in (5):

```
(5) a. [CP XP_i ... [TP clitic_i [T^* ... ]]]
b. [CP ... [TP XP_i [T^* clitic_i + T ... ]]]
c. [CP XP_i ... [TP e [T^* clitic_i + T ... ]]]
```

Option (5a) corresponds in essence to the analysis of e.g. Kayne (1975); Rizzi (1986); Laenzlinger (1998); Belletti (1999), prior to phonological cliticisation of the clitic. It is the only possibility allowed under the assumption that subject clitics are syntactic entities. Option (5b) corresponds in essence to the analysis of e.g. Jaeggli (1982); Roberge (1986); Hulk (1986); Auger (1994). Under the morphological analysis of subject clitics, the possibility of option (5c) (and not just (5b)) has to be acknowledged, on the grounds that a dislocation of the XP is possible (given the existence of structures such as those in (4)).

I will argue on interpretive grounds that, contrary to what is standardly assumed by the proponents of the morphological analysis of subject clitics, option (5b) never arises in spoken French (or at least not in its most widely spoken varieties across Belgium, France and Quebec).²

3. Distributional restrictions

It is well established that peripheral XPs resumed by an element within the sentence are interpreted as the topic of that sentence (see e.g. Gundel 1975; Larsson 1979; Reinhart 1981; Lambrecht 1981, 1994). The most widely accepted definition of topic (which I will adopt here) is that proposed by Reinhart (1981) as 'what the sentence is about' (hence the term 'aboutness topic'). In De Cat (2002), I argue that peripheral topics are dislocated by basegenerated adjunction to a sentential maximal projection (TP or CP). I will therefore refer to the peripheral XPs that occupy us here as 'dislocated phrases'. The observation above can be rephrased as (6), as has been argued for spoken French in De Cat (2002):

(6) Only possible topics can be dislocated in spoken French.

² For a discussion of how to define spoken French, see De Cat (2002), where it is also argued that 'Advanced French' (whose existence was postulated by Zribi-Hertz 1994) does not correspond to any attestable variety of spoken French.

³ Note however that the argumentation in this paper is equally compatible with a feature-driven analysis of peripheral topics, e.g. \hat{a} la Rizzi (1997).

A number of researchers (such as Roberge 1990; Auger 1994; Zribi-Hertz 1994) have taken examples like (7) to indicate that, in dialects allowing such a sentence, subject clitics cannot be syntactic entities bearing a θ -role. Instead, they argue, they are a kind of agreement morpheme on the verb. This depends on the assumption that indefinites like *un enfant* 'a child' in (7) cannot receive a topic interpretation:

(7) Un enfant_i il_i arrive pi il te pose une question. (Auger1994) a child he arrives then he to-you asks a question "A child arrives and asks you a question."

The reasoning behind such a statement is as follows: (i) If subjects clitics are syntactic entities, any XP coindexed with one such a clitic must appear outside of the canonical subject position (by virtue of the Subcategorisation Principle). (ii) Dislocated XPs are obligatorily interpreted as the topic of the sentence. Therefore, if the subject clitic in (7) is a syntactic entity, the coindexed *un enfant* 'a child' must be dislocated. (iii) But indefinites cannot be topics, so a dislocation analysis of *un enfant* 'a child' is impossible. (iv) Therefore the sentence in (7) is representative of a (dialectal) variety of French in which subject clitics are not arguments. (v) The only alternative is that such elements are morphemes in T in that (dialectal) variety.

This reasoning is based on that in Rizzi (1986). However, what is often overlooked by Rizzi's followers is that indefinites *per se* are not banned as topics: it is only under their existential reading that they are incompatible with a topic interpretation. Under a generic reading, indefinites can be topics (Côté 2001). And the sentence in (7) is precisely one that receives a generic interpretation: this sentence is not about a particular child, but about a behaviour that is typical of children in general. If a specific reading is forced (by using a past tense, as illustrated in (8)), this sentence is no longer acceptable for speakers of the main varieties of spoken French (including speakers of Canadian French, to one of whom the sentence in (7) is attributed):

(8) *Un enfant, il est arrivé pi il t' a posé une question. a child he is arrived then he to-you has asked a question

In section 4, I demonstrate that in spoken French, a heavy (i.e. non-weak) element expressing the subject is interpreted as a topic *only if* it is resumed by a subject clitic. The presence of a subject clitic in (7) is therefore not only possible but obligatory (which turns the argument of Auger and others on its head).

4. The presence of a subject clitic forces the topic interpretation of a coindexed XP

Three arguments are presented below to demonstrate that in spoken French, the presence of a subject clitic has a direct impact on the information structure of that sentence. In particular, the XP can only receive a focus reading in the absence of a coindexed clitic (first and second arguments) and the XP can only be interpreted as the topic in the presence of a coindexed clitic (third argument).

4.1 First argument: availability of a focus reading of the XP

The focus is traditionally understood as the most informative part of a sentence (Rochemont 1986). It can be restricted to the subject, as in (9) (as a marked option — cf. Cinque 1993 and Reinhart 1996). The focus here is in capitals, indicating stress prominence.

(9) Q: Who's eaten my porridge?

A: GOLDILOCKS has.

In (9), only the subject conveys new information (which is also clear from the fact that the VP has been elided).

Dislocated DPs cannot be focused. They cannot convey the answer to a *wh*-question. This is illustrated for dislocated objects in (10):

(10) Q: Qu'est-ce qu'il a senti?

"What did he smell?"

A: [LA CHAIR FRAICHE]_i, il *(l'_i) a senti(e). the flesh fresh he it has smelled

If an XP coindexed with a(n adjacent) subject clitic allows for a focus interpretation of that XP, it implies that the XP in question is not dislocated, and hence that it occupies the canonical subject position (yielding a structure like (5b) rather than (5a)). To test whether such an option is allowed in spoken French, a judgement elicitation task was carried out on 14 native speakers from Belgium, Canada and France. The informants were presented with 18 contexts (including 9 distractors), each with three possible follow-ups (pre-recorded on CD, with no transcription provided). The prosody of the sequences 'XP_i - subject clitic_i' was intended to be as close as humanly possible to that of the corresponding construction without clitic (so as to avoid prompting a dislocation analysis of that XP). Each set of possible follow-ups contained one sentence with an XP subject and no coindexed clitic, one sentence with an XP

subject coindexed with an adjacent subject clitic, and one clearly unacceptable distractor (either completely inappropriate in the context in question, or clearly ungrammatical in any variety of spoken French). The contexts all forced a focus interpretation of the subject. In the illustration below, C stands for context and F for follow-up. The distractor has been omitted here.

- (11) C: Qui a fini son travail? who has finished his work "Who's finished their work?"
 - F: (i) CEUX DU GROUPE A ils ont fini leur travail. those of-the group A they have finished their work "Those of the group A have finished their work."
 - (ii) CEUX DU GROUPE A ont fini leur travail. ← those of-the group A have finished their work "Those of the group A have finished their work."

The option where the XP expressing the subject is resumed by a clitic was accepted only 4.7% of the time (corresponding to 6/126 answers — distractors excluded), randomly across speakers and across dialects. Each speaker accepted at most one instance of 'XP_i - subject clitic_i' over the whole test (i.e. out of the 9 test conditions). Most speakers rejected all such configurations in the context provided, which forced a focus interpretation of the XP.⁴ The 6 answers above can thus be treated as noise in the data.

These results are consistent with a dislocation analysis of the XP coindexed with an adjacent subject clitic: dislocated XPs are topics and topics can by definition not be focused (see e.g. Erteschik-Shir 1997).

4.2. Second argument: variable binding

Zubizarreta (1998:11) argues that in several languages (including French, English and Spanish), a QP object *each/every N* may bind a variable contained within the subject if and only if the subject is focused. In spoken French, the binding of a quantifier in the subject position (as in (12)) is only possible in the absence of a resumptive clitic:

(12) *Sa_i mère* (**elle*) *accompagnera chaque enfant_i*. his mother she will-accompany each child

⁴ The informants were allowed to choose more than one option, as long as they indicated which option they preferred. Yet in almost all cases they only allowed one option.

- (13) Il faut encore décider qui rentrera chaque cheval au box.
 - "We still need to decide who will take each horse to its box."
 - a. *Son jockey, il ramènera chaque cheval.* its jockey he will-take-back each horse
 - b. *Son jockey ramènera chaque cheval.* ← its jockey will-take-back each horse

Not all of my 14 informants allowed the variable in the subject position to be bound by the distributive object QP (hence rejecting a wide-scope interpretation of the object). Those who did allow such a binding almost categorically rejected sentences where the DP containing the variable was resumed by a(n adjacent) subject clitic. Out of 28 expected responses (i.e. testing two such sentences), 9 were blank (indicating the impossibility of a wide-scope reading of the object) and only 1/19 corresponded to the string 'DP_i - subject clitic_i'. That speaker abstained from providing a judgement for the other test sentence, which suggests that she only allowed marginally for a distributive object QP to bind a variable in subject position.

Again, these results are consistent with a dislocation analysis of XPs resumed by a subject clitic: such XPs cannot be focused because they are topics.

4.3. Third argument: availability of a topic interpretation of the XP

Not all sentences can take an aboutness topic (as defined in section 3). Whether this is possible depends on the information structure of the sentence. One of the key factors to that respect is the span of the focus, i.e. how much of the sentence is new information. In certain contexts, all the information conveyed by the sentence is new. Such sentences are said to be *thetic* as opposed to categorical. Thetic sentences describe a state of affairs and are typically uttered to answer a question like *What happened?*. Contrary to categorical sentences, they do not predicate something about a referent whose existence is presupposed: their subject is not an aboutness topic. Imagine a situation in which person A sees person B in tears, prompting the following exchange:

- (14) A: Qu'est-ce qui s'est passé? "What happened?"
 - B: Les voisins_i (#, ils_i) ont mangé mon lapin. the neighbours (they) have eaten my rabbit "The neighbours have eaten my rabbit."

40 CECILE DE CAT

In that context, the referent of *les voisins* "the neighbours" is not established, so that referent is not available as a sentence topic. B's response has to describe a state of affairs; it consists entirely of new information and does not permit the presence of a subject clitic (as indicated by the # in the glosses).

What is interesting for our present purpose is that certain predicates can never appear in thetic sentences. Such predicates belong to a (relatively uniform) class defined as *Individual Level Predicates* (Milsark 1974). Individual Level Predicates (henceforth ILPs) generally express (permanent) properties of individuals or types:⁵

(15) A: Ta soeur, elle est musicienne. Your sister she is musician "Your sister's a musician."

B: $MON\ FRERE\ AUSSI\ est\ musicien.$ my brother too $_F$ is musician "My brother's a musician too."

If a topic interpretation is only allowed in the presence of a resumptive clitic in spoken French, one can therefore expect that ILPs will always take a subject clitic, except when there is a narrow focus on the subject. This prediction was confirmed by the analysis of a corpus of spontaneous language production from 12 adult speakers of French (from various regions of Belgium, Canada and France): out of a random sample of 4030 clauses from the York and Cat corpora, ILPs did not appear without a subject clitic, except in the rare instances requiring a narrow focus reading on the subject. Examples of both types are given in (16):

(16) a. La cuisine, c'est le lieu où Maman fait à manger. the kitchen it is the place where Mum makes to eat "The kitchen is the place where Mum cooks."

⁵ Things are more complex than I am suggesting in the text. For an in-depth discussion of the properties of ILPs, see Jäger (2001).

⁶ The York corpus was collected under the direction of Bernadette Plunkett (ESRC grant #R000 22 1972), who has kindly allowed me to use it. The Cat corpus was collected for my doctoral research, which was funded by ESRC grant R00429834373. Details on transcription and coding procedures as well as recording conditions and speakers can be found in De Cat (2002).

b. Luc aussi a les yeux de son père?
 Luc too has the eyes of his dad "Luc too has the eyes of his dad ?"

The quasi-obligatory presence of a subject clitic with ILPs is a direct consequence of fact that the subject of ILPs is interpreted as a topic (except when it is in narrow focus).

4.4 On the pervasiveness of subject clitics in spoken French

It is sometimes argued that the string ' XP_i - subject clitic $_i$ ' cannot possibly require a topic interpretation in all cases, given that it occurs in such a high proportion of sentences in corpora of spontaneous production (e.g. Auger 1994:116). However, one has to bear in mind that crosslinguistically, the topic strongly tends to coincide with the grammatical subject of the sentence (see e.g. Gundel 1975; Givón 1976; Li & Thompson 1976; Reinhart 1981; Lambrecht 1994). Such a high proportion of subject topics is therefore entirely expected.

What has been overlooked in the literature is that French subject clitics, while very frequent, are not omnipresent. A claim such as (17) (or at least its second part) is unwarranted:

(17) Subject markers are true agreement markers and are thus expected to show up on every finite verb (Auger 1994:93)

Under (17), the absence of subject clitic is predicted to be random, or speaker-dependent (as suggested by Auger 1994:13).

Yet, as this section has demonstrated, subject clitics are obligatorily absent in spoken French when the XP expressing the subject is in focus (i.e. when it is in narrow focus or when the sentence is thetic).

Data reported by Auger herself actually corroborate what has been argued in this paper. The Montreal speakers she studies only 'double' their subject with a clitic 70% of the time, which means that the subject clitic is absent in 30% of the cases. What would have to be shown is that in those cases, the subject is in focus. I regard as highly significant the fact that the only example quoted in Auger (1994) in which *personne* "nobody" is the subject does not contain a subject clitic:

CECILE DE CAT

(18) Personne la lisent la revue. (Auger 1994:46) nobody it read the magazine "Nobody reads the magazine."

This is exactly what should be expected under the present analysis, given that *personne* "nobody" cannot be a topic, as it does not correspond to an identifiable discourse referent.⁷

Interestingly, one of the main proponents of the morphological analysis of clitics (Roberge 1990:54) argues that quantifiers and indefinites cannot be doubled by a subject clitic in the language varieties he studied (which he claims are representative of Quebec French). This corroborates the present analysis.

5. Information structure and syntactic structure

5.1 The information structure import of subject clitics

The evidence presented so far indicates that in spoken French, an XP coindexed with an adjacent subject clitic is obligatorily interpreted as a topic. For a focus interpretation of the XP to obtain, the subject clitic must be absent. Under a syntactic analysis of subject clitics, the topic interpretation of the XP follows directly from the syntactic structure of the sentence. That XP is obligatorily dislocated whenever there is a subject clitic because it cannot occupy the canonical subject position (since it is occupied by the clitic). The resulting dislocated structure (given in (5a)) is syntactically transparent to what Erteschik-Shir (1997) calls f(ocus)-structure, the grammatical level where the scope of topic and focus is defined and which mediates between syntax and PF / LF. At f-structure, the topic has to take scope over the rest of the sentence. The information structure contrast between sentences with and without subject clitic is thus accounted for straightforwardly.

Under a morphological analysis of subject clitics, the topic interpretation of the XP expressing the subject is not forced to follow from the presence of a subject clitic, unless such clitics are endowed with special properties. One can for instance imagine a mechanism that would force the presence of a pro in the canonical subject position whenever there is a subject clitic. Subject clitics could (perhaps in addition to that) be argued to be topic markers (e.g. by endowing them with a topic feature). This would account for the fact that such clitics are only realised when the subject is interpreted as topic. An interesting consequence would be the blurring of the distinction between topics and pronoun-like elements (at least those associated with the subject): if subject

⁷ A possible exception to this would be if *personne* "nobody" refers to "nobody out of a set defined in the discourse context". The contexts are generally not given in Auger (1994).

clitics bear a topic feature, any sentence with such a clitic would force a topic interpretation of the subject (i.e. in (19a) as well as (19b)). This idea is compatible with the claim that topics can be covert (cf. e.g. Gundel 1975).

- (19) a. Tim_i, il_i a retrouvé ses framboisiers.

 Tim he has retreived his raspberrycanes
 "Tim has got his raspberrycanes back."
 - b. *Il a retrouvé ses framboisiers* . he has retreived his raspberrycanes

Note that under such an analysis, French subject clitics would not be mere agreement morphemes but would have a nontrivial information structure status. This would account for the obligatory absence of subject clitc in sentences like (20) — which remains unexplained under current versions of the morphological analysis of French subject clitics:

- (20) a. $Qui_i(*il_i)$ veut du gouda? who he wants some goudacheese
 - b. $Quels \ soldats_i \ (*ils_i) \ sont \ partis ?$ which soldiers they are left (i.e. have left)

5.2 Prosodic evidence for the dislocation of the coindexed XP

Irrespective of whether the information structure import of the subject clitic is due to an inherent feature or to its argumental status, there are good reasons to consider that the XP coindexed with such a clitic is dislocated in all cases.

First, as pointed out in the introduction, that XP is clearly dislocated at least when an element intervenes between it and the clitic (as shown by examples (4a) and (4b)).

Second, the prosodic characteristics of that XP indicate that it is dislocated even in the absence of intervening material. This has been argued for Quebec French by Deshaies, Guilbault, & Paradis (1993) and Guilbault (1993) and by myself for varieties of Canadian and European French (De Cat 2002). Contrary to what is often assumed, the decisive criterion in identifying left-dislocation prosody is not the presence of a pause between the XP in question and the rest of the sentence, but a combination of factors of which the most important are the presence of stress (i.e. increased intensity) on the last syllable of the

dislocated element, and pitch (i.e. melodic) prominence on that syllable, as compared with the pitch of what follows the dislocated element.⁸

6. Conclusion

I have demonstrated that in the most widely spoken varieties of French, subject clitics can only appear when the subject is interpreted as the topic of the sentence and are banned whenever the subject is in focus. Prosodic evidence also suggests that an XP coindexed with a subject clitic is always dislocated in spoken French. This entails that the following structure never arises in spoken French: $\begin{bmatrix} \operatorname{CP} XP_i \dots \end{bmatrix}_{TP} e \begin{bmatrix} \operatorname{T} : \operatorname{clitic}_i + T \dots \end{bmatrix}_{T}$.

Any analysis of French subject clitics, whether or not it endows them with argument status, has to provide a principled explanation for these facts.

References

- Auger, Julie. 1994. Pronominal Clitics in Québec Colloquial French: A Morphological Analysis. University of Pennsylvania: PhD thesis.
- Belletti, Adriana. 1999. "Italian/Romance Clitics: Structure and Derivation". *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 543-579. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 1965. Aspects of the Theory of Syntax. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo. 1993. "A null Theory of Phrase and Compound Stress". *Linguistic Inquiry* 24.239-298.
- Côté, Marie-Hélène. 2001. "On the Status of the Subject Clitics in child French". *Proceedings of the 8th Conference of the IASCL*, ed. by Margareta Almgren, Andoni Barreña, María-José Ezeizabarrena, Itziar Idiazabal & Brian Mac Whinney, 1314-1330. Somerville, Mass.: Cascadilla Press.
- De Cat, Cécile. 2002. French Dislocation. PhD thesis. University of York.
- Deshaies, Denise, Christian Guilbault & Claude Paradis. 1993. "Prosodie et dislocation à gauche par anaphore en français québécois spontané". *Actes du XVe Congrès international des linguistes*, ed. by André Crochetière, Jean-Claude Boulanger, & Conrad Ouellon, 31-34.
- Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.

⁸ Space restrictions prevent me from going into much detail. For a definition of Intonation Group and the importance of this notion for the prosodic analysis of French left-dislocation, see Mertens, Goldmann, Wehrli, & Gaudinat (2001).

- Givón, Talmy. 1976. "Topic, Pronoun and Grammatical Agreement". *Subject and Topic*, ed. by Charles Li, 149-211. New-York: Academic Press.
- Guilbault, Christian. 1993. "Dislocation à gauche et prosodie en québécois spontané". *Dialangue* 4.34-38.
- Gundel, Jeannette. 1975. "Left Dislocation and the role of Topic-Comment Structures in Linguistic Theory". *OSU WPL* 18.72-131.
- Hulk, Aafke. 1986. "Subject Clitics and the Pro-drop Parameter". *Going Romance: Formal Parameters of Generative Grammar II*, ed. By Peter Coopmans, Ivonne Bordelois & Bill Dotson-Smith, 107-120. Dordrecht: Foris.
- Jaeggli, Oswaldo. 1982. Topics in Romance Syntax. Dordrecht: Foris.
- Jäger, Gerhard. 2001. "Topic-Comment Structure and the contrast between Stage Level and Individual Level Predicates". *Journal of Semantics* 18.83-126.
- Kayne, Richard. 1975. French Syntax. Cambridge, Mass.: MIT Press.
- Laenzlinger, Christophe. (1998). *Comparative Studies in Word Order Variation*. Amsterdam & Philadelphia: John Benjamins.
- Lambrecht, Knud. 1981. *Topic, Antitopic and Verb Agreement in non-standard French.* vol. II:6 of *Pragmatics & Beyond.* Amsterdam & Philadelphia: John Benjamins.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form. Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.
- Larsson, Eva. 1979. *La dislocation en français: étude de syntaxe générative.* Lund: CZK Gleerup.
- Li, Charles & Sandra Thompson. 1976. "Subject and topic: A new Typology of Language". *Subject and Topic*, ed. by Charles Li, 457-489. New-York: Academic Press.
- Mertens, Piet, Jean-Philippe Goldmann, Eric Wehrli & Arnaud Gaudinat. 2001. "La synthèse de l'intonation à partir de structures syntaxiques riches". *TAL* 42(1).142-195.
- Miller, Philippe & Paola Monachesi, (in press). "Les pronoms clitiques dans les langues romanes". *Les langues romanes: Problèmes de la phrase simple* ed. by Danièle Godard. Paris: Editions du CNRS.
- Milsark, Gary. 1974. Existential Sentences in English. PhD thesis. MIT.
- Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics". *Philosophica* 27(Special issue on Pragmatic Theory).53-94.
- Reinhart, Tanya. 1996. "Interface economy: Focus and Markedness". *The Role of Economy Principles in Linguistic Theory*, ed. by Chris Wilder, Hans MartinGaertner & Manfred Bierwisch, 146-169. Berlin: Akademic Verlag.

- Rizzi, Luigi. 1986. "On the Status of Subject Clitics in Romance". *Studies in Romance Linguistics* ed. by Oswaldo Jaeggli & C. Silva-Corvalan. vol. 24 of Publications in Language Sciences. 391–419. Dordrecht: Foris.
- Rizzi, Luigi. 1997. "The fine Structure of the Left Periphery". *Elements of grammar. A handbook of Generative Syntax* ed. by Haegeman, L., 281-337. Dordrecht: Kluwer.
- Roberge, Yves. 1986. "Subject Doubling, Free Inversion, and Null Argument Languages". *Canadian Journal of Linguistics* 31.55-79.
- Roberge, Yves. 1990. *The syntactic recoverability of Null Arguments*. Montréal: McGill-Queen's University Press.
- Rochemont, Michael. 1986. Focus in Generative Grammar. Amsterdam: John Benjamins.
- Zribi-Hertz, Anne. 1994. "The Syntax of Nominative Clitics in Standard and Advanced French". *Paths Towards Universal Grammar* ed. by Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Rafaella Zanuttini. Georgetown Studies in Romance Linguistics, 453-472. Washington, DC: Georgetown University Press.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus and Word-Order*. vol. 23 of LI Monographs. Cambridge: MIT Press.

A RESTRICTED VIEW OF HEAD MOVEMENT*

HELES CONTRERAS University of Washington

1. Introduction

In recent versions of the minimalist program (Chomsky 2000, 2001), the conditions on syntactic movement are much more stringent than in previous models. In addition to agreement between a *probe* and a *goal*, the former must provide an adjunction site for the latter, and they must both be *active*, i.e. have uninterpretable features. Agreement alone is not sufficient to trigger movement, since uninterpretable features may be deleted in situ.

Consider the well-known case of existential sentences like (1):

(1) There are many books on the table.

Previous approaches to this type of construction required covert adjunction of *many books* or its relevant features to the functional category Tense. Under the restricted view of movement of the current model, such movement is unnecessary. Tense and *many books* agree in their person and number features, which are valued and interpretable for the DP, but unvalued and uninterpretable for Tense. The DP has an uninterpretable Case feature. Since *many books* is in the domain of Tense, they can delete each other's uninterpretable features without any movement. The feature which sometimes requires movement is an independent uninterpretable property of Tense usually called an EPP feature. In our example, this feature is deleted by merging in the expletive *there*.

Given this restricted view of movement, it is possible to speculate, as Chomsky (2001) does, that the so-called *displacement property* of language may not be an imperfection at all but it may instead be required by the conceptual interface. In other words, displacement serves the function of constructing structures that are 'legible' by the thought systems. While this is

^{*} Thanks to two anonymous reviewers for their very useful comments. Remaining errors are my responsibility.

plausible in the case of phrasal movement, it does not seem to apply to head movement at all. As Chomsky (2001) points out, the semantic effects of head movement are slight or nonexistent. Since head movement is anomalous in several other respects, Chomsky has suggested that it should be reanalyzed as a phonological operation.

Logically, it is possible to retain the restricted view of movement of the current minimalist model without treating all cases of head movement as phonological operations. In this paper I want to present empirical evidence in support of the view that the standard cases of head movement fall into three classes: a) phonological operations; b) internally driven head movement; c) no movement. In previous work (Contreras 2001, 2003) I have argued that T-to-C involves no movement, and that some cases of English V-to-T are phonological (Halle and Marantz 1993, Lasnik 1999), while the Spanish cases are either type b or type c. I will review this work here and in addition I will consider 'short' and 'long' N-movement as suggested by Cinque (1994) and Longobardi (1994, 1996) respectively, and show how these operations fit within the proposed framework. I will leave the case of incorporation (Baker 1988) and other potential instances of head movement, such as clitic climbing, for future research.

An important consequence of this new perspective on head movement is that parametric variation must be reexamined. I conclude the paper by suggesting how this can be accomplished.

2. Standard treatment of the data

The standard instances of head movement include the following subcases: inversion (T-to-C), verb raising (V-to-T), and 'short' and 'long' N-movement. These are illustrated in (2)-(5) respectively:

- (2) a. Will Chris write the letter?
 - b. What will Chris write?¹

¹ T-to-C appears to extend to cases other than interrogatives, as shown in (i).

⁽i) a. Should Chris call, please take a message.

b. Seldom have I heard such nonsense.

However, conditional and negative inversion is much more idiosyncratic than interrogative inversion, as shown by the ungrammaticality of (ii).

⁽ii) a. *Could Mary speak French, she would have shown up. (from Pesetsky 1989)

b. *Never can I pin it down.

These facts make it difficult to treat all these cases as resulting from the same operation.

- (3) Pedro examinó cuidadosamente los datos.

 Peter examined carefully the data

 "Peter examined the data carefully."
- (4) el examen cuidadoso de los datos the examination careful of the data "the careful examination of the data"
- (5) a. L'antica Roma fu la città più importante del Mediterraneo. the ancient Rome was the most important city of the Mediterranean
 - b. *Antica Roma fu la città più importante del Mediterraneo. ancient Rome was the most important city of the Mediterranean
 - c. Roma antica fu la città più importante del Mediterraneo. Rome ancient was the most important city of the Mediterranean
 - d. La Roma antica fu la città più importante del Mediterraneo. the Rome ancient was the most important city of the Mediterranean (examples from Longobardi 1994)

All four of these cases have been analyzed as adjunction of a head to a higher functional head: T-to-C in (2), V-to-T in (3), N to some intermediate head, possibly Number, in (4), and N-to-D in (5c). The fact that languages like English seem to lack the overt movement shown in (3) through (5) has been attributed to differences in the strength of the attracting functional head.

3. Theoretical problems

Under a view of movement where certain functional heads attract other elements in their domain, the analysis just sketched seems reasonable. It is less so, however, under the view of movement as a last resort operation triggered by a specific EPP feature. To illustrate one theoretical problem that the head-movement analysis faces within a restricted view of movement, consider examples like (2b):

(2) b. What will Chris write?

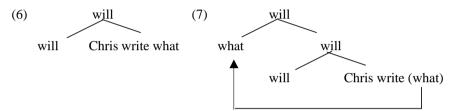
Movement of *what* from its underlying direct object position is triggered by an EPP feature associated with the interrogative complementizer with which *what* shares a feature, say [+Q]. What about movement of *will* from T to C? It would also have to be triggered by an EPP feature on C, different from the one responsible for movement of *what*, and formulated in such a way as to be deletable only by a head, not by a higher projection.² In addition to the obvious

.

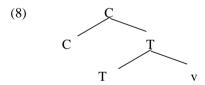
² For an interesting analysis along these lines, see Pesetsky & Torrego 2001.

problem of feature proliferation, this type of analysis is incompatible with the bare phrase structure approach adopted in the minimalist program which dispenses with bar levels, and thus makes it impossible to distinguish heads from their projections in the usual way.

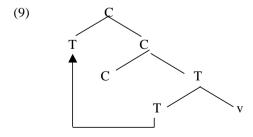
Another theoretical problem common to all cases of head movement is that they are countercyclical, as pointed out in Chomsky 2001. Unlike cases of phrasal movement, which involve merger at the top of the structure, head movement inserts an element *inside* a structure previously formed. Thus, when *what* raises in (2b), it merges with structure (6), yielding (7):

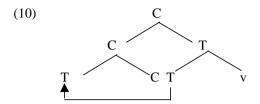


In other words, this type of movement obeys Chomsky's (1995:190) Extension Condition, which disallows operations that are internal to a structure already formed. Contrast this with the movement of T to C. In order for this operation to apply, there must be a structure like (8):



If raising of T were cyclical, the resulting structure would be (9). But the correct structure is (10), in violation of the Extension Condition:





The countercyclical property of head raising is, of course, not a decisive argument against the existence of this operation. It could be that the theory is too restrictive, and countercyclical operations do exist. The empirical evidence, however, suggests that this is unlikely.

In addition to these general theoretical problems, there are problems that are specific to the different subcases. I will review here the ones pertaining to inversion and verb raising, which I have discussed in detail in previous work (Contreras 2001, 2003).

4. Inversion

4.1 Specific problems

The correct application of Inversion requires a couple of adhoc stipulations under the standard account. First, the operation must be prevented from applying in embedded contexts in standard English but not in other dialects, for instance Belfast English or Black English, to account for the different judgments concerning structures like (11):

(11) John wondered had Bill got the letter. (from Henry 1995) (Belfast English)

Second, in those dialects that allow embedded inversion, the operation must be restricted to instances where the attracting C is phonetically null, to prevent structures like (12):

(12) *John wondered if/whether had Bill got the letter.3

4.2 *Alternative analysis*

The problems discussed may be overcome by assuming that there is no T-to-C operation at all. This becomes possible if we make the following reasonable assumptions for English:

³ For alternative proposals to deal with these issues, see Rizzi 1995 and Henry 1995.

- (13) a. Nominative Case is checked in situ if T has no EPP feature.
 - b. The feature [+Q] may be associated with C or T. (Rizzi 1995)
 - c. A [-Q] T always has an EPP feature.
 - d. A [+Q] T may optionally have an EPP feature.
 - e. T and its potential goal must agree in +/-O.
 - f. Matrix sentences are TPs.

Consider matrix sentences. If T is [-Q], it has an EPP feature. Unless there is an expletive in the array, the subject must move, as in the standard account. If T is [+Q], there are two possibilities: a) it has no EPP feature; b) it has an EPP feature. If it has no EPP feature, no movement occurs, and the subject checks its Case in situ. Thus, a structure like (2a), repeated here, is generated simply by successive applications of Merge:

(2) a. Will Chris write the letter?

If T has an EPP feature, it will attract the closest wh-element:

- (14) a. Who will (who) write the letter?
 - b. What will Chris write (what)?

The well-known contrast between subjects, which show no Inversion, and non-subjects, which do, is thus automatically captured.

Consider now embedded clauses. We make the standard assumption that in the unmarked case, predicates select for CP, not TP. In addition, we assume that even though by (13b) the feature [+Q] may be associated with either C or T, clauses are not doubly-marked for [+Q], so if both C and T are present, this feature attaches only to the highest head.

Consider the structures in (15):

- (15) a. We wonder if Chris will write the letter.
 - b. We wonder what C Chris will write (what).

By assumption, the feature [+Q] is in the complementizer *if* in (15a), and in the null complementizer represented by C in (15b). T is represented by *will* in both cases and since it lacks the feature [+Q], it has an EPP feature by (13c).⁴ This forces *Chris* to move from its VP-internal position. Since by

⁴ As a reviewer points out, it must also be ensured that if has no EPP feature but the null C does.

assumption there is no T-to-C, the problem does not arise as to how to prevent this operation from applying.

Let us now look at Belfast English, which allows two possibilities for embedded questions (Henry 1995):

- (16) a. John wondered whether/if Bill will get the letter.
 - b. John wondered will Bill get the letter.

We can derive the properties of Belfast English by assuming that embedded interrogatives can be either CP or TP, and that there is no null C with the feature [+Q] in this language. If the embedded clause is CP, as in (16a), the derivation proceeds as in standard English. For cases like (16b), the embedded clause must be TP under our assumptions. Its head T carries the feature [+Q] and it can optionally have an EPP feature. If it does not have such a feature, no movement is triggered, and the embedded subject checks its Case in situ. If it has an EPP feature, only an agreeing wh-element can be attracted. This is shown in (17):⁵

- (17) a. John wondered who will (who) get the letter.
 - b. *John wondered Bill will get the letter.

As for the Belfast English counterpart to (15b), Henry (1995) reports that some, but not all, speakers exhibit Inversion. The proposal just sketched predicts Inversion for all speakers. It is not clear how to deal with the variation reported by Henry.

5. Verb Raising

5.1 Specific problems

Let us now review some problems specific to verb raising. The verb-raising account of examples like (3) goes back to Emonds (1976). Pollock (1989) reformulates the analysis within the Principles and Parameters model,

(ii) Will Chris write what?

I will assume, following Chomsky (1995: 290), that this is a convergent structure that receives a deviant interpretation, since a question cannot be simultaneously a yes-no question and an information question.

⁵ Since by assumption T with the feature [+Q] has an optional EPP feature, structure (i) is generated:

⁽i) John wondered will who get the letter.

Similarly for main clauses:

and Chomsky (1995) further refines it in minimalist terms. In Chomsky's version, V adjoins to T(ense) overtly in languages like Spanish because T(ense) has a strong V-feature. Since by assumption the V-feature of T in English is weak, there is no overt V-raising.⁶

There are several problems with this account, however. Empirically, it is not true that English entirely disallows the order Verb-Adverb-Complement, as noted by Pesetsky (1989), Ouhalla (1990), Johnson (1991) and others. These authors point out that adverbs may intervene between the verb and its complement as long as the complement is not a DP. The following examples from Pesetsky (1989) and Johnson (1991) illustrate this:

- (18) a. Chris walked quickly down the street.
 - b. Mickey talked slowly to Gary.
 - c. Sam said suddenly that we must all leave.
 - d. Betsy spoke loudly to everyone.
 - e. Mary tried diligently to leave.

On the basis of these facts, Pesetsky (1989) suggests that the verb does raise out of VP in English, but to a lower position than in French (and Spanish). Since there are additional problems with the verb-raising account that Pesetsky's proposal does not obviate, I will not pursue this solution.

Another empirical problem with the verb-raising account of (3) comes from the fact that *be* and auxiliary *have* do seem to raise out of VP, as originally suggested by Emonds (1976), and illustrated in (19):

- (19) a. The director has carefully explained the proposal.
 - b. *The director carefully has explained the proposal.
 - c. The director was carefully explaining the proposal.
 - d. *The director carefully was explaining the proposal.

The problem is: If the attracting feature in T is weak in English, why do *be* and *have* raise overtly? Pollock (1989) proposes a solution in terms of thetarole assignment:

(20) a. English Agr is opaque to theta-role transmission because it is not morphologically rich. If a verb with theta-roles to assign raises, it is unable to assign them, resulting in a Theta Criterion violation.

⁶ For an alternative account that analyzes V-raising as an operatrion that satisfies EPP, see Alexiadou and Anagnostopoulou (1998).

b. *Be* and *have* do not assign any theta-roles, so nothing prevents them from raising.

This solution, however, is not compatible with the minimalist assumption that movement is a last-resort operation.

Chomsky (1993) suggests that *have* and *be* are semantically vacuous, hence not visible to LF operations. Thus, unless they raise overtly, they will not be able to raise at all. There are empirical and theoretical problems with this proposal, however. Wexler (1994) notes that in Swedish auxiliaries pattern with main verbs in that they do not raise out of VP, as shown in (21):

- (21) a. om hon inte ofte har sett honom whether she not often has seen him
 - b. *om hon har inte ofte sett honom whether she has not often seen him
 - c. *om hon inte har ofte sett honom whether she not has often seen him

Presumably, if English auxiliaries are devoid of semantic content, Swedish auxiliaries are also, and under Chomsky's account they should be unable to raise at LF, and they should undergo overt raising instead. From a theoretical point of view, as pointed out by Lasnik (1999) it is questionable that syntactic operations, even those applying at LF, should be sensitive to semantic considerations. It seems fair to say, then, that the behavior of *be* and auxiliary *have* remains mysterious under the verb-raising account of examples like (3).

Baker (2002) has recently pointed out another problem with the standard view of V-raising. Since this operation is independent of whether a language requires the subject to move out of VP, the theory predicts that there should be four types of languages:

- (22) a. V-raising and subject raising.
 - b. V-raising and no subject raising
 - c. Subject raising and no V-raising
 - d. No V-raising and no subject raising

According to Baker, (22a) is represented by French and Italian, (22b) by Welsh and Irish, (22c) by English, but there do not seem to exist any languages of type (22d). Since there is nothing in the theory that would prevent V-raising across a subject that remains *in situ*, this gap is unaccounted for.

5.2 *Alternative analysis*

- 5.2.1 *Merge instead of Move*. I will sketch the alternative presented in Contreras 2003 concerning verb raising. First, I suggest that the cases where the verb appears to have raised across an adverb, as in (3) repeated here, involve Merge, and not Move:
- (3) Pedro examinó cuidadosamente los datos.

 Peter examined carefully the data "Peter examined the data carefully."

Following Contreras and Masullo (2002), I assume that Merge of arguments is triggered by an uninterpretable feature of a specific head. For instance, a transitive verb like *examine* has an uninterpretable [D] feature that is checked when the verb merges with a DP. I further assume that these selectional features may be strong or weak, and that only weak features percolate. If we assume that the D-feature is strong in English but weak in Spanish, we obtain the desired results: *examine* must merge with its object before merging with any adjuncts, otherwise its D feature will not be checked. In Spanish, on the other hand, since the feature is weak and it has the capacity to percolate, it can be either checked by a first merge with the DP or by a second merge after the verb has merged with the adjunct. This predicts correctly that in addition to (3), the version in (23) is well formed in Spanish:

(23) Pedro examinó los datos cuidadosamente.

"Peter examined the data carefully."

Recall that English allows the order Verb-Adverb-Object when the object is not a DP, as shown in (18) above. Under the suggested alternative, this requires that the selectional features for PP and CP be weak in English.

- 5.2.2 Internally-driven Head Movement. Not all of the cases that have been analyzed as resulting from V-raising can be analyzed this way though. In particular, the Merge account proposed in the previous section does not extend to cases where the verb precedes the subject, as in (24):
- (24) (No sabemos si) corre Juan. not we-know if runs John "We don't know if John runs."

For such cases I have suggested that there is a type of verb raising that applies cyclically, and that is triggered by the selectional requirements of Tense, which in Spanish, I claim, is not a separate lexical item but a feature of the verb. I have called that type of raising *syncretic category movement*, but a better label for it might be *internally driven head movement*.⁷

I will now sketch the basis for this analysis. The claim that Tense is not a separate lexical item in Spanish is based on an extension of Warner's (1986) observations concerning the different behavior of main and auxiliary verbs in English. Warner noted that English main verbs count as "identical" for the purposes of VP ellipsis even when their tense inflection differs, but auxiliary verbs require identity of tense, as shown in (25):

- (25) a. John slept, and Mary will [...] too.
 - b. *John was here, and Mary will [...] too.

Lasnik (1999) interprets these facts as suggesting that the lexical entries for main verbs are their bare forms, not the inflected ones as suggested in Chomsky 1995, and that Tense is an independent syntactic element. Under his proposal, Tense and main verb are joined at PF by an updated version of the old Affix Hopping rule. Auxiliary verbs, on the other hand, are entered in the lexicon with their full inflection.

In Spanish all verbs require identity of tense (but not of phi-features) for the purposes of ellipsis, as shown in (26) and (27):

- (26) a. *María escribió una carta ayer, y Pedro [...] mañana.

 Mary wrote a letter yesterday, and Peter [...] tomorrow
 (Cf. Mary wrote a letter yesterday, and Peter will tomorrow.)
 - b. *María estaba escribiendo una carta ayer, y Pedro [...]

 Mary was writing a letter yesterday, and Peter leyendo el diario mañana.

 reading the paper tomorrow
- (27) a. *María escribe poemas, y nosotros* [...] *novelas*. Mary writes poems, and we novels
 - b. *María está escribiendo poemas, y nosotros* [...] *leyendo novelas*. Mary is writing poems, and we reading novels

These facts suggest, following Lasnik's (1999) logic, that Spanish verbs enter the numeration inflected for Tense, although perhaps not for phi-features.

.

⁷ For similar views, see Ackema et al. (1993) and Koeneman (2000).

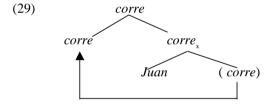
If this is the case, there is no motivation for a separate lexical item Tense. This empirical result can be derived if we assume a strong condition on the existence of phonetically empty lexical items such as (28):

(28) Lexical Parsimony

A lexical item α can lack phonetic features in L, where L is the lexicon of a particular language, only if L includes an item β nondistinct from α which has phonetic features.

Since there is no overt lexical item marking Tense in Spanish, there cannot be null Tense. In English, the existence of *do*, arguably just a Tense marker, makes the existence of null Tense possible.

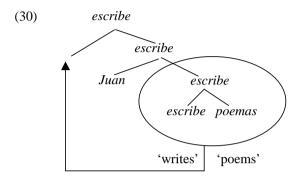
I suggest, however, that Tense universally subcategorizes for vP, regardless of whether it is a separate lexical item or not. Since, by assumption, Tense is not a separate lexical item in Spanish, the only way it can satisfy its selectional restriction is by pied-piping the verb that includes it to a position of sister of vP. This is shown in (29):



The verb *corre* 'runs', with its features [+V, +T], is copied and merged as a sister of the structure labeled $corre_x$ and it projects. The label for the whole structure is that of a complex category with the features +V and +T.

This version of head movement, unlike the standard one, is cyclical, i.e. compatible with Chomsky's (1995) Extension Condition. What about the correct identification of the target? Recall that under a bare phrase structure theory, heads are indistinguishable from their projections in terms of their label. In our example, the question is why raising targets only the head *corre* and not the larger structure labeled $corre_x$. Consider what would happen if this higher projection were to undergo raising: the structure would merge with itself. If we assume that this is not an available option for Merge, we get the correct result.

However, we still incorrectly allow movement of an intermediate projection, as in (30):



A possible solution to this problem is to assume that the [__vP] feature of T is strong, and consequently does not percolate, so only the lowest occurrence of *escribe* 'writes' has it.⁸

5.2.3 Revisiting 'no movement' cases. If this account is correct, we must revise our proposal concerning structures like (3) or (23). It must be the case that in those instances, the verb also undergoes internally-driven movement, given our assumptions. The effects of this movement are not immediately detectable, of course, since the verb is VP-initial. There is empirical evidence in favor of the idea that even in these cases the verb moves out of VP. Consider the fact that Spanish, unlike English, does not admit manner adverbs in VP-initial position, as shown in (31):

- (31) a. *Pedro cuidadosamente examinó los datos.
 - b. Peter carefully examined the data.

Recall that the merger of 'examine' with 'the data' is triggered by the strong D-feature on the verb. The question now is: Couldn't the verb check this strong feature 'late', i.e. after merger with 'carefully' by resorting to internally-driven head movement? The answer is that, since internally-driven movement is cyclical, the verb could only merge with 'examine carefully the data', not with 'the data'. Thus, there is no way for the verb to check its selectional feature after merging with an adjunct.

⁸ This would appear to cause a problem for structures like (i), but it does not.

⁽i) John examined the data carefully.

This contrast follows from the analysis. The Spanish verb *examinó* "examined" must undergo internally-driven movement, so it can only surface in VP-initial position. Since, under Lasnik's (1999) analysis, English main verbs do not include Tense, such movement is not necessary, and by economy conditions, not allowed. English auxiliary verbs, on the other hand, do include Tense, so under our assumptions must undergo internally driven movement. This predicts correctly that a manner adverb cannot precede an auxiliary verb:

- (32) a. Peter has carefully examined the data.
 - b. *Peter carefully has examined the data.

The remaining question is why English auxiliaries and Spanish verbs in general do not always surface to the left of the subject. For English, we must assume that (declarative) T always has an EPP feature, regardless of whether it is a separate lexical item or incorporated within an auxiliary. For Spanish, I assume with many researchers that preverbal subjects are topics, so at no point in the derivation are they constituents of vP.

5.3 Solution to Baker's (2002) problem

To conclude the discussion of V-raising, notice that Baker's (2002) observation concerning the apparent non-existence of languages with no V-raising and no Subject-raising follows from the analysis proposed. If such a language existed, it would exhibit the order Tense/Auxiliary-Subject-Verb-Object. If Tense is a separate lexical item from the Verb, there cannot be internally-driven movement, so the only way for Tense to be associated with the Verb is by PF-adjunction, which requires adjacency. The presence of the intervening subject prevents such an operation. If Tense is a feature on the main verb, the verb must undergo internally-driven movement, so the order VSO is generated, not the order SVO.

6. Short N-movement

Let us now deal with cases like (4), repeated here:

(4) el examen cuidadoso de los datos the examination careful of the data "the careful examination of the data"

⁹ Adjacency must be defined in such a way that adverbs are invisible for PF adjunction of T to V, as suggested by Bobaljik (1994). Otherwise, structures like *Mary silently left the room* could not be generated.

Cinque (1994) and others have analyzed such cases as involving adjunction of N to a functional category below D, but some problems have been pointed out by Lamarche (1991). The feature responsible for this movement is assumed to be strong in Spanish, but weak in English. These cases bear a strong resemblance to the putative cases of V-raising illustrated by (3):

(3) Pedro examinó cuidadosamente los datos.

This suggests that both cases should be analyzed in a similar way. The contrast between (4) and its English counterpart follows immediately if we assume that the selectional feature of *examination* is strong while that of *examen* is weak. Since by assumption only weak features percolate, Spanish has the option of merging *examen* directly with its object or delaying such merger until after the noun merges with the adjective. Under the latter option, (4) is generated. If the first option is taken, we generate either (33a) or (33b):

- (33) a. el cuidadoso examen de los datos
 - b. el examen de los datos cuidadoso

While (33a) is much more natural than (33b), I will assume that they are both convergent, since (33b) improves considerably once we replace the simple adjective with a heavier phrase:

- (34) el examen de los datos más cuidadoso que conocemos the examination of the data most careful that we-know "the most careful examination of the data that we know"
- 6.1 Differences between V-raising and short N-raising
- 6.1.1 English. There are a couple of differences between the putative cases of V-raising and those of 'short' N-raising. First, English N, unlike V, does not distinguish between different categories of object. Objects headed by of, which are sometimes analyzed as DP with of as a Case marker, behave the same as objects headed by true prepositions. Unlike the verbs in (18), the nouns in (35) do not tolerate an adjunct between them and their complement:

- (35) a. the sudden announcement that Bill was leaving
 - b. *the announcement sudden that Bill was leaving
 - c. the direct appeal to the mayor
 - d. *the appeal direct to the mayor

This difference follows directly from the assumption that the selectional features of N are always strong in English.

- 6.1.2 *Spanish*. Another difference is illustrated by the contrast in Spanish between (31a) and (33a) repeated here:
- (31) a. *Pedro cuidadosamente examinó los datos.

 Peter carefully examined the data
- (33) a. el cuidadoso examen de los datos the careful examination of the data

This follows from the natural assumption that, unlike verbs, Spanish nouns have no feature comparable to Tense which would trigger internally-driven movement. Consequently, nothing forces a Spanish noun to move to the NP-initial position.

Finally, consider the account of the variation in (36):

- (36) a. el cuidadoso examen de los datos
 - b. el examen cuidadoso de los datos

Under the N-raising account, the existence of these two variants poses a problem, since it suggests that N-movement is optional. To circumvent this problem, proponents of N-movement have suggested that adjectives like *cuidadoso* must be allowed to merge in two different positions, either to the left of NP or to the left of NumP. Under the alternative proposed here, it is not necessary to posit two different insertion sites for adjectives like *cuidadoso*. The options are an automatic result of the two orders of Merge which the analysis makes available.

7. Long N-movement

Longobardi (1994, 1996) has suggested that proper names and some exceptional common nouns undergo N-to-D movement in Italian, as shown in (5), repeated here:

- (5) a. *L' antica Roma fu la città più importante del Mediterraneo.* the ancient Rome was the most important city of the Mediterranean
 - b. *Antica Roma fu la città più importante del Mediterraneo. ancient Rome was the most important city of the Mediterranean
 - c. Roma antica fu la città più importante del Mediterraneo. Rome ancient was the most important city of the Mediterranean
 - d. La Roma antica fu la città più importante del Mediterraneo. the Rome ancient was the most important city of the Mediterranean (examples from Longobardi 1994)

As (5a) and (5d) show, when the determiner is present, the adjective may either precede or follow the noun, but when there is no overt determiner (5b, 5c), only the order noun+adjective is allowed. Longobardi (1994) accounts for this paradigm by assuming that these structures must contain a determiner, and that in Italian the determiner position must be filled in order to avoid a quantificational interpretation which would be in conflict with the features of proper names.

In English, however, the order proper name+adjective is ill-formed, as shown in (37):

- (37) a. Old John came in.
 - b. *John old came in.

Longobardi suggests that English tolerates empty determiners, so there is no need for an N-to-D operation.

Since this account is also incompatible with the restricted view of movement assumed in current versions of the minimalist program, we must explore an alternative. I will suggest that this is another example of internally-driven head movement. The following assumptions would appear to be sufficient to account for the basic facts:

- (38) a. Proper names are [+N, +D].
 - b. The feature [+D] is uninterpretable in N.
 - c. [+D] can be checked in the domain of an agreeing Determiner.
 - d. If there is no independent determiner, [+D] can only be checked by Merge (with N).
 - e. English has null determiners, but Italian does not.

Under these assumptions, (37a) has a null determiner, which checks the [+D] feature in *John*, so movement is unnecessary. In (5a) and (5d), the overt determiner checks the [+D] feature of *Roma*, and movement is unnecessary. However, in (5c), since by assumption there is no null determiner, the only way the [+D] feature of *Roma* can be checked is by internally-driven head movement, which allows [+D] to be checked by merging with a constituent of type N, that is, *Roma antica*.

8. Parametric variation

Finally we must address the question of parametric variation. In previous versions of the minimalist program, one source of parametric variation was the relative strength of the features responsible for head movement. Having argued against the standard version of head movement, we must look for parameters elsewhere.

Chomsky (2000) has suggested two sources of parametric variation: first, different languages select different subsets of features from the pool made available by Universal Grammar; second, different languages may organize the selected features into lexical items in different ways. As an example of the first type of parameter, a language like English does not select the feature [+honorific] while a language like Japanese does.

A possible example of the second type has been argued for in the preceding discussion: both English and Spanish select the feature Tense from the universal pool, but they treat it differently. In English, Tense may be a separate lexical item or, in the case of auxiliaries, it may be included among the features of the verb; in Spanish, Tense is always a feature of the verb, not a separate lexical item.

Our discussion of Merge suggests another possible parameter: a selectional feature may be strong or weak, determining the order of Merge. Thus, we have claimed that the D feature of English transitive verbs is strong while in Spanish it is weak.

Finally, languages may vary in their linearization possibilities: in both English and Spanish, nouns may combine with adjective phrases to yield nominal expressions. In both languages, when the adjective contains a complement, the order is N+AP. However, when the adjective has no complement, English requires the order A+N, while Spanish established no fixed linear order, as shown in (39) and (40):

- (39) a. people proud of their achievements
 - b. *proud of their achievements people
 - c. gente orgullosa de sus logros
 - d. *orgullosa de sus logros gente
- (40) a. proud people
 - b. *people proud
 - c. gente orgullosa
 - d. orgullosa gente

It is not always easy to distinguish cases involving different orders of Merge from those involving different linearization. Consider the examples in (41):

- (41) a. We consider Mary very smart.
 - b. *We consider very smart Mary.
 - c. Consideramos a María muy inteligente.
 - d. Consideramos muy inteligente a María.

An account in terms or order of Merge would be as follows: In English, the small clause must be formed first, before it merges with the higher predicate; in Spanish, you either form the small clause first or construct a complex predicate first. An account in terms of linearization would claim that in both cases the small clause is formed first, and that at PF, English requires the order subject+predicate while Spanish does not establish a linear order. The fact that the interpretation of both Spanish versions is the same suggests that the difference is a PF matter.

9. Conclusions

The restricted view of syntactic movement suggested in Chomsky 2000 and 2001 makes the standard view of head movement anomalous. A close look at Inversion, V-Raising, short N-movement and long N-movement suggests that they fall into three classes:

- (42) a. No movement (English Inversion, short N-movement in Romance)
 - b. PF merger (English Affix Hopping (Halle and Marantz 1993, Lasnik 1999))
 - Internally-driven head movement (Spanish V-Raising, long N-movement in Romance (Longobardi 1994, 1996))

Internally-driven head movement, unlike the standard version, obeys the Extension Condition and can correctly identify its target within the restrictions of a bare phrase structure.

References

- Ackema, Peter, Ad Neeleman & Fred Weerman. 19 93. "Deriving Functional Projections". *Proceedings of the North East Linguistic Society* 23. 17-31.
- Alexiadou, Artemis & Elena Anagnostopoulou. 1998. "Parametrizing AGR: Word Order, Verb-Movement, and EPP Checking". *Natural Language and Linguistic Theory* 16. 491-539.
- Baker, Mark. 1988. Incorporation. Chicago: University of Chicago Press.
- Baker, Mark. 2002. "Building and Merging, not Checking: the Nonexistence of (AUX)-S-V-O Languages". *Linguistic Inquiry* 33. 321-328.
- Bobaljik, Jonathan David. 1994 "What Does Adjacency Do?" MIT Working Papers in Linguistics 22. 1-32.
- Chomsky, Noam.1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. "Minimalist Inquiries: the Framework". *Step By Step. Essays on Minimalist Syntax in Honor of Howard Lasnik* ed. by R. Martin et al., 89-155. Cambridge, Mass.: MIT Press..
- Chomsky, Noam. 2001. "Derivation by Phase". *Ken Hale. A Life in Language* ed. by M. Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo . 1994. "On the Evidence for Partial N-Movement in the Romance DP". *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne* ed. by G. Cinque et al., 85-110. Washington, D.C.: Georgetown University Press.
- Contreras, Heles. 2001. "A New Perspective on Inversion". To appear in the Proceedings of the 11th Colloquium on Generative Grammar. Universidad de Zaragoza, Spain.
- Contreras, Heles. 2003. "Verb Raising as Syncretic Category Movement". *Proceedings of the 13th Western Conference on Linguistics* ed. by Lesley Carmichael, Chia-Hui Huang, and Vida Samiian, 86-103. Department of Linguistics, California State University, Fresno, California.
- Contreras, Heles & Pascual José Masullo. 2002. "Motivating Merge". *Current Issues in Generative Grammar* ed. by Manuel Leonetti, Olga Fernández Soriano & Victoria Escandell Vidal. Universidad Alcalá de Henares. Servicio de Publicaciones.
- Emonds, Joseph. 1976. A Transformational Approach to English Syntax. New York: Academic Press.

- Halle, Morris & Alec Marantz. 1993. "Distributed Morphology and the Pieces of Inflection". *The View From Building 20* ed. by K. Hale & S.J. Keyser, 111-176. Cambridge, Mass.: MIT Press.
- Henry, Alison. 1995. *Belfast English and Standard English*. New York and Oxford: Oxford University Press.
- Johnson, Kyle. 1991. "Object Positions". *Natural Language and Linguistic Theory* 9:4. 577-636.
- Koeneman, Olaf. 2000. *The Flexible Nature of Verb Movement*. LOT, Netherlands Graduate School of Linguistics.
- Lamarche, Jacques. 1991. "Problems for N⁰-Movement to Num-P". *Probus* 3. 215-236.
- Lasnik, Howard. 1999. *Minimalist Analysis*. Malden, Mass. and Oxford: Blackwell Publishers
- Longobardi, Giuseppe. 1994. "Reference and Proper Names". *Linguistic Inquiry* 25. 609-665.
- Longobardi, Giuseppe. 1996. "The Syntax of N-Raising: a minimalist theory". *OTS Working Papers*, Utrecht.
- Ouhalla, Jamal. 1990. "Sentential Negation, Relativised Minimality and the Aspectual Status of Auxiliaries". *The Linguistic Review* 7:2. 183-231.
- Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP". *Linguistic Inquiry* 20:3. 365-424.
- Pesetsky, David. 1989. "Language-particular Processes and the Earliness Principle". Ms., MIT, Cambridge, Mass.
- Pesetsky, David & Esther Torrego. 2001. "T-to-C Movement: Causes and Consequences". *Ken Hale. A Life in Language* ed. by M. Kenstowicz, 355-426. Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1995. "Residual Verb Second and the Wh Criterion". *Parameters and Functional Heads* ed. by L. Rizzi, & A. Belletti, 63-90. New York: Oxford University Press.
- Warner, Anthony. 1986. "Ellipsis Conditions and the Status of the English Copula". *York Papers in Linguistics* 12.153-172. University of York, Heslington, England.
- Wexler, Kenneth. 1994. "Optional Infinitives, Head Movement, and the Economy of Derivations". *Verb Movement* ed. by David Lightfoot & Norbert Hornstein, 305-350. Cambridge: Cambridge University Press.

THE EFFECTS OF PHONOLOGICAL CUES ON THE SYNTAX OF FOCUS CONSTRUCTIONS IN SPANISH*

LAURA DOMINGUEZ

Boston University

1. Introduction

In this paper I am interested in the mechanisms that constrain focus and word order in Spanish. In particular I investigate the possibility that prosody may place its own constraints on the realization of focus in this language. Accordingly, I will argue in favor of an analysis of focus as a category represented both in Syntax and Phonology. In some Romance languages the focused element must receive stress which is assigned by the Nuclear Stress Rule (NSR) (Chomsky and Halle 1968, Cinque 1993, Reinhart 1997) to the rightmost (or most embedded, constituent in a sentence. In English, however, relocation of stress is an available mechanism to mark with stress a constituent which is not final (Reinhart 1995). Since relocation of stress is not available in Romance, Spanish uses movement of the nonfocal material as a primary focus marking device (Zubizarreta 1998) to ensure that focused elements are aligned with nuclear stress. In contrast to prior work, in this paper I present evidence that Spanish uses more than one type of movement for marking a constituent with focus.

In recent years, arguments have been given in favor of analyzing focus in Romance as a phonological category. ¹ Taking these studies as a starting point, I propose that focus is both phonologically and syntactically realized in Spanish as well. Therefore I present an analysis of focus marking in Spanish which is

^{*} Thanks to Victor Manfredi for the suggestion to study the correlations between word order, pitch and focus.

¹ See also Jun and Elordieta (1997) for Basque, Frota (2000) for Portuguese, Frascarelli (2000) for Italian. See also Büring and Gutiérrez-Bravo (2001) for an OT analysis of the prosodic constraints of word order in Spanish, German and English.

constrained by prosody (Nespor and Vogel, 1986) as well as Syntax. I argue that focus prominence correlates to pitch accents in Spanish, providing an alternative to the NSR which is heavily dependent on the linear ordering of constituents. I will also argue that an analysis of the sentence at the phonological phrase level is required in order to provide an accurate account of the realization of Focus in Spanish and its effects on the order of constituents. Therefore, I argue that movement of constituents in Spanish can be motivated by phonological factors, in particular alignment with stress and high pitch contours in initial and final positions. It is also possible that a particular word order is preferable in certain contexts in order for a phrase to avoid being aligned with a high pitch peak, as in the case of the subject in VSO orders. This will be presented as evidence that focus in Spanish is realized both syntactically and phonologically and that word order in focus constructions is affected by prosodic and well as syntactic constraints. Section 2 presents an analysis of the syntactic properties of focus in Spanish. Section 3 presents an analysis of new experimental data in order to establish a correlation between pitch contours and the ordering of focused phrases. Section 4 provides the discussion of findings and the conclusions.

2. The syntactic properties of Focus in Spanish

Focus and word order are closely interrelated in Spanish and although SVO is the canonical word order, other configurations, such as VSO, OVS and VOS are possible in this language as well. It may seem *a priori* that subjects are allowed to appear freely in different positions. However, it has been argued that subject position is highly constrained by focus marking requirements (Casielles, 1998, Contreras, 1976, Zubizarreta, 1998). The sentences in (1) illustrate the different focus marking mechanisms available in Spanish:²

(1)	a.	[Susana abrió el <u>libro]</u>	SVO
		"Susana opened the book."	
	b.	Susana abrió [el <u>libro</u>]	SVO
	c.	[El libro] abrió <u>Susana</u>	OVS
	d.	Abrió el libro [<u>Susana]</u>	VOS
	e.	#[Susana] abrió el <u>libro</u>	#SVO
	f	(El libro), lo abrió [<u>Susana</u>]	O-Cl-V-S
		the book, it opened Susana	

² In this paper focused phrases appear in brackets and underlined phrases represent phrases which receive nuclear stress.

In the sentences illustrated in example (1) the main stress falls on the final constituent of the sentence. However, the object, el libro, in example (1b), already in its canonical final position, cannot be marked with unambiguous narrow focus. This is the reason why sentences (1a) and (1b) are ambiguous with respect to whether they show a broad and a narrow focus interpretation on the object.³ However, in sentence (1c) the preposed object is marked with narrow focus. Therefore, it is possible for the focused element not to be aligned with nuclear final stress and still be marked with focus as example (1c) illustrates with the preposed object. When focused subjects or objects appear in initial position they are obligatorily associated with an emphatic or contrastive reading, that is why example (1b) is preferred over example (1c), although (1b) is ambiguous with respect of the scope of focus (narrow or broad). In the same sense, example (1e) with the focused subject in its canonical position cannot be the answer to Who opened the book? which has narrow focus on the subject, and requires the subject to appear in final position as in examples (1d) and (1f). However, (1e) would be felicitous in a context where the subject is contrastive or has a corrective interpretation, such as the answer to the question 'Did Juan open the book?' Finally, example (1f) illustrates a common strategy used in Spanish and other Romance languages (see Vallduví (1992) for Catalan). Spanish uses clitic-left dislocations very often in order to leave the presupposed information out of the sentence. This common strategy is preferred by most speakers over a sentence like (1d) where the old information remains in the sentence. Notice that both sentences (1d) and (1f) are felicitous to a question such as Who opened the book? and in both sentences the focused subject appears in final position. This seems to indicate that non-focal phrases have the option to remain in the sentence or move out of the intonational

According to Zubizarreta the subject in (iv) is not marked by the NSR but by the Emphatic/Contrastive Stress. Therefore the subject can be marked with focus in an SVO configuration if and only if it is associated with a contrastive meaning:

It is also important to note that Zubizarreta derives the VOS word order from VSO; therefore, VSO, and not SVO is the canonical word order in Spanish in her analysis. Thus, in order to obtain a VOS configuration the subject stays in situ [Spec, VP] and the object is moved right to the left of the subject via scrambling. See Ordóñez (1997) for an alternative account of VOS in Spanish.

³ The subject cannot appear in a SVO word order and be marked with narrow information focus (i) since it must always appear in final position (ii): What has happened?

^{*[}Susana] abrió el libro.

Abrió el libro [Susana]

Did John open the book? iii.

No, [Susana] abrió el libro.

phrase that contains the main focus. As we see in these examples, Spanish uses syntactic movement to realize focus and more than one movement type is available.

There is a long tradition in the literature of Romance languages to characterize focus as a primarily syntactic phenomenon where defocalized (or presupposed) constituents move out of their canonical position in order to escape being part of the focus set (see Vallduví (1992) for Catalan, Zubizarreta (1998) for Spanish and Costa (1998) for Portuguese). Following the idea that movement is a primary device in focus marking in Spanish, Zubizarreta presented an analysis of focus which shows the narrow relationship between the order of constituents and their alignment with nuclear stress, which always falls in final position. Therefore, constituents move so that the focused phrase always appears in final position. This type of movement is then prosodically motivated (p-movement):⁴

- (2) Canonical Word Order (S-V-O-PP):
 - a. Susana leyó el libro en la <u>biblioteca</u>
 "Susana read the book in the library"
- (3) Application of p-movement (S-V-PP-O):

What did Susana read in the library?

- a. Susana leyó en la biblioteca_i [el <u>libro</u>] t_i
- b. #Susana leyó [el <u>libro</u>] en la biblioteca.

In example (3a) the prepositional phrase, *en la biblioteca*, must move out of its final position so that the object receives stress. This operation is necessary in Spanish since a non-final constituent cannot be marked with information focus *in situ* (3b). Zubizarreta's analysis is crucial to understand the availability of VOS in Spanish. In particular, it accounts for the fact that subjects which carry the new (non contrastive) information must appear in final position. However, we find that in addition to the new information, an element marked with contrastive focus can appear in final position as well:

(4) Was it Antonio that saw the accident?

No, lo vio [Carlos]

no it saw Carlos

"No, Carlos saw it."

⁴ It is important to note that in Zubizarreta's analysis it is the defocalized material that undergoes movement and not the focused phrase. Therefore, the motivation of this type of movement cannot be checking a feature.

Interestingly, the contrastive subject has the option of appearing in initial position as well:

(5) No, [Carlos] vio el accidente. no Carlos saw the accident "No, Carlos saw the accident."

Therefore we see that the marking of information focus in Spanish does not correlate to a structural position of a sentence (such as final position) but to a position where prominence is assigned. These examples seem to pose a problem for an account of focus which is based on the NRS. We see that in example (5) the subject is focused and it is not in final position. This is not the only case where a similar discrepancy between nuclear stress and narrow focus is found. In certain type of questions subjects can appear in initial position (6a)⁵ as well as in their usual postverbal position (6b). Only the subject in final position can have a contrastive reading:

- (6) a. ¿[Juan] qué quiere?

 Juan what wants

 "What does Juan want?"
 - b. Qué quiere <u>Juan</u>?

In the following examples, the final constituent is associated with a contrastive reading as well:

- (7) a. Al terminar Juan, le dijo que se fuera Antonio (no Carlos). upon finishing Juan, him said that him go Antonio "When Juan finished, Antonio (not Carlos) asked him to leave."
 - b. Al terminar Juan, Antonio le dijo que se <u>fuera</u> (no que se quedara). "When Juan finished, Antonio asked him to leave (not stay)."

We see that only in some cases contrastive focus may overlap with the syntactic positions usually available for new narrow focus. Therefore, it is not possible to characterize all cases of narrow focus (new and contrastive) as being aligned with main stress or appearing in a determined position in the sentence.

⁵ Sentence (6a) must be read without a pause between the subject and the question word.

We have already seen that subjects appear in VOS word orders when they are marked with information focus. Another case where subjects appear in final postverbal position is with intransitive verbs. In these cases the subject has the option to appear in initial or final position, but as we would expect, subjects can only be narrow-focused when they are postposed (example 8b). As expected, example (8c) with a focused subject is not allowed as an out-of-the-blue statement:

(8)	a.	[El correo ha <u>llegado</u>]	S-V
		"The mail has arrived"	
	b.	Ha llegado [el <u>correo</u>]	V-S
	c.	#[El <u>correo</u>] ha llegado	S-V

Some speakers allow (8b) as a neutral declarative which would be felicitous as an answer to *What happened?* It seems, however, that having a postposed subject in this context carries an additional meaning (contrastive or exhaustive) which is not found in sentences like (8a) with the subject in initial position. Thus, (8b) may imply that it is only the mail which has arrived and nothing else (it shows exhaustiveness), whereas we do not find this interpretation in (8a). In any case, subjects in V-S configurations are subject to have only a narrow focus reading whereas S-V orders are ambiguous between a contrastively focused subject and an all-sentence focus interpretation.

As we see in VS and VOS orders the subject is always marked with narrow focus. In contrast, VSO is the only word order where a postverbal subject is not associated with a narrow focus interpretation. Therefore, the next example, with a subject in postverbal position can only be felicitous to a question that can be answered with a broad focus sentence or with narrow focus on the object:

```
(9) Who read the book?

#Leyó [Susana] el libro.

read Susana the book."

"Susana read the book."
```

In the literature VSO has also been considered to be a canonical word order in Spanish, since sentences with this word order are usually associated

with broad focus readings.⁶ An interesting fact concerning subjects in VSO is that they are not aligned with stress and therefore the fact that they appear in postverbal position cannot be motivated by focus.⁷ This contrasts with the case of postverbal subjects in VS constructions (example (8)). Belletti (1999) argues that free inversion of the subjects in VS word orders in Italian is licensed by an 'internal focus' in the low IP area. This is easily corroborated by the Italian data since all cases of free inversion are associated with focus. However, Spanish poses a problem for the theory of internal focus since in VSO word orders the inverted subject cannot be marked with narrow focus, and therefore not all cases of inversion of the subject can be licensed by focus. Incidentally, VSO is not a possible word order in Italian. So the status of postverbal subjects remains open in Spanish.

In the following section I will show that subjects appear in these different configurations in order to satisfy certain prosodic requirements. In order to account for this fact we need to look at the prosodic characteristics of Spanish.

3. Prosodic characteristics of Focus in Spanish

In this section I argue that an analysis of focus prominence based on pitch, rather than sentence stress, can account not only for sentences with new narrow focus (information focus), but also with contrastive and emphatic focus. Also important is the fact that an account of focus based on pitch provides a better understanding of the prosodic requirements of focus constructions since it relieves it from the heavy syntactic dependency of the NSR.

The existence of mismatches between syntactic constituents and the domain where phonological rules apply is taken to be evidence in favor of considering a separate level of representation: the prosodic structure. It mediates between the syntactic and phonological structures and interacts with them via mapping rules (Inkelas and Zec, 1990, Nespor and Vogel, 1986, Selkirk, 1984). This prosodic structure is a hierarchy composed by different constituents (syllable, foot, phonological word, the clitic group, the phonological phrase, the intonational phrase and the phonological utterance). In this paper I will analyze the phonological ($_{\phi}$) and intonational ($_{I}$) phrases. (10) and (11) are examples of the analysis of the prosodic structure in English and in Spanish:

⁶ See Suñer (1982), Groos & Bok-Bennema (1986), Hernanz and Brucart (1987), Demonte (1994), and Zubizarreta (1998) for a discussion on the status of the canonical word order in Spanish.

⁷ It is often assumed that the subject in VSO stays in [Spec, VP], whereas the subject in SVO moves to [Spec, IP]. See Uribe-Etxebarria (1992) and Zagona (2002) for discussion.

```
(10) [[John]_{\varphi} [bought]_{\varphi} [candy]_{\varphi}]_{I} (11) [[Juan]_{\varphi} [compro)_{\varphi} [caramelos]_{\varphi}]_{I}
```

Languages can restructure phonological phrases so that certain elements appear with their heads. In example (10) the verb and its complement may appear in the same phonological phrase after restructuring has applied. In general, in Spanish each phonological phrase bears its own pitch accent (Büring and Gutiérrez-Bravo, 2001) although the absence of a pitch accent may imply that the phrase is not new information. Thus, it has been proposed that a language such as Spanish uses pitch accents in order to mark different focus structures (Sosa, 1999). In order to understand the prosodic characteristics of focus in Spanish we need to look at two important features: the structure of phonological phrases and the analysis of pitch contours.

3.1 Phonological Phrases in Spanish

In his study of Spanish intonation, Sosa (1999) argues that Spanish favors short phonological phrases due to the fact that the main stress, which always falls on final position, cannot fall on any other phrase in the sentence. Since the word that is situated immediately before a pause between phrases always receives some prominence Spanish uses restructuring and short phonological phrases to give some prominence to words that are not in final position. The following example illustrates different possible structures of the same sentence. The underlined words are the ones which appear to the left of a phonological phrase boundary and therefore are able to receive prominence. Interestingly, in example (12b), the object, *reloj*, cannot receive any prominence since it is not at the edge of a boundary:

- (12) a. [Juan] [miró] [su reloj] [con preocupación]
 - b. #[Juan] [miró] [su reloj con preocupación] "Juan looked at his watch with concern."

Thus, restructuring may be used in cases where elements are not in focus in order to escape the effects of prominence, as in the case of the subject in a VSO word order:

(13) [Miró] [Juan su <u>reloj</u>] looked Juan his watch "Juan looked at his watch" In any other configuration (SVO, SV, VOS) the subject will be in a position to the left of a phrasal boundary and therefore will be subject to receive some prominence.

3.2 Intonation in Spanish

Several studies on Spanish intonation (Face, 2001, Sosa, 1999) have found that stessed syllabes are usually associated with a pitch accent.. Face's work on contrastive focus also reveals the use of early peak alignments as an active device for signaling contrastive focus in Spanish. Face also shows how words in broad focus in initial position have higher peaks than narrowly focused words in other positions. Accordingly, we should expect phrases marked with focus to be associated with high pitch accents in different positions in a sentence (either in final position if they are non-contrastive or in any position if they are emphatic or contrastive). At the same time, we should also expect non-focused phrases not to be associated with pitch accents, as in the case of subjects in VSO orders.

To investigate this I analyzed the characteristics of focalized and non-focalized subjects in an array of word orders to look for correlations between the position of focal and non-focal subjects in the prosodic phrase and their alignment with high pitch accents. The corpus consisted of five sets of two sentences with focal and non focal phrases in final and non final position. A perception test was administered prior to the experiment to validate the focus interpretation. The subjects were three female native speakers of Iberian Spanish. They were asked to read each sentence three times in a natural way and each sentence was introduced by a short paragraph to contextualize the sentences. The analysis of the data corroborates the existence of one initial and one final high rise on the stressed syllables. Figure 1 shows the analysis of *El técnico instaló la lavadora* "The technician installed the washing machine". We see how the highest peak is associated with the subject, *técnico*, in initial position, and how the verb does not show a focal pitch accent:

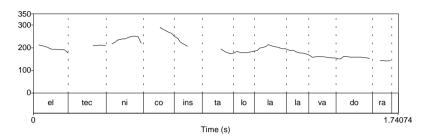


Fig. 1: Pitch analysis of "El técnico instaló la lavadora"

The data also show that defocalized constituents appear in the plateau between peaks as illustrated In Figure 2 where the stressed syllabe in la lavad'ora does not bear a pitch accent:

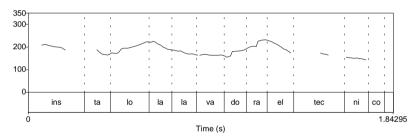


Fig. 2: Pitch analysis of "Instaló la lavadora el técnico"

As expected, the stessed syllabe of the subject in a VSO word order does not bear a pitch accent as illustrated in Figure 3:

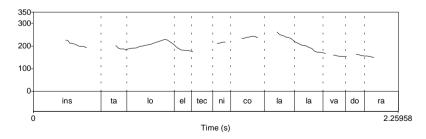


Fig. 3: Pitch analysis of "Instaló el técnico la lavadora"

The previous results contrast with the analysis of contrastive phrases. As expected, in the sentence *Al terminar Juan*, *llegó Antonio* ("When Juan finished, Antonio arrived") *Juan*, which is in focus, bears the highest pitch of the sentence. It is also important to note that the high rise is aligned with the stressed syllable of *Juán* showing a L+H* contour:

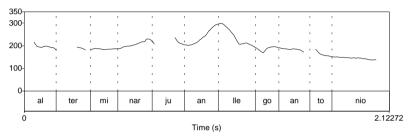


Fig. 4: Pitch analysis of "Al terminar Juan, llegó Antonio"

Thus, the data indicate that focused phrases are marked prosodically by high pitch contours and that narrow focused phrases (new and contrastive) can appear in final position where, besides receiving main prominence, they bear a focal pitch accent (L*+H). As expected, contrastive phrases can appear in a position which is not final and they bear a L+H* focal accent as proposed by Face (2001). On the other hand, defocalized elements escape the prosodic marking of focus by not bearing focal pitch accents. Thus, the analysis of the data shows that subjects appear in positions which are licensed by prosodic requirements. This can account for the wide variety of positions of subjects in Spanish and provides evidence for direct effects of prosody on word order.

4. Conclusions

In this paper I have argued that focus is both a syntactic and a phonological phenomenon in Spanish. I have analyzed the different types of movement available in Spanish and I have shown that there is more than one syntactic movement available. I have also demonstrated the existence of correlations between phonological requirements and word order which provides evidence that word order is affected by both syntactic an prosodic requirements. Finally, I have presented evidence that an analysis of focus in terms of pitch can provide a unified account for all cases of narrow focus, both new and contrastive.

References

- Belletti, Adriana. 1999. "Inversion as Focalization and Related Questions". *Catalan Working Papers in Linguistics* 7.9-45.
- Büring, Daniel, & Rodrigo Gutiérrez-Bravo. 2001. "Focus Related Word Order Variation Without the NSR: A Prosody-Based Crosslinguistic Analysis". *Syntax at Santa Cruz.* 41-58. University of California, Santa Cruz: Linguistic Research Center.
- Casielles, Eugenia. 1998. "Focus Preposing (it is called)". *Proceedings of the Workshop on Focus 21*, ed. by Elena Benedicto, Maribel Romero and Satoshi Tomioka, 51-64. University of Massachussetts Occasional Papers.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper.
- Cinque, Guglielmo. 1993. "A Null Theory of Phrase and Compound Stress". Linguistic Inquiry 24.239-267.
- Contreras, Heles. 1976. A theory of Word Order with Special Reference to Spanish: North-Holland linguistic series 29. Amsterdam & New York: North-Holland.
- Costa, João. 1998. *Word Order Variation: A Constraint-Based Approach*. The Hague: Holland Academic Graphics.
- Demonte, Violeta. 1994. *Teoría Sintáctica: de las Estructuras a la Rección*. Madrid: Síntesis.
- Face, Timothy. 2001. "Focus and Early Peak Alignment in Spanish Intonation". *Probus* 13.223-246.
- Frascarelli, Mara. 2000. The Syntax-Phonology Interface in Focus and Topic Constructions in Italian. Dordrecht: Kluwer.
- Frota, Sonia. 2000. *Prosody and Focus in European Portuguese*. PhD Dissertation. Lisboa.
- Groos, Anneke & Reineke Bok-Bennema. 1986. "The Structure of the Sentence in Spanish". *Generative Studies in Spanish Syntax*, ed. by Ivonne Bordelois, Heles Contreras and Karen Zagona, 67-80. Dordrecht: Foris.
- Hernanz, Maria Lluisa & José Maria Brucart. 1987. *La Sintaxis (1)*. Barcelona: Editoral crítica.
- Inkelas, Sharon & Draga Zec. 1990. "Prosodically Constrained Syntax". *The Phonology-Syntax Connection*, ed. by Sharon Inkelas and Draga Zec, 365-378. CSLI Publications and the University of Chicago Press.
- Jun, Sun-Ah & Gorka Elordieta. 1997. "Intonational Structure of Lekeitio Basque". *Intonation: Theory, Models, and Applications. Proceedings of the Esca workshop*, ed. by Antonis Botinis, Georgios Kouroupetroglou and George Carayannis, 193-196. ESCA/University of Athens.
- Nespor, Marina & IreneVogel. 1986. Prosodic Phonology. Dordrecht: Foris.

- Ordóñez, Francisco. 1997. Word Order and Clause Structure in Spanish and other Romance Languages. PhD Dissertation. CUNY.
- Reinhart, Tanya. 1995. "Interface Economy: Focus and Markedness". *Studia Grammatica* 40.47-169.
- Selkirk, Elisabeth. 1984. *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.
- Sosa, Juan Manuel. 1999. La entonación del español. Madrid: Cátedra.
- Suñer, Margarita. 1982. *Syntax and Semantics of Spanish*: Washington: Georgetown University Press.
- Szendröi, Kriszta. 2001. *Focus and the Syntax-Phonology Interface*. Doctoral Dissertation. University College London.
- Uribe-Etxebarria, Myriam. 1992. "On the Structural Positions of the Subject in Spanish, their Nature and their Consequences for Quantification". *Syntactic Theory and Basque Syntax*, 447-491. Donostia.
- Vallduví, Enric. 1992. The Informational Component. New York:Garland.
- Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.
- Zubizarreta, Maria Luisa. 1998. *Prosody, Focus, and Word Order*. Linguistic inquiry monographs 33. Cambridge, Mass.: MIT Press.

OPTIONAL INFINITIVES OR SILENT AUXES? NEW EVIDENCE FROM ROMANCE*

CRISTINA DYE Cornell University

1. Introduction

In this paper I investigate the range of variation in ostensibly nonfinite matrix verbs in child language through a comparison of English, Spanish, French and Italian naturalistic speech samples. Such a comparison brings to light certain patterns which are not visible by examining one individual language alone, nor by focusing on only one nonfinite form type. I show that there is a correspondence between the occurrence of a certain type of ostensibly nonfinite verb form in a given child language and the existence in the corresponding target language of a periphrastic construction embedding the same type of nonfinite verb. On the basis of my results, I propose that child utterances with ostensibly nonfinite matrix verbs are modeled on utterances with periphrastics in the target language, but that the inflected element remains unpronounced; this lends support to previous studies arguing for a null auxiliary (e.g., Boser, Lust, Santelmann & Whitman 1992; Josefsson 1999).

2. Ostensibly nonfinite matrix verbs in child language

Across several languages, two-year old children have been reported to use ostensibly nonfinite verbs in matrix sentences in contexts where the target

^{*} I would like to thank Carol Rosen, Barbara Lust, John Whitman, and Yasuhiro Shirai for their invaluable suggestions and support. I am also grateful to Elma Blom, Maria Blume, Wayles Browne, Allan Dye, Claire Foley, James Gair, Kleanthes Grohmann, Wayne Harbert, Ingeborg Lasser, Gunlög Josefsson, Yumiko Nishi, Margarita Suñer, the audiences at the Cornell Linguistics Colloquium (November 21, 2002) and at Going Romance 2002, and to an anonymous reviewer. I thank Andres Lema-Hincapie for coding the Spanish data and Anastasia Riehl for assistance with editing.

¹ Here I report some preliminary results. See Dye (forthcoming) for the full study.

 $^{^2}$ I use 'periphrastic' as a cover term for constructions involving *auxiliary* + *nonfinite verb* and certain combinations of *main verb* + *nonfinite verb*, many of which are treated by Rizzi (1982) under the rubric 'restructuring predicates.'

(1) Monsieur conduire.

(7) Why dey sidding in de water?

grammar requires finite verbs (while also using finite forms that conform to the adult language). Initial research on this phenomenon focused on ostensible matrix infinitives (OMI's) illustrated in (1)-(3). Not all apparent matrix nonfinite forms are infinitives, however. Child language also shows ostensible matrix past participles (OMP's), e.g., (4) and (5). Additionally, ostensible matrix present participles/gerunds (OMG's) have been reported; these are exemplified in (6) and (7):

French (Pierce 1992)

English (Santelmann et al. 2000)

	"Mister drive-INF"	
(2)	Eve sit ³ floor.	English (Hoekstra &Hyams1996)
(3)	Muchas recoger.	Spanish (Ezeizabarrena 2002)
	"Many pick-INF"	
(4)	Reh gelauf.	German (Boser et al. 1992)
	"Deer run-PRT"	
(5)	Perdut llapis.	Catalan (Bel 2001)
	"Lose-PRT pencil"	
(6)	Aquest dormint.	Catalan (Bel 2001)
	"this one sleep-GER"	

Within a UG framework, two opposing approaches have been posited to account for children's ostensibly nonfinite matrix sentences. Proponents of the *Root/Optional Infinitive Hypothesis (ROIH)* claim that children may optionally use infinitives in contexts where the adult grammar requires finite verbs, perhaps because in children's grammar certain functional projections might be missing/optional/underspecified, subject to maturation (e.g., Rizzi 1994; Wexler 1994; Hoekstra & Hyams 1996). According to the second approach, which I will refer to as the *Silent Aux/Modal Hypothesis (SAMH)*, in children's ostensibly nonfinite matrix utterances the nonfinite verb is licensed by an auxiliary/modal that remains silent/unpronounced (e.g., Boser et al. 1992; Josefsson 1999). These two views entail contradictory interpretations of examples (1)-(7): under ROIH they are syntactically non-finite, having non-adult representations, while under SAMH they are syntactically finite, having adultlike representations with regard to UG, but have an unpronounced element.

³ This verb form is ambiguous between an infinitive and a bare stem. Here I consider it an infinitive, based on the arguments in Phillips (1995) and Borer & Rohrbacher (2002), but I do not exclude the possibility that some such forms may be bare stems.

3. Data and methods

I reviewed natural speech samples from eleven monolingual, normally developing children from the CHILDES database (longitudinal samples) and from the Cornell Language Acquisition Lab/CLAL database (cross-sectional samples) (Table 1). For the quantitative analyses of OMP's and OMI's (Tables 6 and 8) a smaller number of sessions was used, which is indicated in the column titled 'subset.' In selecting this subset I aimed at having similar ages and comparable numbers of utterances across the four languages.

Language	Child and age	Subset	Corpus	Database
	(# of sessions)	(# of sessions)		
English	Eve 1;.9.0 & 2;3.0 (2)	Eve 1;9.0 & 2;3.0 (2)	Brown	CHILDES
Spanish	Magín 1;7-2;3.2	Magín 1;10.16, 1;10.27 & 2;3.2 (3)	Aguirre	CHILDES
	Alfonso 2;3.7-2;10.22 (3)	-	Marrero/ Albala	CHILDES
	Maria 1;7-2;7 (13)	-	Lopez-Ornat	CHILDES
	Eduard 1;4-3;4 (11)	-	Serra /Sole	CHILDES
Italian	Raffaello 1;7.07-	Raffaello 1;7.07, 1;9.7,	Calambrone	CHILDES
	2;11.20 (17)	1;10.20, 2;0.10, 2;1.15, 2;3.14 & 2;4.29 (7)		
	Camilla 2;2.6-3;4.9 (7)	Camilla 2;2.6 (1)	Antelmi	CHILDES
French	Grégoire 1;9.28-2;5.27		Champaud	CHILDES
	(10) Phillipe 2;1.19-3;3.12 (33)	2;3.0 (3)	Leveille	CHILDES
	BM 2;6 (1)	-	Foley	CLAL
	ED 2;8 (1)	-	Foley	CLAL
	KD 2;10 (1)	-	Foley	CLAL

Table 1: Subjects

Utterances were coded following the procedures described in the Cornell University Virtual Linguistics Lab Research Methods Manual. The basis for the present analyses were all utterances with overt verbs, except for unclear utterances and imitations of adult utterances. ⁴

My hypothesis is that if the target language has a periphrastic construction involving a given type of non-finite verb, then the child language may evidence ostensible matrix verbs involving that particular non-finite form. If a given type

⁴ In French past participle and infinitive forms of first conjugation verbs are homphonous, e.g., /kupe/ could be either *couper* or *coupe* 'cut'. I counted as infinitives those forms which were transcribed as infinitives and as participles those forms which were transcribed as participles. Further details regarding coding criteria can be found in Dye (forthcoming).

of nonfinite verb is not used in periphrastic constructions in the target language, then it will not occur as an ostensible matrix verb in child speech, even if the language does have the morphological form. This is because the child may allow the aux/modal in the periphrastic to remain unpronounced (see previous findings regarding empirical evidence for positing unpronounced auxes/modals, e.g., Josefsson 1999; Dye et al. 2002).

4. Results and analysis

I present the results regarding each of the three types of nonfinite forms, i.e., OMG, OMP, OMI, in subsections 4.1, 4.2, and 4.3, respectively. For each form type, my main objective is to determine in which child language it does/does not occur and how this does/does not correlate with the types of periphrastics available in the target language. Secondarily, I discuss the occurrence rates of a given nonfinite form and (for OMI) its semantics across the four child languages as further support for the link with periphrastics.

4.1 OMG's

In the data that I have examined, examples of OMG's are found in child English, Spanish or Italian but *not* in child French (Table 2):

Language	Periphrastics w/ gerund	Child OMG's	
English	I am writing	(8) Eve writing.	(Eve 1;9.0)
Spanish	Estoy escribiendo	(9) Haciendo así a enanito "Do-GER so dwarf"	(Alfonso 2;10.22)
Italian	Sto scrivendo	(10) <i>Io mi metto qui, e ascoltando questo.</i> "I me put here and listen-GER this"	(Camilla 2;9.4)
French	NONEXISTENT		(Phillipe, BM, KD Gregoire, ED)

Table 2: Distribution of periphrastics with gerund in the target languages and distribution of OMG's in the child languages

My results are corroborated by previous studies attesting the existence of OMG's in English (e.g., Blom & Krikhaar 2002), Spanish (e.g., Bel 2001), and Italian (e.g., Pizzuto & Caselli 1992). ⁵ In contrast, *none* of the previous studies on child French mention OMG's, as far as I know, including studies that take into account both OMI's and OMP's, such as Levow (1995). The non-occurrence of OMG's in French, as opposed to the other three languages, corresponds to the fact that English, Spanish and Italian have periphrastics with gerunds while French does not. Notice that (adult) French does have gerund morphology, e.g., *repondant* 'answer-GER'. The crucial difference between

⁵ The authors only mention the occurrence of gerunds (p.520), without giving examples.

French and the other three languages thus lies not in the absence of the gerund form but rather in the absence of periphrastics with gerund.

Table 3 shows the proportions of OMG's in the three child languages in which they occur, i.e., in child English, Spanish and Italian.⁶ The data show that the proportion of OMG's is higher in child English than in child Spanish or Italian:

English	15.7%
Spanish	< 1%
Italian	< 1%

Table 3: Proportions of OMG's out of matrix utterances with overt verbs

An inspection of the three target languages with regard to the use of the periphrastic with gerund (i.e., the present progressive) reveals an interesting pattern. The present progressive is in fact used quite differently across the three languages. To refer to an ongoing activity simultaneous with the moment of speech, English requires the periphrastic with gerund, while Italian and Spanish allow either the simple present tense or the present progressive. The simple present is restricted to a habitual reading in English. In contrast, in Italian and Spanish, the simple present encompasses both the continuous imperfective interpretation and the habitual reading, the periphrastic with gerund being mainly employed to stress imperfectivity and durativity (Ledgeway 2000:99-100). The present progressive/the periphrastic with gerund is thus more common in English, where it is the basic present tense, than in Spanish or Italian where it has aspectual value.

It is therefore apparent that there is a correspondence, across languages, between the proportion of child OMG's and the distribution of periphrastics with gerunds in the respective target language⁷. OMG rates are highest in child English where periphrastics with gerund are very common in the target language, lower in child Spanish and Italian where periphrastics with gerunds

⁶ Due to space limitations, in this paper the number of occurrences of a given form is collapsed over sessions and children. For a fuller analysis see Dye (forthcoming).

⁷ It is important to note here that this correspondence is visible when one looks across several languages. I do not expect the same percentage for child OMG's and adult periphrastics with gerund within a given language, because all children also have periphrastics with overt auxes and because there is no reason to expect that child discourse is identical to adult discourse. Additionally, there may be cross-linguistic differences in the omission rates of the aux, possibly due to phonological factors. These comments also apply to the discussions on OMP's and OMI's.

are less common in the target language, and nonexistent in child French where there are no periphrastics with gerunds in the target language.

4.1.1 Further evidence on OMG's Additional support for my hypothesis regarding OMG's comes from data from published studies on child Catalan, Basque, German and Dutch. Table 4 opposes Catalan and Basque against German and Dutch:

Language	Periphrastics w/	Child OMG's	
	gerunds		
Catalan	Estic escribint	(11) Aquest dormint.	(G 2;4, Bel 2001)
		"This one sleep-GER"	
Basque	Idazten ari naiz	(12) Txoria jaten.	(DG 3;1, Austin 2001)
-		"Bird eat-IMP(GER)"	
German	NONEXISTENT	NONEXISTENT	(Lasser p.c., Lasser 1997)
Dutch	NONEXISTENT	NONEXISTENT	(Blom & Krikhaar 2002)

Table 4: Distribution of periphrastics with gerunds in the target languages and distribution of OMG's reported in the child languages

OMG's have been reported in child Catalan and Basque, corresponding to the fact that Catalan and Basque have periphrastics with gerund. In contrast, OMG's are not attested in child German or Dutch, corresponding to the fact that German and Dutch do not have periphrastics with gerunds. Notice that, similar to French, both German and Dutch do have gerund morphology, e.g., German *kommend* 'come-GER' and Dutch *helpend* 'help-GER.'

To summarize, the data reveal that the distribution of OMG's across the child languages correlates with the distribution of periphrastics with gerunds in the corresponding target languages.

4.2. OMP's

I found occurrences of OMP's in all four child languages. Examples of OMP's for each child language are given in (13)-(16) in Table 5:

Language	Periphrastics w/	Child OMP's	
	past participle		
English	I have written	(13) It gone away.	(Eve 2;3.0)
Spanish	He escrito	(14) <i>Lo roto</i> . "It break-PRT"	(Magín 1;10.27)
Italian	Ho scritto	(15) Messo latte. "Put-PRT milk"	(Raffaello 2;3.14)
French	J'ai écrit	(16) Auto parti. "Car go-PRT"	(BM 2;6)

Table 5: Distribution of periphrastics with past participles in the target languages and distribution of OMP's in the child languages

My findings are supported by previous studies which have reported OMP's in English (e.g., Radford 1994), Spanish (e.g., Bel 2001), Italian (e.g., Giannelli & Manzini 1996), and French (e.g., Levow 1995; Dye et al. 2002). The occurrence of OMP's in all four child languages corresponds to the existence of periphrastics with past participle in all four target languages.

An analysis of the rates of OMP's indicates that there are differences across the four child languages (Table 6). According to my data, OMP's are more common in child French or Italian than in child English or Spanish:

English	1.1%
Spanish	1.9%
Italian	6.7%
French	5.3%

Table 6: Proportions of OMP's out of matrix utterances with overt verbs

This state of affairs in the child languages corresponds to an important difference between French or Italian and English or Spanish with regard to the use of periphrastics with past participles. Namely, in order to refer to past events, colloquial varieties of Italian and French rely (almost) exclusively on the analytic past tense which consists of the auxiliary *be/have* and a past participle, e.g., French *j'ai vu* and Italian *ho visto*. (In these varieties the synthetic past/the simple past is no longer in use, the analytic past having become an all-purpose past tense.) In contrast, English and Spanish use both the synthetic past (*I saw*, *vi*) and the analytic past (*I have seen*, *he visto*). The analytic past/the periphrastic with past participle is thus more common in French or Italian than in English or Spanish ⁹ (Harris 1982).

To summarize, I have demonstrated that OMP's occur in all four languages. The proportions of OMP's are higher in child French and Italian where periphrastics with past participles are quite common in the target language, and lower in child Spanish and English where periphrastics with past participles are somewhat less common in the target language. The distribution of OMP's across the child languages therefore correlates with the distribution of periphrastics with past participles in the corresponding target languages.

⁸ Since the figures in Tables 6 and 8 are based on a subset of my data (see Table 1), further calculations are needed to confirm them.

⁹ There are dialectal differences for both Spanish and English regarding the distribution of the periphrastic past. For both languages, the European variety relies more on the periphrastic past than does the American variety (but still less than do French or Italian). Since my data are from Peninsular Spanish and American English, one might have expected a slightly higher OMP rate for Spanish than for English (though still lower than for French and Italian).

4.3 *OMI's*

Similar to OMP's, I found OMI's to also occur in all four child languages. In Table 7 I provide examples of OMI's from each child language:

Language	Child OMI's	
English	(17) Eve make tower.	(Eve 1;9.0)
Spanish	(18) Ya me a acostar. "Now me to sleep-INF"	(Magín 2;3.2)
Italian	(19) <i>Io anae foi</i> . "I go-INF outside"	(Raffaello 2;3.14)
French	(20) Toi venir. "You come-INF"	(Grégoire 2;3.0)

Table 7: Distribution of OMI's across the child languages

My findings are in agreement with previous studies which have documented OMI's in English (e.g., Blom & Krikhaar 2002), Spanish (e.g., Bel 2001; Ezeizabarrena 2002), Italian (e.g., Giannelli & Manzini 1996), and French (e.g., Levow 1995). The occurrence of OMI's in all four child languages corresponds to the existence of periphrastics with infinitive in all four target languages (see Table 9 below).

Table 8 shows the rates of occurrence of OMI's in the four languages.

English	18.4%
Spanish	8.3%
Italian	3.5%
French	10.8%

Table 8: Proportions of OMI's out of matrix utterances with overt verbs

The results in Table 8 are in line with what has been previously reported in the literature, though, of course, due to methodological differences, only broad comparisons are possible with other studies. Proportions of OMI's are usually reported to be higher for English and lower for Italian than for other languages, which is also the case here. The rates for French, which are comparable to those reported in recent studies (e.g., Jakubowicz & Rigaut 2000:12; Levow 1995:286), are somewhat closer to the rates for Spanish than for English (cf. Bel 2001:1299 who also points out that rates for French and Spanish may be similar).

Differences in OMI rates across the four child languages may be related to differences among the four target languages with regard to the number of construction types involving periphrastics with infinitive. As summarized in

Table 9, English has the most construction types involving periphrastics with infinitive, French and Spanish have fewer, and Italian has the least.

Language	Modal periphrastics	Future periphrastics	Do-support
English	e.g., I can write	8	I do write
	e.g., Puedo escribir	-	NONEXISTENT
Italian	e.g., Posso scrivere	NONEXISTENT	NONEXISTENT
French	e.g., Je peux écrire	Je vais écrire ¹⁰	NONEXISTENT

Table 9: Distribution of periphrastics with infinitives in the target languages

Furthermore, not only does English have the most construction types involving periphrastics with infinitives, but crucially, one of these types, i.e., *do*-support, is in fact overgeneralized in child speech (e.g., Hollebrandse & Roeper 1996). This phenomenon, by now widely documented, is illustrated in (21) where a semantically vacuous *do* is inserted before the main verb:

(21) I do have juice in my cup. (Tim 2;11) (Hollebrandse & Roeper 1996)

The pattern now emerging is this: the highest OMI rates are found in English, the language with the most opportunities for infinitive use (i.e., the most construction types involving periphrastics with infinitive) and the lowest OMI rates are found in Italian, the language with the least contexts for infinitive use (the least construction types involving periphrastics with infinitives).

There is an additional piece of evidence in support of this interpretation of the data. German and Dutch are usually reported to have relatively high rates of OMI's. This is directly accounted for under the present analysis: (colloquial) German and Dutch pattern with English, having modal and future periphrastics as well as *do*-support, and crucially, in both languages children are known to overuse pleonastic *do*. Example (22) illustrates *tun*-insertion in child German and example (23) illustrates *doen*-insertion in child Dutch. Notice that what is important for my argument here is the fact that children overuse pleonastic *do* in English, German and Dutch, and not so much the distribution of *do*-support in the three adult languages. I am suggesting that this (nonadult) use of *do*-support in children may be responsible for a number of OMI's when *do* remains unpronounced. (For further discussion of the aux-insertion phenomenon in child German and Dutch, see Boser et al. 1992 and Zuckerman et al. 2000, respectively.)

 $^{^{\}rm 10}$ Colloquial French is known to prefer the analytic future to the synthetic future.

(22) Suzanne tat den Schneeball werfen. (age 3;5) (Boser et al. 1992) Suzanne did the snowball throw-INF.

(23) *Ik doe ook verven.* (Niek 3;10) I do also paint-INF

(Zuckerman et al. 2000)

4.3.1 Independent evidence: The semantics of OMI's as a clue to their identity Semantic analyses provide independent evidence for the link between OMI's and periphrastics with infinitives. In recent work, Blom & Krikhaar (2002) have convincingly shown through an experimental study that OMI's in Dutch and English may have both modal/future reference and past reference, depending on the discourse context. The initial analyses I have undertaken with regard to French ostensibly nonfinite matrix verbs show a different picture. As shown in Table 10, there is a semantic difference between verb forms which are morphologically unambiguous infinitives, i.e., 2nd/3rd conjugation verbs, and 1st conjugation verb forms which are ambiguous between a past participle and an infinitive (and which I refer to as E-forms):

Morphological form	Future/ modal	Past	Undeterminable
	reference	reference	reference
2 nd /3 rd conjug. infinitives	67% (2/3)	0% (0/3)	33% (1/3)
E-forms	63% (5/8)	37% (3/8)	0% (0/8)

Table 10: Semantics of French 2nd/3rd conjugation infinitives and E-forms¹¹

2nd/3rd conjugation infinitives occur with modal/future, but not past reference. In contrast, *E*-forms occur with either past (24) or future/modal (25) reference:

- (24) Il reveillE. [il s'est reveillé] (after making the toy stand up) (BM 2;6) "he wake-E" ["he woke up"]
- (25) Attends, je envE. [attends, je vais enlever] (before removing pants) (ED 2;8) "Wait I remove-E" ["wait, I'm gonna remove"]

This difference between $2^{nd}/3^{rd}$ conjugation infinitives and E-forms is due to the fact that the latter include both past participles that are part of periphrastics with $avoir/\hat{e}tre$ expressing past tense, and infinitives that are part of periphrastics with aller/vouloir/devoir/falloir expressing future/modality.

I now return to Blom & Krikhaar's (2002) facts for English and Dutch OMI's and compare them to my present findings for French (Table 11):

¹¹ Thus far, quantitative analyses on the semantics of OMI's have been performed on three children, namely the Foley/CLAL subjects (which were also discussed in Dye et al. 2002).

Semantics	English (or Dutch)	Periphrastics w/ infinitive	French ¹²	Periphrastics w/infinitive
Modal/future	+	I 'll/must write	+	Je vais/dois écrire
Past	+	I did write	-	-

Table 11: Correspondence between the semantics of OMI's and the existence of periphrastics with infinitives expressing the same meaning

In English, OMI's occur with both modal/future and past meaning corresponding to the fact that English has periphrastics with infinitive expressing both meanings. In French, OMI's occur mainly in modal/future contexts, corresponding to the fact that French has periphrastics with infinitives for expressing modal/future meaning but not for expressing past meaning. Semantic analyses of OMI's thus offer additional support for the connection between OMI's and periphrastics with infinitives. Some recent studies confirm this. Giannelli & Manzini (1996:213) found that Italian OMI's occur in modal contexts, which is expected based on the facts in Table 9. Ezeizabarrena (2002) reports that while Spanish OMI's tend to occur with modal/future reference, Basque OMI's additionally occur with past reference; this corresponds to the fact that in Basque, as opposed to Spanish, a periphrastic with infinitive is used to express past tense. ¹³

To conclude, the distribution of OMI's across the child languages appears to correlate with the distribution of periphrastics with infinitives across the corresponding target languages.

5. Summary and discussion

OMG's occur in child English, Spanish and Italian but do not occur in child French; OMP's and OMI's occur in all four child languages. This pattern corresponds to the existence of periphrastics with gerund in English, Spanish, Italian but not French, and to the existence of periphrastics with past participles and infinitives in all four target languages. The occurrence of a given form type therefore appears to depend on the existence in the target language of a periphrastic involving that nonfinite form. Differences in the rates of ostensibly nonfinite matrix forms across the four child languages, as well as the crosslinguistic variation in the meaning of OMI's, are also suggestive of the connection with periphrastic constructions. These facts cannot be explained under the ROIH where nonfinite utterances are expected to occur in free variation with finite ones. Instead, they support the hypothesis that children's ostensibly nonfinite utterances are modeled on utterances with periphrastics in

¹² Considering only unambiguous infinitives.

¹³ See Bel (2001) for somewhat different results for Spanish.

which the aux/modal remains unpronounced. OMI's, OMP's, & OMG's are *different* facets of the *same* phenomenon, i.e., aux/modal omission. The type of nonfinite verb form that will occur (as a matrix verb) in a given child language may thus be predicted based on the type of nonfinite forms used in periphrastic constructions in that language.

This conclusion corroborates other kinds of evidence from previous proposals regarding a silent aux/modal, e.g., for German OMI's & OMP's (Boser et al.1992), Swedish OMI's & OMP's (Josefsson 1999), English (interrogative) OMI's & OMG's (Guasti & Rizzi 1996), French OMI's & OMP's (Ferdinand 1996; Dye et al. 2002), Spanish and Basque OMI's (Ezeizabarrena 2002). The place/form correlation observed across different child languages (i.e., finite forms raise, nonfinite forms do not) is directly predicted under the SAMH. (Under checking theory, this correlation is not logically entailed by ROIH, rather as Borer & Rohrbacher 2002 point out, a Tense deficit predicts *random* inflection errors).

The weight of the evidence, based on many and varied sources, seems to now point to the SAMH. Adopting this paradigm not only eliminates fundamental problems such as the issue of optionality in UG (which is incompatible with the Minimalist Program and which raises learnability issues), but also opens the door to a new set of issues to explore. For example, is the modal/aux sometimes left unpronounced because the child is in the process of learning the specific morpho-phonological forms, or are there certain contexts for aux/modal omission?

Analyzing children's utterances with ostensibly nonfinite verbs as being covertly finite eliminates the need to posit different representations^{14,15}. Regardless of what the exact reason(s) behind children occasionally not pronouncing the aux/modal may be, it is not due to a deficit in UG but rather to development in language-specific knowledge (e.g., see Demuth 1994 for possible phonological factors, or Dye et al. 2002 for some pragmatic factors

¹⁴ OMI's have been claimed to have certain 'properties' that allegedly distinguish them from finite sentences, e.g., non-occurrence with subject clitics or wh-questions. These claims need to be approached with caution. Such 'properties' seem to also hold for OMP's (Levow 1995), suggesting that if they indeed exist, they may be characteristic of null-aux structures in general. 5-8 year old L2 learners' OMI's are also said to display them, suggesting that they cannot be due to maturation (Prevost & White 2000). Some have been disconfirmed, e.g., the claim that OMI's tend to occur with null subjects (e.g., Jonas 1995). See also Phillips (1995), Ezeizabarrena (2002), and for a comprehensive discussion, Dye (forthcoming).

¹⁵ It is possible, in principle, that although in regard to the issues under investigation here children's representations are similar to adults, in other areas of UG children differ from adults; this remains to be investigated. The present study shows that children's ostensibly non-finite utterances are not a reason for positing differences in child UG representations.

involved). This coheres with the theory of acquisition proposed in Lust (1999), Santelmann et al. (2000), where development in child language is seen as due to lexical learning and/or to the integration of language-specific knowledge with UG modules.

6. Conclusion

This study begins to fill the existing gap regarding comparative studies on children's ostensibly nonfinite forms. It has contributed new evidence in support of the SAMH. The data analyzed here are not compatible with the ROIH where the existence of children's ostensibly nonfinite forms is assumed to be biologically determined. The present results indicate that, on the contrary, the existence of these forms depends on (the existence of periphrastics in) the target/input language.

Current work is under way to extend the analyses reported here to larger samples and to test them through experimental procedures (see Dye forthcoming). Among the advantages of the present line of research, compared to the ROIH paradigm, are that it provides a unified account of the different types of ostensibly non-finite matrix verbs in the four child languages (giving support to previous attempts at a unified explanation, e.g., Josefsson (1999), Bel (2001)) and that it makes it possible to explain i) the differences in rates and meaning of these forms across the different child languages; ii) why a given child language may demonstrate more than one type of form; iii) why certain child languages show certain types of forms and not others. Most importantly, it eliminates the problematic issue of optionality in regard to UG.

References

- Austin, Jennifer. 2001. Language Differentiation and the Development of Morphosyntax in Bilingual Children Acquiring Basque and Spanish. PhD. Dissertation, Cornell University.
- Bel, Aurora. 2001. "The Projection of Aspect: A key in the Acquisition of finiteness?" *Research on Child Language Acquisition. Proceedings of IASCL* ed. by M. Almgren et al.,1297-1313. Sommerville Mass.: Cascadilla Press.
- Billerey, Roger, & Lillehaugen, eds. 2000. *Proceedings of WCCFL 19*. Somerville, Mass.: Cascadilla Press.
- Blom, Elma & Evelien Krikhaar. 2002. "Where to Use a Nonfinite Clause? An Experiment on the Use of Verb Forms in Early Dutch and English Child Language". *Proceedings of GALA 2001* ed. by J. Costa & J. M. Freitas.
- Borer, Hagit & Bernhard Rohrbacher. 2002. "Minding the Absent: Arguments for the Full Competence Hypothesis". *Language Acquisition* 10.123-175.

- Boser, Katarina, Barbara Lust, Lynn Santelmann, & John Whitman. 1992. "The Syntax of CP and V-2 in Early child German: The Strong Continuity Hypothesis". *Proceedings of NELS* ed. by K. Broderick, 22.51-66. Amherst: GLSA.
- Brown, Roger. 1973. *A first language : The early steps*. Cambridge, Mass. Harvard University Press.
- Cipriani et al. 1989. *Protocolli diagnostici e terapeutici nello sviluppo e nella patologia del linguaggio* (1/84 Italian Ministry of Health) Stella Maris.
- Demuth, Katherine. 1994. "On the Underspecification of Functional Categories in Early Grammars". Lust et al. 1994. 119-134.
- Dye, Cristina. Forthcoming Ph.D. Dissertation, Cornell University.
- Dye, Cristina, Claire Foley & Barbara Lust. 2002. "Dissociating Finiteness from Morphology: New Evidence from the Acquisition of Verbal Inflection in French". Paper presented at LSRL 32, Toronto, April 19-21.
- Ezeizabarrena, Maria J. 2002. "Root Infinitives in Two Pro-Drop Languages". *The Acquisition of Spanish Morpho-syntax: The L1/L2 Connection* ed. by A. T. Perez-Leroux & J. Liceras. Dordrecht: Kluwer.
- Giannelli, Grazia & Rita Manzini 1996. "The Prefunctional Stage in the Light of Minimalism". *Proceedings of GALA*, Groningen, 7-9 Sept. 1996 ed. by C. Koster & F. Wijnen, 207-216. Groningen: Centre Language & Cognition Groningen.
- Guasti, Maria Teresa & Luigi Rizzi 1996. "Null Auxes and the Acquisition of Residual V2". *Proceedings of BUCLD 20* ed. by A. Stringfellow, D. Cahana, E. Hughes & A. Zukowsly, 284-295. Somerville, Mass.: Cascadilla Press.
- Harris, Martin. 1982. "The Past Simple and the Present Perfect in Romance". *Studies in the Romance Verb* ed. by N. Vincent & M. Harris, 42-70. London: Croom Helm.
- Hoekstra, Teun & Nina Hyams 1996. "Missing Heads in Child Language" Koster & Wijnen 1996. 251-260.
- Hollebrandse, Bart & Thomas Roeper. 1996. "The Concept of DO-Insertion and the Theory of Infl in Acquisition". Koster & Wijnen 1996. 261-270.
- Jakubowicz, Celia & Catherine Rigaut. 2000. "L'acquisition des clitiques nominatifs et des clitiques objets en français". *Canadian Journal of Linguistics* 45.119-157.
- Jonas, Dianne. 1995. "On the Acquisition of Verb Syntax in Child Faroese". *MIT Working Papers in Linguistics*, 26. 280-285.
- Josefsson, Gunlög. 1999. "Non-Finite Root Clauses in Swedish Child Language". *Proceedings of BUCLD 23* ed. by A. Greenhill et al., 313-324. Somerville Mass.: Cascadilla Press.

- Koster, Charlotte & Frank Wijnen, eds., 1996. *Proceedings of the Groningen Assembly on Language Acquisition*, Netherlands: Center Language and Cognition Groningen
- Lasser, Ingeborg 1997. Finiteness in Adult and Child German. PhD Dissertation, CUNY.
- Ledgeway, Adam. 2000. A Comparative Syntax of the Dialects of Southern Italy: A Minimalist Approach. Boston & Oxford: The Philological Society.
- Levow, Gina-Anne. 1995. "Tense and Subject Position in Interrogatives and Negatives in Child French: Evidence For and Against Truncated Structure". *MIT Working Papers in Linguistics* 26.281-304.
- Lust, Barbara, Maria Blume & Tina Ogden. In prep. "Cornell University Linguistics Lab Research Methods Manual: Scientific Methods for the Study of Language Acquisition. Cornell University.
- Lust, Barbara 1999. "Universal Grammar: The Strong Continuity Hypothesis in First Language Acquisition". *Handbook of Child Language Acquisition* ed. by W. Ritchie and T. Bhatia, 111-156. New York: Academic Press.
- Lust, Barbara, Margarita Suner & John Whitman, eds. 1994. *Syntactic Theory and First Language Acquisition* vol. 1. Hillsdale: Lawrence Erlbaum.
- MacWhinney, Brian & Catherine Snow. 1990. "The Child Language Data Exchange System: An update". *Journal of Child Language* 17.457-472.
- Phillips, Colin 1995. "Syntax at Age Two: Cross-linguistic differences". *MIT Working Papers in Linguistics* 26.325-382.
- Pierce, Amy E. 1992. *Language Acquisition and Syntactic Theory*. Dordrecht: Kluwer Academic Publishers.
- Pizutto, Elena & Maria Caselli. 1992. "The Acquisition of Italian Morphology: Implications for Models of Language Development". *Journal of Child Language*, 19.491-557.
- Prevost, Philippe & Lydia White. 2000. "Accounting for Morphological Variation in Second Language Acquisition: Truncation or Missing Inflection". *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* ed. by M. Friedmann & L. Rizzi, 202-235. Essex: Pearson Education.
- Radford, Andrew 1994. "Tense and Agreement Variability in Child Grammars of English". Lust et al. 1994. 135-158.
- Rizzi, Luigi 1982. "A Restructuring Rule". *Issues in Italian Syntax*, 1-48. Dordrecht & Cinnaminson: Foris.
- Rizzi, Luigi 1994. "Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives". *Language Acquisition* 3.371-394.

- Santelmann, Lynn, Sthephanie Berk, & Barbara Lust. 2000. "Assessing the Strong Continuity Hypothesis in the Development of English Inflection: Arguments for the Grammatical Mapping Paradigm". Billerey & Lillehaugen 2000. 439-452.
- Suppes, P., R. Smith & M. Leveille. 1973. "The French Syntax of a Child's Noun Phrases". *Archives de Psychologie* 42.207-269.
- Wexler, Ken 1994. "Optional Infinitives, Head Movement and the Economy of Derivations". *Verb Movement* ed. by David Lightfoot & Norbert Hornstein, 305-362. London: Cambridge University Press.
- Zuckerman, Shalom, Roelien Bastiaanse & Ron van Zonneveld. 2000. "Auxiliary Insertion in Child Dutch". Billerey & Lillehaugen 2000. 631-644.

DISLOCATION, CLITIC RESUMPTION AND MINIMALITY A COMPARATIVE ANALYSIS OF LEFT AND RIGHT TOPIC CONSTRUCTIONS IN ITALIAN *

MARA FRASCARELLI Università degli Studi di Roma Tre

1. Introduction

This paper is concerned with the syntax of clitic-resumed dislocated constituents, which represent 'given' information. We will refer to such constituents as 'Topics' (as is now common in the literature) and, consequently, the relevant construction will be referred to as 'topicalization'.

Topic constituents in Italian can be realized in either periphery of the sentence. Nonetheless, left and right dislocation (henceforth, LD and RD) show crucial structural asymmetries. The purpose of this paper is to compare and discuss the properties of LD and RD, in order to provide a comprehensive account of these constructions and shed light on the connection between syntax and information structure.¹

In the generative literature a movement analysis is generally assumed for Topics in Italian, where Topic constituents are merged in argument position and then moved into an extrasentential maximal projection, the 'Topic Phrase' (TopP), for which different locations have been proposed. Rizzi (1997) indicates two Topic fields in the left periphery of the sentence, one above and the other below the Focus Phrase (FocP), while Benincà (2001) claims that there is only one TopP position in the CP layer. As for right-dislocated constituents, Cecchetto (1999) and Belletti (2001) support the existence of a lower TopP node, just above the VP. So, according to a movement analysis, a

^{*} I wish to express special thanks to Annarita Puglielli and Roland Hinterhölzl, for their precious comments and support. Many thanks also to Adriana Belletti, Anna Cardinaletti, Carlo Cecchetto, Cécile De Cat, Caterina Donati, Luigi Rizzi and two anonymous reviewers for their helpful suggestions and criticism. General disclaimers apply.

¹ Indeed, a major problem in the understanding of Topic constructions is a limited comparative analysis and the widespread assumption of a 'mirror approach', according to which what is known about LD is also assumed for RD (cf., for instance, Vallduvì's 1990 'mirror image').

Topic is generated within IP and can reach one of the positions indicated in (1) below (the asterisk indicates that the TopP projection can be recursive):

(1)
$$[T_{OPP} * TOPIC_k [F_{OCP} [T_{OPP} * TOPIC_k [IP [T_{OPP} * TOPIC_k [VP t_k]]]]]]$$

Challenging this leading approach, this paper shows that an analysis of topicalization in terms of movement from an argument position – in which LD and RD are only distinguished by a different landing site – can account neither for the syntactic properties of Topics nor for their interface interpretation.

The first important claim of this paper is that clitic LD in Italian is the effect of Merge, rather than Move. This proposal thus restates the claim put forth in Cinque (1990), though with basic differences concerning Local Merge and TopP-to-TopP movement. The second major claim is that RD is derived from LD by IP-inversion. These claims will be supported by theoretical as well as empirical arguments.²

2. Topicalization as Merge

2.1 *The proposal* (I)

According to the analysis we are going to discuss in the following sections, topicalization is a *Merge operation*. Specifically, a Topic is base-generated in the TopP projection which immediately dominates the FocP; hence, at no point of the derivation is the Topic in argument position.

We also propose that the relevant TopP projection is *the only Topic field* available in the sentence (cf. (2) below), so that we exclude the presence of additional TopP projections, either below the FocP or above the VP. This means that RD Topics are also merged in the TopP projection located in the left periphery of the sentence and that *RD constructions are derived through IP-inversion*, as shown in (3):

- (2) LD $[T_{OPP} *TOPIC_k [F_{OCP} [IP [VP [DP cl_k [NP pro_k]]]]]]]$
- (3) RD $\left[_{GP}\right]_{IP}\left[_{VP}\right]_{DP} cl_k\left[_{NP}\right]_{Pro_k}\left[_{TopP}\right]_{TOPIC_k} t_{IP}$

The validity of IP-inversion to derive RD will be supported in this paper by sound evidence concerning scope, binding and Minimality facts, while the analysis of its nature and

² For reasons of space we will concentrate on Italian data here, postponing the comparative issue to future work (for important crosslinguistic differences, the interested reader is referred to Browning 1996, Culicover 1996 and Delfitto 2002).

³ IP-inversion to derive RD in Italian was previously put forth in Frascarelli (2000) and a similar proposal is also present in Cardinaletti (2002), though within a movement approach to topicalization. For the purposes of the present paper, the GP node simply indicates a functional projection in the C-system hosting IP-movement.

In the next sections we will first consider the common properties of LD and RD, in order to discuss the most immediate problems connected with a movement analysis.

2.2 WCO, parasitic gaps, multiple realizations and free word order

If we assume that Topic constituents are derived through A'-movement from an IP-internal position, then we must immediately exclude a wh-movement analysis. Indeed, Topic constituents do not show any of the typical Operator-variable effects. Specifically, Topics in Italian do not show Weak Cross Over (WCO) effects and do not license Parasitic Gaps (PG):

(4) a. Gianni, sua_i madre l'_i ha sempre apprezzato.

Gianni his mother him have .3sG always appreciated

the scopal features involved in this operation are left for future work. It is worth mentioning, however, that the interpretation of LD and RD clearly shows that the position of the sentence is crucial in determining the discourse properties of the Topic. Consider, for instance, the following sentence (taken from the LIP corpus of spoken Italian (cf. De Mauro, Tullio et al. ed. 1993, Lessico di Frequenza dell'Italiano Parlato, Etaslibri, Milano):

(i) Nun è questione che il tempo i non te li'ho dato, not be-3sG point that the time not to.you it have-1sG given io te li'ho dato il tempoi.
 I to.you it have-1sG given the time

"The point is not that I didn't give you time, I did give you time."

In the first part of (i) the Topic DP *il tempo* is the point at issue (the *questione*), that is to say, the subject of predication (in Reinhart's 1981 sense). On the other hand, in the second part of the utterance, the RD Topic is only resumed as a sort of 'memo'. As for the IP *te l'ho dato*, it is a broad Focus sentence in the first clause, while there is some contrastive emphasis on the verb *dato* in the second one. It is then plausible to assume that IP-movement targets a projection in the C-system to meet an interface requirement (cf. Chomsky 2002) concerning contrast.

IP-inversion is also supported by interpretative evidence at the PF interface. Consider a Focus construction like (ii) below, in which the postverbal subject *Maria* is focused:

(ii) L' odia MARIA, Gianni.

him hates Maria Gianni

"It is MARIA who hates Gianni."

In a movement analysis which wants a lower TopP to host the RD Topic, *Gianni* should take the main stress (since it is the most embedded constituent). Hence in a sentence like (ii) the Nuclear Stress Rule (Cinque 1993) is violated and an ad hoc stress assignment (Zubizarreta 1998) is invoked for *Maria*. On the other hand, if we assume that *Gianni* is base-generated in the left peripheral TopP, after IP-inversion *Maria* is the most embedded constituent and, as such, it is correctly identified at PF for Focus assignment (while the RD Topic is destressed, as expected). IP-movement can thus provide a connection between syntax and interface interpretation in a straightforward and elegant way.

"(as for) Gianni, his, mother has always appreciated him,."

- b. Sua_i madre l'_i ha sempre apprezzato, Gianni_i.
- (5) a. *Quel libro_i, l_i 'ho cercato [senza trovare e_i] that book it have 1SG looked for without to find
 - b. $*L_i$ 'ho cercato (quel libro_i) [senza trovare e_i] (quel libro_i)

Since a Topic is not an Operator-like constituent (regardless of its location), it cannot be connected with a trace-variable in argument position. Indeed, scholars generally agree that topicalization is *another kind of A'-movement*. However, even assuming this alternative option, we still have problems in defining *what kind* of A'-movement topicalization can be. Major problems for classification derive from data concerning multiple realization and free word order. Consider, for instance, the following contrast between Topics and fronted Foci in Italian:

(6) a. A me, Gianni, di vestiti, in quel negozio non me ne ha to me Gianni, of dresses, in that shop not to-me of-them have 3SG

mai comprati.

never bought.PL

"Gianni has never bought dresses for me in that shop."

- b. Non me ne ha mai comprati, a me, Gianni, di vestiti, in quel negozio.
- (7) *A MARIA IL TUO LIBRO devi dare. to Maria the your book must. 2SG to give

As we can see, there is no limit on the number of Topic constituents that can be realized in Italian (in either periphery) – contrary to fronted Foci – and Topics can be located in any order, independently of their (alleged) original argument position. This represents a serious drawback for a movement analysis (either in syntax or in LF⁴).

Multiple Topics also imply a 'trigger problem' since Move is a costly operation and, from Chomsky (1995) on, it must be triggered by the checking of (semantic or formal) features. However, topicalization in Italian is not apparently connected to a head and, even considering the fact that topicalization is connected with a head in some languages, a feature can be checked only once and in no language are multiple Topics associated with multiple overt heads.

⁴ Indeed, if RD were only the intonational/interpretative consequence of an LF operation (as proposed in Kayne 1994), a fixed order would be expected for dislocated phrases (determined by their being generated as complements), contrary to facts.

Finally, a movement approach must deal with other shortcomings concerning Minimality effects and Locality. Indeed, as Rizzi (2001) points out, "Topics form a separate class from other A'-dependencies", insofar as they do not intervene *in other kind of dependencies of the same type* (i.e., wh-constituents and QPs)⁵ and they do not seem to conform to the locality property of chains, unlike any other kind of A'-movement relation. In a movement analysis, Topic constituents thus assume an 'exceptional character', so that a comprehensive analysis is very far from being reached.

2.3 Clitic doubling and preliminary conclusions

While the derivation of Topic constituents seems to provide different problems for the analysis, their A'-location is not an issue. Clitic resumption provides the clearest argument in favour of this claim.

Clitic pronouns in Italian govern and identify a *pro* in argument position.⁶ Since pronouns are subject to Condition B of Binding, a complement sitting in argument position cannot be doubled by a clitic within the same binding domain. This is typically the case of objects which represent new information, as in (8B):

(8) A: Cosa ha comprato Maria al mercato? "What did Maria buy at the market?"

B: *Maria l_k' ha comprata [una gonna]_k al mercato Maria it have. 3SG bought.F a skirt at.the ma rket

This is clear evidence that (standard) Italian is not a clitic-doubling language. Hence, a Focus or a wh-constituent cannot be doubled by a clitic pronoun either:

(9) a. *Che cosa_k l_k'hai comprato t_k?
what it have.2SGbought

b. *QUESTA GONNA_k l_k 'ho comprata t_k .

this skirt it have. 1SG bought.F

⁵ It is important to notice that this syntactic peculiarity does not distinguish between categories of Topics (i.e., between DPs and PPs). This casts doubts on the possibility of a 'mixed' analysis of topicalization, as proposed in Cecchetto & Chierchia (1999), according to which DPs are moved from an IP-internal position while PPs are base-generated extrasententially.

⁶ Following Uriagereka (1994) and Kayne (1994) we assume "that a clitic does not originate as a pure head that is the complement of the verb, but rather as a subpart of the complement" (Kayne 1994:61). Hence, a clitic pronoun enters the computation as the Determiner of some sister NP, where NP is a *pro*.

Ξ

On the contrary, clitic resumption is a main feature of topicalization:

(10) Questa gonna, *(l)'ha comprata Maria. this skirt it have. 3SG bought.F Maria "This skirt, Maria has bought it."

Topics are thus subject to *different binding conditions* with respect to both DPs located in argument position (where the Topic should be merged, in a movement approach) and constituents moved into an A'-position in the CP system (where the Topic should finally land). This discrepancy is also hard to explain in a movement analysis.

Summing up, the common properties shown by LD and RD provide clear evidence that: a) Topics are located in A'-position; but, b) Topics do not show the effects of A'-movement; and, c) a Topic is not bound to a trace-variable. These common properties cast serious doubts on the feasibility of a movement analysis supporting, instead, a Merge approach. Such an approach will be strongly supported in the following section by the analysis of data concerning some sharp asymmetries in the syntax and interpretation of LD and RD.

3. Reconstruction, Binding and Scope Properties

3.1 LD, RD and reconstruction for binding

Cecchetto (1999) provides sentences like (11) below to show that reconstruction effects with Condition C are attested for LD:

(11) *Il libro di [Leo]_i pro_i l'ha letto volentieri. the book of Leo (he) it have .3sG read with pleasure "Leo_i's book, he_i read it with pleasure."

According to a movement analysis, sentence (11) is ungrammatical because, after reconstruction, the object Topic [*il libro di Leo*] sits in argument position. The R-expression *Leo* is thus c-commanded by the subject *pro* and the relevant binding determines a violation of Condition C, according to which an R-expression must be free in its *Complete Functional Complex* (CFC).⁷

This analysis seems to be supported by the fact that, if we consider (11) in its basic order, the relevant binding is also ungrammatical (cf. (12a)). Moreover, the relevant binding is not allowed for RD either (cf. (12b)):

⁷ The 'Complete Functional Complex' of α is the minimal domain containing α , its governor and an accessible subject/SUBJECT (i.e., a subject DP or a finite Agr head).

- (12) a. *pro $_i$ ha letto volentieri il libro di [Leo] $_i$
 - b. *pro_i l'ha letto volentieri, il libro di [Leo]_i

According to these data, LD and RD seem to share the same binding properties, which apparently depend on LF interpretation after reconstruction.

Let us consider, however, sentences like the following:

(13) a. Il libro che mi ha dato [Leo]_i the book that to.me have.3sg given Leo pro_i *l' ha* scritto da giovane. (he) it have 3sg written when young "The book that Leo; gave me, he; wrote it when he was young." Il libro che ho dato a [Leo]_i h the book that have. 1sg given to Leo pro_i *l' ha* le tto volentieri. (he) it have .3sg read with pleasure "The book that I gave Leo_i, he_i read it with pleasure."

As we can see, in (13a-b) coindexing between the Topic-internal R-expression *Leo* and the pronominal subject obtains grammatical results. However, if we reconstruct the relevant Topics in argument position, the sentences that we obtain are crucially *ungrammatical*:

- (14) a. *proi ha scritto da giovane il libro che mi ha dato [Leo]i.
 - b. *pro_i ha letto volentieri il libro che ho dato a [Leo]_i.

This challenges reconstruction for topicalization. Indeed, if Topics were derived via movement from an IP-internal position and then reconstructed for interpretation, the relevant coindexing should be equally admitted or excluded.

Interestingly, moreover, the relevant binding is also excluded in RD topicalization. So, binding relations for RD Topics seem to conform to basic word order interpretation, but they differ from LD Topics:

- (15) a. *pro_i l'ha scritto da giovane, il libro che mi ha dato [Leo]_i.
 - b. *pro_i l'ha letto volentieri, il libro che ho dato a [Leo]_i.

⁸ As is well known, judgements can differ according to speakers' variety, so that for some speakers the sentences in (13) may sound slightly marginal. What is crucial, however, is the contrast between (11) and (13) that – as confirmed by several informants– cannot be d enied.

This asymmetry can be hardly explained in a movement analysis since, in that case, reconstruction effects should equally apply for LD and RD.

On the other hand, if we assume a Merge analysis, then Principle C is no longer an issue in the relevant sentences because the DP *Leo* will never be in the condition to be c-commanded by the subject *pro*. On the contrary, c-commanding conditions *on* the subject *pro* must be considered. In this perspective, we can see that what makes Topic binding different in (11) with respect to (13) is the CFC to which the R-expression *Leo* belongs. Compare the internal structures of the relevant Topic DPs:

In (16) the Topic DP [*il libro di Leo*] does not contain a(n accessible) subject/SUBJECT, hence it cannot qualify as a CFC for the R-expression *Leo* that it contains. This means that the DP *Leo* belongs to the CFC defined by the sentential subject *pro*. On the contrary, in (17) the R-expression *Leo* is part of the Topic-internal CFC defined by the relative clause [*che ho dato a Leo*]. Hence, it belongs to a different binding domain with respect to the matrix pronominal subject. We can thus conclude that sentences like (16) are ruled out by Principle B, because the pronominal subject is coindexed with an R-expression within its CFC, while sentences like (17) are grammatical because no binding violation arises.

the file $\,$ that Leo have.3SG prepared, (he) it have.3SG put on the our desk $\,$ b. $\,$ *L'affermazione [che Leo_i è $\,$ un ladro] $\,$ pro_i $\,$ l'ha $\,$ contestata con forza.

the claim that Leo be.3sG a thief (he) it have.3sG contested forcefully The argument/adjunct asymmetry, however, suffers from several shortcomings. As convincingly argued in Lasnik (1998), Lebeaux's effects are rather spurious and most of his examples can be plausibly ruled out on independent grounds, namely, because of their pragmatic oddity; once the pragmatic factor is controlled, many examples become totally fine.

-

⁹ The CFC of the object *pro* is also defined by the IP node (which contains in its Spec an 'accessible subject') but it excludes the Topic. Their coindexing is thus allowed.

¹⁰ An anonymous reviewer points out that the difference between (11) and (13) is also compatible with a Move analysis, if we assume (following Lebeaux 1989) a 'late adjunction' for the relative clause which, combining with the antecedent after Move, will not be lowered for reconstruction. Indeed, Cecchetto & Chierchia (1999) use the argument/adjunct asymmetry to prove antireconstruction effects through examples like the following:

⁽i) a. La scheda [che Leo_i ha preparato] pro_i l'ha messa sulla nostra scrivania.

It could be objected that the pronominal subject is not c-commanded by the R-expression *Leo* in either construction, so that in a sentence like (16) no violation of Condition B can be at stake. This objection, however, is not germane if we assume, following Kayne (1994:26-27), that constituents located in the highest Specifier within a DP can c-command out of it. So, since possessor phrases move into Spec,DP at LF, then the DP *Leo* can c-command out of the Topic DP [*il libro di Leo*] and bind the subject *pro* (*within its CFC*), determining a violation of condition B. 11,12

On the other hand, it is worth noticing – in favour of a Merge analysis – that in a sentence like (ia) the relevant coindexing is excluded if we add a possessor phrase:

(ii) *La scheda di Luigi [che Leo_i ha preparato] pro_i l'ha messa sulla nostra scrivania.

the file of Luigi that Leo have. 3sg prepared, (he) it have.3sg put on the our desk This asymmetry is immediately caught by Kayne's (1994) suggestion about a higher Operator position for subject constituents in the DP: in (ii) such a position is reached at LF by the possessor phrase *di Luigi*, blocking the raising of *Leo* and excluding the relevant coreference. ¹¹ This explanation accounts for the ungrammaticality of English sentences like (i) and for the anaphoric binding in Italian sentences like (ii), without depending on reconstruction effects (cf. Kayne 1994). In Hinterhölzl (2002) movement of the possessor phrase is also argued to be responsible for the inverse reading in English sentences like (iii) and, interestingly, the same type of inverse reading is available in Italian (cf. *qualche abitante di ogni città*, *odia il suo*;

(i) *John;'s book, he; does not read it.

traffico):

- (ii) il suo_i libro, pro_i lo legge sempre.the his book (he) it read.3sG always"His book, he always read it."
- (iii) Some inhabitant of every city, hates its, traffic.
- ¹² An anonymous reviewer points out that other sentences similar to (11) except for the fact that the possessor phrase is replaced by another non-clausal constituent are as degraded as (11). Consider, for instance:
- (i) *Il libro per Leo_i pro_i l'ha letto volentieri. the book for Leo (he) it have 3SG read with pleasure

Given the evidence provided up to now, however, we think that our argument cannot be challenged by sentences like (i), for which another explanation must be found, independent of reconstruction and possessor phrase raising. Specifically, it is plausible to assume that *Leo* in (i) is not a sufficiently prominent referent for *pro* (cf., in this respect, Calabrese 1986). Evidence supporting this argument comes from the fact that, if we substitute *pro* with an open pronoun both in (11) and in (i), sentence (11) stays ungrammatical (cf. (iia)) while coreference in (iib) is acceptable (we thank Valentina Bianchi for this suggestion):

- (ii) a. *Il ritratto di Leo; lui; l'ha visto ieri.
 the picture of Leo he it have3sG seen yesterday
 - b. ²Il ritratto <u>per Leo_i lui</u> l'ha visto ieri. the picture for Leo he it have3sG seen yesterday

As for RD topicalization, the present analysis can explain its binding properties in the light of IP-movement, which *inverts scope properties*. So, after movement, the IP c-commands the TopP projection and, crucially, the subject DP occupies the highest Spec position within the relevant IP. Hence, it can ccommand out of it. For this reason, in (15a-b) the Topic-internal DP Leo is always bound by the pronominal subject, causing a violation of Condition C:

(18)*
$$[_{GP}[_{IP}pro_k l'ha scritto pro_j da giovane] [_{TopP}[_{DP} il libro che mi ha dato [_{Leo}]_k]_i t_{IP}]^{13}$$
 (=15a)

In the next section additional evidence for a Merge analysis will be provided from the interaction between scope and binding properties.

3.2 LD, RD and scope properties

As claimed in Bobaljik (2002), LF is 'coherent', that is to say, a given element may not take scope from one position and yet be interpreted in another position for the purposes of Binding Conditions. Hence, in a movement

Sentence (ia), in which the extraposed IO al suo autore is not clitic-resumed, is ambiguous since it admits both a reading in which all the books of a given list will be given to a single author and a distributive reading. This is evidence that the non clitic-resumed IO is 'marginalized' (cf. Frascarelli 2000, Cardinaletti 2002) and, as such, reconstructed within IP. Hence, either scope relation is possible between the QP ogni and the possessive suo.

On the other hand, the distributive interpretation is excluded in sentence (ib), showing that the clitic-resumed Topic has scope over the DO, but not vice-versa and providing strong support to the present Merge analysis.

Along the same lines, it is feasible to assume that subject dislocation is derived by movement (since subject clitics are not available in standard Italian) and, for this reason, reconstructed at LF. So, in a sentence like (iia) the subject position is not occupied by a ccommanding pro (like in (18)), but by a trace (we thank Petra Sleeman for having pointed this fact out to us):

 $[GP]_{IP} t_k$ è molto interessante [questo libro] $[E]_{IP} t_k$

¹³ It is important to underline that the Merge analysis supported here only refers to cliticresumed Topics. We therefore do not exclude that other forms of 'dislocation' might be derived through movement (and recent analyses have shown that such a distinction should be made: see Cardinaletti 2002, Hinterhölzl & Pili 2002), Indeed, contrasts like the following support the necessity of such a distinction:

Darò ogni libro, al (i) a. suo autore. give.FUT.1SG every book to.the his author

h. Glidarò ogni libro, al suo autore. to.him give.FUT.1SG every book to.the his author

⁽ii) a. E'molto interessante, questo libro. be.3sg very interesting this

analysis, Topic scope properties must also depend on reconstruction. As a consequence, the scope properties of a QP should not be affected by topicalization, because QPs always assume scope over the sentence and Topics are always reconstructed in argument position. Let us then consider a sentence like (19) below:

(19) Maria_i ha presentato [ad ogni ragazzo]_k [il suo_{i/k} professore] Maria have.3SG introduced to every boy her/his teacher "Maria has introduced every boy to her/his teacher."

In (19) both the QP *ad ogni ragazzo* and the object DP *il suo professore* sit in argument position (assuming a basic IO-DO order; cf. Larson 1988). In this basic order, the possessive *suo* can be bound both by the subject DP *Maria* and by the QP *ad ogni ragazzo*. This is expected since both constituents c-command the object DP, hence they both qualify as possible binders¹⁴. Let us now analyse the interpretation of the possessive *suo* after the topicalization of the object DP:

- (20) a. Maria_i, [il suo_{i/*k} professore], l'ha presentato [ad ogni ragazzo]_k
 Maria her/*his teacher him have.3SG introduced to every boy
 - b. *Maria*_i *l'ha presentato* [ad ogni ragazzo]_k, [il suo_{i/?*k} professore]¹⁵
 Maria him have.3SG introduced to every boy her/^{**}his teacher

So, binding and scope properties are completely different from those examined in the basic structure (19) and, what is even more interesting, LD and RD also differ from each other. In particular, in the LD structure the object Topic can only be bound by the DP *Maria*, while in the RD construction the subject DP *Maria* is the preferred binder, even if the QP shows a (marginal) possibility of coreference. This is strong evidence that a Topic is *not* reconstructed for interpretation since, in that case, scope properties should always be reduced to the unmarked sentence reading, contrary to facts.

Scope properties of Topics, however, are immediately explained in the light of a Merge analysis. Consider the following structures:

(20a')[TopP Maria_i [TopP il suo_i/*k professore [IP l'ha presentato [ad ogni ragazzo]k]]]

¹⁴ A binder in the extralinguistic context is, of course, also possible.

¹⁵ As is well known, an Operator can only be the antecedent of an anaphor it c-commands. Hence, a QP must have scope over pronouns to bind them, whereas this is not necessary for non quantified DPs (cf.: *Maria_k l'ha reso a Gianni_j, il suo_{k/j} libro*. "Maria has given it back to Gianni, her/his book.").

(20b')[GP[P Maria; l'ha presentato [ad ogni ragazzo]k] [TopPil suoi/?*k professore] tp]

As we can see, in (20a') the object Topic *il suo professore* is only c-commanded by another Topic, namely the DP *Maria*. In no point of the derivation can the QP assume scope over these two constituents, hence it cannot qualify as a possible binder for the possessive *suo*. On the contrary in the RD construction, after IP-inversion, the Topic is c-commanded by the whole sentence, which contains both the subject *Maria* and the QP. However, since the DP *Maria* is the highest nominal constituent within the sentence, it qualifies as the closest potential binder for *suo* and induces the preferred reading, while the QP only maintains a marginal possibility of binding the possessive.

These data provide strong empirical support in favour of the present analysis. They exclude that a RD Topic is simply moved into a lower TopP projection because, in that case, the QP in (20b) would qualify as the closest potential binder for the object Topic, contrary to facts. The evidence provided also excludes a PF account for RD topicalization since, in that case, scope and binding conditions should be kept constant with respect to basic order.

4. Pronominal Variables, Local Merge and TopP-to-TopP Movement 4.1 The proposal (II)

The data presented up to now strongly support a Merge analysis of topicalization. This means that Topics do not head a movement chain, but a *binding chain* and, considering the inventory of empty categories, the tail of this chain can only be a *pro*. However, since the Topic is merged in extrasentential position, this is a particular kind of *pro*, namely a *pronominal variable* (an A'-bound *pro*). As extensively discussed in Rizzi (1994), an A'-bound *pro* is a *pro* which acts as a variable ¹⁶ and, for this reason, it must meet both *pro* licensing conditions (i.e., government by a lexical head) and the conditions imposed on variables (i.e., it must have a *local A'-binder* to be identified). ¹⁷

¹⁶ A pro in Italian can be a variable (bound by a DP or a QP), unlike pronouns or full DPs:

 $⁽i) \quad [\textit{Mario}]_k \, / \, [\textit{ogni} \, uomo]_k \, \, \textit{pensa} \quad \, \textit{che} \quad pro_k / * \textit{lui}_k / * \textit{Mario}_k \, \, \textit{vincer\`a}.$

Mario every man think.3sG that pro / *he / *Mario win.Fut.3sG

 $^{^{17}}$ A semantic analysis of Topic is far beyond the purposes of the present paper. In this respect, an interesting discussion is provided in Delfitto's (2002) work on clitic-constructions, according to which clitic pronouns are bound variables and the Topic is "simply the argument of the λ -abstract encoded by means of cliticization". Hence, "there is no syntactic dependency between the dislocated Topic and the argument position to which the Topic is related" (p. 54).

Though we share with Delfitto the basic assumption that a Topic is bound to a pronominal variable, we cannot conform with his 'extra-syntactic' analysis of Topic

The presence of a pronominal variable in argument position is crucial to understand the syntax and interpretation of Topic constructions. In particular, the 'local A'-binder' requirement crucially implies that a Topic must be merged in the *TopP projection which locally c-commands* the sentence containing the coindexed *pro*. In other words, it means that Topic Merge *depends on the position of the pronominal variable* the Topic is coindexed with.

Local Merge is forced by a *double formal requirement*: on the one hand, the Topic provides a reference for (i.e., it *identifies*) the *pro* and, on the other, the Topic – an XP in non-argument position – is *licensed* by the argument pronoun through the binding chain (in the spirit of Baker 1996). ^{18, 19}

So, when a Topic is merged in a simple sentence, Merge location is straightforward: the Topic is located in the only TopP projection available:

On the other hand, when the DP containing the pronominal variable is in a subordinate clause, then the Topic is *also* generated in the subordinate TopP:

constituents. Indeed, syntactic analysis – the basic concern of this paper – provides substantial evidence that topicalization is not a purely interpretative fact. Scope, binding and phonological data show that Topic constituents cannot be simply 'attached' in the left periphery and that a specific maximal projection – having non-quantificational A'-properties (cf. Lasnik & Stowell 1991 and the data in section 3.2) - must be assumed to host these constituents. It is otherwise very hard to understand, among other things, the sharp asymmetries between LD and RD and the fixed order between Topics and Operators in general (cf. Rizzi 1997, Frascarelli 2000). Also, a free generation for Topics could not explain why contrastive Topics (Büring 1999) can only be found in the left periphery of the sentence, while the right periphery is only accessible to familiar, destressed elements. Thus, according to the present analysis, Topics occupy the Spec position of a recursive TopP projection and they define a kind of "high predication" within the C-system (as argued in Rizzi 1997). The possibility of TopP recursion can be attributed (as suggested in Delfitto 2002) to the fact that Topics are not connected with [-interpretable] features that trigger Move for checking. Merge is therefore the most economical and comprehensive way to account for the syntax of Topic constructions.

 $^{^{18}}$ The argument pro 'transfers' Case assignment to the Topic in virtue of the chain. This accounts for Topic Case marking in languages in which morphological Case is available.

¹⁹ Of course, not only does a Topic provide referential features to the argument *pro*: its discourse role with respect to the rest of the sentence is a basic element of information structure. In particular, we can distinguish (at least) three types of Topic, namely Aboutness (Reinhart 1981), Familiarity (Pesetsky 1987) and Contrast (Büring 1999) Topics. The interplay between syntax and information structure is the subject of ongoing research.

(22) $[_{TopP}[_{IP}ti\ ho\ già\ detto\ [_{CP}che\ [_{TopP}G._k[_{IP}lo\ vedrò\ [_{DP}[_{NP}\ pro_k]]\ domani\]]]]]$ "I already told you that (as for) Giovanni, I will see him tomorrow."

This analysis makes Topic interpretation *completely conform* to the general properties ruling over the formation of chains, differently from what is generally assumed in a movement approach.

Local Merge is also supported by immediate empirical facts. As we know, a Topic can be found in the left periphery of a subordinate clause (cf. (22) above), or it can be found in the matrix TopP:

(23) $[T_{OPP}G_{.k}]_P ti \ ho \ già \ detto [C_P che[T_{OPP}t_k]_P lo \ vedrò \ [D_P[N_P Pro_k]] \ domani]]]]]$ "(as for) Giovanni, I already told you that I will see him tomorrow."

If we refute a local Merge analysis and assume that Topics are always inserted in the highest TopP projection of a complex sentence, what operation should we suppose to account for its presence in a subordinate TopP? Since lowering operations are excluded (for independent reasons), we should admit an optional location for Topic Merge in complex sentences. However, to posit optionality as a property is not an achievement for the theory. On the other hand, if we assume that a Topic is always merged locally with respect to its tail, then non-local positions are simply the effect of 'TopP-to-TopP movement' (i.e., movement from local TopP into a higher TopP projection). ²⁰

²⁰ Local Merge and TopP-to-TopP movement also provide a straightforward explanation for data concerning complex Topic constructions. Consider (i) and the two Topic structures in (ii):

⁽i) [IP non credo [CP che Gianni sia convinto [CP di conoscere Maria]]]
not think.1SG that Gianni be.SUBJ.3SG convinced of to know Mary
"I don't think that Gianni is sure that he really knows Mary."

⁽ii) a. Maria; non credo che, [di conoscerla; t;] Gianni ne sia convinto.

b. *[di conoscerla; t;] non credo che, Maria; Gianni ne sia convinto.

Cecchetto (1999) attributes the ungrammaticality of (iib) to a "massive violation of the Proper Binding Condition" (PBC) since (in a movement analysis) *Maria* does not properly c-commands its trace. The validity of such a claim, however, is immediately refuted by the grammaticality of examples like (iii) below, which should also imply a PBC violation:

⁽iii) [di conoscerla_i t_i] Maria_i non credo che Gianni ne sia convinto.

On the other hand, the contrast between (iia) and (iib) can be easily explained through Locality of Merge, which wants the two relevant Topics to be inserted as shown in (iv):

⁽iv) $[P_{\mathbb{P}} non\ credo\ [CP\ che\ [TopP\ Maria_i\]\ di\ conoscerla_i\]\ [P_{\mathbb{P}} Gianni\ ne\ sia\ convinto\]]]$

In the following section we will concentrate on the distinction between local and non-local Topics, showing that it can account for additional data (and asymmetries) concerning LD and RD Topic constituents.

5. Topics and Minimality

Rizzi (2001) argues that Topic constituents in Italian never interfere with other types of A'-moved elements, and that only a 'mild degradation' is produced by the interaction with (argument and adjunct) wh-movement:

- (24) a. Non so come_j pensi che [$_{TopP}$ a G. [$_{FocP}$ t_j [$_{IP}$ gli dovremmo parlare]]] "I don't know how you think that, to Gianni, we should talk to him."
 - b. 'Non so a chi_j pensi che[TopPqueste cose[FocPt_j [Iple dovremmo dire]]] "I don't know to whom you think that, these things, we should say them."

Minimality, however, does not produce 'mild degradations', so that the effect noticed in (24a-b) is hardly explained in a movement analysis in which Topics are generated in argument position.

On the contrary, the lack of Minimality effects is fully expected in a Merge analysis, since binding chains are not affected by them and the relevant degradation can be attributed to a weak subjacency effect, determined by the movement of the wh-constituent across a non L-related (and lexically filled) projection (Chomsky 1995:64, 196) as the TopP projection is, in this case.

However, since we have assumed the possibility of TopP-to-TopP movement, then we expect an asymmetry between local and non-local (i.e., Topics moved from their Merge position) Topics. That is to say, while local Topics are not affected by Minimality (heading a binding chain), we expect *non-local* Topics to show Minimality effects with A'-constituents *of the same type*. ²¹ Let us then consider the relevant data.

5.1 Non-local Topics, Wh-constituents and QPs

The analysis of data show that non-local Topics do not interfere with the interpretation of Operator-like constituents, such as wh-elements and QPs,

Then, starting from (iv), we obtain (iia) moving *Maria* to matrix TopP, while to derive (iib) the CP [di conoscerla] should be moved leaving *Maria* in situ, but this is not possible, since *Maria* is part of the relevant Topic CP. The derivation of (iii), finally, implies movement of *Maria* and remnant movement (an operation which is exempted from the PBC) of the Topic CP.

²¹ Indeed, Minimality effects have only been tested for *local* Topics in the relevant literature.

providing additional (and conclusive) evidence that Topics and Operators are different types of A' constituents:

- (25) Queste $cose_k$, non so a chi pensi che t_k le dovremmo dire pro_k .
 - "These things, I don't know to whom you think that we should say them."
- $(25')[_{TopP}\ queste\ cose_k\ [_{IP}\ non\ so\ [_{FocP}\ a\ chi_j\ [_{IP}\ pensi\ [_{CP}\ che\ [_{TopP}\ t_k\ [_{FocP}\ t'_j\ [_{IP}\ le\ dovremmo\ dire\ t_j\ [_{DP}t_{cl}\ [_{NP}\ pro_k\]]]]]]]]]$
- (26) Sono sicura che Mario ha detto questa storia a qualcuno. be.1SG sure that Mario have.3SG told this story to someone "I am sure that Mario told someone this story."
- (26')a. A qualcuno, questa storiak, sono sicura che tk M. l'ha detta prok.
 - b. Questa storia_k, a qualcuno sono sicura che t_k M. l'ha detta pro_k.
 - c. Sono sicura che M. l'ha detta, questa storia, a qualcuno $t_{\mathbb{P}}$.
 - d. Sono sicura che M. l'ha detta a qualcuno, questa storia t_{IP}.

5.2 Local and non-local multiple Topics

Crucial evidence for a definition of the syntax and interpretation of Topic constituents emerges, on the other hand, from their interaction with other Topics.

It is commonly agreed that multiple Topics are always allowed in Italian, regardless of linear order. The following data will show that this claim needs a refinement, that is to say: multiple *local* Topics *are always allowed*, while *non-local* Topics *are affected by Minimality*. Consider local Topics first:

- (27) a. *M.i*, *questa storia*_j, *a Luigi*_k pro_i *gliel'ha detta* [PP[DPPro_j]] [DPPro_j] Mario this story to Luigi to.him-it have3SG told
 - b. $A\ Luigi_k,\ M._i,\ questa\ storia_j\ pro_i\ gliel'ha\ detta[pp[Dppro_k]]\ [Dppro_j]$
 - c. $\mathit{Questa\ storia_j},\ a\ \mathit{Luigi_k},\ M_i,\ \mathsf{pro_i}\ \mathit{gliel'ha\ detta}\ [_{PP}[_{DP}\ \mathsf{pro_k}]]\ [_{DP}\mathsf{pro_j}]$

Multiple Topics are thus always allowed in a simple sentence, regardless of linear order and of the syntactic role of the pro they are coindexed with. ²²

Consider now topicalization in a *complex* sentence, like (28) below, in which an IO Topic is merged into the matrix TopP, while a DO Topic is located in the embedded TopP position:

²² This means that Italian Topics do not show *superiority effects*, which are present, on the contrary, in wh-questions, providing additional evidence against Move from an argument position:

⁽i) a. Chi non sai a chi ha parlato? who not know.2sg to whom have.3sg spoken

b. *A chi non sai chi ha parlato?

(28) A Luigi _k gli_k ho detto che Mario _j lo_j vedrò domani. to Luigi to.him have.1SG told that Mario him meet.FUT.1SG tomorrow

As (28) shows, double topicalization does not affect interpretation. This is expected by the present analysis since the two Topics are merged locally with respect to their coindexed *pros*, they form binding chains and, consequently, they cannot determine Minimality effects on each other:

$$(28')[_{\text{TopP}}a\ L._k\ [_{\text{IP}}gli\ ho\ detto\ [_{\text{PP}}[_{\text{DP}}\ pro_k]]]_{\text{CP}}che[_{\text{TopP}}M._j\ [_{\text{IP}}lo\ vedr\grave{o}\ [_{\text{DP}}\ pro_j]]]]]]$$

Let us consider now the effects of TopP-to-TopP movement, by raising the lower DO Topic into the matrix TopP (while *a Luigi* sits in its Merge position):

- (29) $^{?}Mario_{j}$, a $Luigi_{k}$ gli_{k} ho detto che t_{j} lo_{j} vedrò domani. (29') $[TopPM._{j}[TopPA L._{k}[PBli..detto[PP[DPPTO_{k}]]][CPChe[TopPt_{j}[Plo vedrò[DPPTO_{j}]]]]]]]$
- (29) show that after TopP-to-TopP movement, the relevant Topic construction is still grammatical (though marginal, for some speakers). This is also expected, because the (nested) chains formed by the two Topics are differently constituted: the highest Topic is connected with a trace, while the lowest sits in its Merge position and binds a *pro*. Hence, each Topic is the closest potential A'-binder for the empty category it is connected with and no Minimality effect can arise.

Let us finally consider the movement of the matrix Topic (i.e., *a Luigi*) into a higher TopP position:

$$(30)*A\ \textit{Luigi}_k,\ \textit{Mario}_j,\ t_k\ \textit{gli}_k\ \textit{ho\ detto\ che}\ t_j\ lo_j\ \textit{vedr\`o}\ \textit{domani}.$$

$$(30')*[{}_{\text{TopP}}\textit{A}_j\ [{}_{\text{TopP}}\textit{M}_j\ [{}_{\text{TopP}}\textit{t}_k[{}_{\text{IP}}..[{}_{\text{PP}}[{}_{\text{DP}}\textit{pro}_k]][{}_{\text{CP}}\textit{che}[{}_{\text{TopP}}\textit{t}_j[{}_{\text{IP}}\textit{lo\ vedr\'o}[{}_{\text{DP}}\textit{pro}_j]]]]]]]]]$$

The ungrammaticality of (30) is fully consistent with our hypothesis. What makes the difference between this construction and the (grammatical) topicalizations in (28') and (29') is that only in (30) do both Topics head a movement chain and, crucially, these chains intersect. In this configuration, therefore, the Topic moved from the subordinate clause, *Mario*, qualifies as the most local A'-binder for the trace left by *a Luigi* in the matrix TopP.

Minimality effects for non-local Topics are also found in RD structures:

- (31) a. A Luigi_k, gli_k ho detto che lo_j vedrò, Mario_j.
 - b. Gli_k ho detto a Luigi_k, che lo_j vedrò, Mario_j.
 - c. ??Gli_k ho detto che lo_i vedrò, Mario_i, a Luigi_k.
 - d. *Gli_k ho detto che lo_i vedrò, a Luigi_k, Mario_i.

Sentences (31a) and (31b) are fully grammatical. This is to be expected since both Topics are merged in their local TopP projections and, after IP-inversion, each RD Topic c-commands the relevant IP-trace. Hence, each Topic qualifies as the closest A'-binder for its pronominal tail. The marginality of sentence (31c), on the other hand, can be profitably attributed to the presence of nested IP-inversions. Minimality effects are, finally, clearly present in sentence (31d), as expected. Indeed, according to the present analysis, the relevant RD construction can only be derived from the (ungrammatical) LD construction shown in (30'). In other words, to obtain the [a Luigi, Mario] final order, we must first raise the embedded DO Topic, then move a Luigi across it and, finally, apply IP-inversion. This final movement operation can, of course, only increase the ungrammaticality of (30).

We can thus conclude that non-local Topics – unlike local ones – determine Minimality effects on each other's traces. Since Topic linear order is never an issue in simple sentences, these data provide conclusive evidence that local Topics are not subject to any kind of movement operation, while non-local Topics are A'-moved constituents. Since Minimality effects only arise between constituents of the *same type*, it is fundamental to notice that non-local Topics only show Minimality effects with other Topics, that is to say, with constituents merged in *non-argument position*.

We consider the data analysed in this paper as sound evidence against a theory of argument generation for (clitic-resumed) Topic constituents. A Merge analysis can indeed provide a uniform account for both LD and RD properties, without ad hoc stipulations. As for IP-inversion, we can tentatively assume that it is designed to check a scopal feature in the C-system, in line with Chomsky's (2002) suggestion that dislocation may be implemented by the computational system to satisfy some specific interface conditions.

References

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Belletti, Adriana. 2001. "Aspects of the Low IP Area". *The Structure of IP and CP. The Cartography of Syntactic Structure* 2 ed. by Luigi Rizzi. Oxford: Oxford University Press.

- Benincà, Paola. 2001. "The Position of Topic and Focus in the Left Periphery". *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi* ed. by Guglielmo Cinque & Giampaolo Salvi, 39-64. Amsterdam & New York: Elsevier.
- Bobaljik, Jonathan David. 2002. "A-Chains at the PF-Interface: copies and 'covert' movement". *Natural Language and Linguistic Theory* 20.197-267.
- Browning, Marguerite A. 1996. "CP Recursion and *that-t* Effects". *Linguistic Inquiry* 27:2.237-255.
- Büring, Daniel. 1999. "Topic". Focus Linguistic Cognitive and Computational Perspectives ed. by Peter Bosch & Rob van der Sandt, 142-165. Cambridge: Cambridge University Press.
- Calabrese, Andrea. 1986. "Some Properties of the Italian Pronominal System". Papers in Theoretical Linguistics ed. by Naoki Fukui, Tova Rapoport & Elizabeth Sagey. MIT Working Papers in Linguistics, vol. 8, MIT Press.
- Cardinaletti, Anna. 2002. "Against Optional and Zero Clitics. Right Dislocation vs. Marginalization". *Studia Linguistica* 56:1.29-57.
- Cecchetto, Carlo. 1999. "A Comparative Analysis of Left and Right Dislocation in Romance". *Studia Linguistica* 53.40-67.
- & Gennaro Chierchia. 1999. "Reconstruction in Dislocation Constructions and the Syntax/Semantics Interface". WCCFL XVII Proceedings, ed. Susan Blake, Eun-Sook Kim & Kimary Shahin, 132-146. Stanford University: CSLI publications.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- —— 2002. "An Interview on Minimalism". *On Nature and Language* ed. by Adriana Belletti & Luigi Rizzi, 92-161. Cambridge: Cambridge University Press.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.
- Culicover, Peter W. 1996. "On Distinguishing A'-Movements". *Linguistic Inquiry* 27:3.445-463.
- Delfitto, Denis. 2002. "On the Semantics of Pronominal Clitics and some of its Consequences". *Catalan Journal of Linguistics* 1.41-69.
- Frascarelli, Mara. 2000. The Syntax-Phonology Interface in Focus and Topic Constructions in Italian. Dordrecht: Kluwer Academic Publishers.

- Hinterhölzl, Roland. 2002. "Parametric Variation and Scrambling in English". *Studies in Comparative Germanic Syntax* ed. by Jan-Wouter Zwart & Werner Abraham, 131-150. Amsterdam & Philadelphia: John Benjamins.
- —— & Diana Pili. 2002. "Argument shift across language types". Talk presented at *XXVIII Incontro di Grammatica Generativa* (28 February 2 March 2002, Lecce).
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Larson, Richard. 1988. "On the Double Object Construction". *Linguistic Inquiry* 19.335-391.
- Lasnik, Howard. 1998. "Some Reconstruction Riddles". *University of Pennsylvania Working Papers in Linguistics*, *Vol 5.1* ed. by Alexis Dimitriadis, Hikyoung Lee, Christine Moisset & Alexander Williams, 83-98.
- —— & Tim Stowell. 1991. "Weakest Crossover". *Linguistic Inquiry* 22.687-720.
- Lebeaux, David. 1989. *Language Acquisition and the Form of Grammar*. Ph.D. Dissertation, University of Massachusetts, Amherst.
- Pesetsky, David. 1987. "Wh-in-Situ: Movement and Unselective Binding". *The Representation of (In)definiteness* ed. by Eric Reuland & Alice ter Meulen, 98-129. Cambridge, Mass.: MIT Press.
- Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics". *Philosophica* 27.53-94.
- Rizzi, Luigi. 1994. "Early Null Subjects and Root Null Subjects". *Language Acquisition Studies in Generative Grammar* ed. by Teun Hoekstra & Bonnie Schwartz. Amsterdam & Philadelphia: John Benjamins.
- ——— 2001. "Locality and Left Periphery". Ms., University of Siena.
- Uriagereka, Juan. 1994. "Aspects of the Syntax of Clitic Placement in Western Romance". *Linguistic Inquiry* 26.79-123.
- Vallduví, Enric. 1990. *The Informational Component*. Ph.D. Dissertation, University of Pennsylvania.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge, Mass.: MIT Press.

FOCUS ON NEGATIVE CONCORD

DANIELA ISAC *UOAM/ Concordia*

1. The Problem

Negative concord languages express negation more than once in a clause, but negation is nevertheless interpreted only once. Typically, negative concord (NC) is manifested when a clause contains the sentential negative marker and one or several instances of the so called N-words (Laka 1990). Another variety of NC involves the co-occurrence of two (or possibly more) N-words without a negative marker ('negative spread', den Besten 1986). A proper characterization of NC thus involves the need for a clear definition of N-words. Languages like Italian, Spanish, European (E) Portuguese and Romanian pose the following problem for the attempt to offer a uniform characterization of N-words: in Italian, Spanish, and E Portuguese, N-words can co-occur with the negative marker if they are postverbal, but not if they are preverbal, as in (1):

(1)	a.	Maria*(non) ha visto nessuno.	(I)
		Maria Neg has seen N-word	
		*(<u>No</u>) conozco <u>a nadie</u> .	(S)
		Neg know to N-word	
		O Pedro <u>não</u> viu ninguém	(EP)
		Pedro <u>Neg</u> saw <u>N-word</u>	
	b.	<u>Nessuno</u> * <u>non</u> ha visto Maria.	(I)
		N-word neg has seen Maria	
		<u>Nadie</u> * <u>no</u> ha visto a Maria.	(S)
		N-word neg has seen Maria	
		<u>Ninguém*não</u> viu o Pedro.	(EP)
		N-word neg saw Pedro	

In contrast, in Romanian, N-words co-occur with the negative marker irrespectively of the postverbal or preverbal position, as illustrated in (2).

- (2) a. <u>N-</u> a venit <u>nimeni</u>. <u>Neg</u> has come <u>N-word</u> "Nobody has come."
 - b. <u>Nimeni</u> <u>n</u>- a venit. <u>N-word</u> <u>Neg</u> has come "Nobody has come."

In view of (1)-(2), one might conclude that N-words in Romanian have a consistent behavior, since they always co-occur with a negative marker, whereas in Italian, Spanish and E Portuguese, N-words are lexically ambiguous: they are negative quantifiers when in preverbal position, since they have enough inherent negative force to make the whole sentence negative, without the 'help' of a negative marker, as in (1b), and they are Negative Polarity Items when in postverbal position, since in postverbal positions they require the presence of a negative marker, as in (1a) (see Rizzi 1982, van der Wouden and Zwarts 1993, Dowty 1994).

In this paper we propose that the lexical ambiguity of N-words in (1) is only apparent and that N-words in these languages are unambiguously indefinites or pure variables. The quantificational feature of preverbal N-words is not to be identified with the negative feature within the N-word, but is instead the result of the interaction of the N-word with Focus. More precisely, we propose that preverbal N-words are in a Focus position. Since Focus is quantificational, an N-word in Focus acquires quantificational features. The differences between Italian, Spanish, and E Portuguese, on the one hand, and Romanian, on the other hand, are accounted for by the particular way in which [neg] features cluster together with [foc] features in these languages. In Spanish, Italian, and E Portuguese, both [neg] and [foc] are features of the same functional head, whereas in Romanian they are features of independent functional heads. This creates a competition between Merge and Move in the case of Spanish, Italian, and E Portuguese, but not in the case of Romanian.

The paper is organized as follows: in section 2, we consider previous accounts of N-words and point out their strengths and weaknesses; in section 3, we formulate our proposal and we present arguments for the feature content of N-words, and arguments for a focus analysis of preposed N-words; in section 4, we propose an analysis for the asymmetry between Italian, Spanish and E Portuguese, on the one hand, and Romanian, on the other hand; and finally, in section 5, we present the conclusions.

2. N-words. Previous Accounts

As mentioned in the introduction, a clear definition of N-words is crucial for a proper characterization of NC. There has been a lot of debate in the literature on whether N-words are negative quantifiers or NPIs. (Laka 1990, Zanuttini 1991, Haegeman 1995, among others). In this section we review these positions and we also discuss a third, intermediary, position, which holds that N-words are ambiguous between an interpretation as negative quantifiers or as NPIs.

2.1 N-words are quantificational

This view analyses N-words as inherently negative elements and as quantificational. N-words thus have two crucial features: [neg] and [qu]. Not all proponents of this view agree as to the exact nature of the quantificational feature of N-words; some hold that it is instantiated by the existential quantifier, while others hold that it is expressed by a universal quantifier. We will ignore this distinction in this paper, since it is irrelevant for what we have to say.

The challenge that this type of analysis must face is the fact that a sentence containing several N-words becomes semantically non-compositional. If each N-word has a negative quantificational feature, then the presence of, say, two such N-words in one sentence, or of one N-word and a negative marker, should result in an affirmative interpretation. However, this is not the interpretation that one gets for a sentence like (3a,b):

- (3) a. <u>Nessuno</u> ha letto <u>niente</u>. (I) nobody has read nothing "Nobody has read anything."
 - b. Non ha letto niente.

 neg has read nothing

 "(S)he hasn't read anything."

One way to deal with sentences like (3b) is to propose the generalization in (4), by analogy with the Wh-criterion proposed by Rizzi (1990):

(4) The NEG criterion (Haegeman & Zanuttini 1991; Zanuttini 1991)

- a. A NEG-operator must be in Spec-head agreement relation with an $X^{\circ}[NEG]$.
- b. An X°[NEG] must be in Spec-head agreement relation with a NEG-operator.

122 DANIELA ISAC

The Neg criterion requires that every element bearing negative features form a constituent with a negative head at some level of representation. Since N-words are assumed to be negative, they must undergo movement to satisfy the Neg criterion. NPI's, in contrast, remain in situ. In (3b), therefore, *niente* moves to the Spec of the negative projection headed by *non* in order to satisfy the Neg Criterion. Since the relation between the head and the Spec of NegP is one of agreement, no cancellation of the two negative features occurs. On the other hand, sentences like (3a) are accounted for by allowing multiple specifiers for NegP and by the application of a rule of semantic absorption in SpecNegP, similar to the one proposed by Higginbotham and May 1981 for multiple questions. Absorption is an operation that creates a unique n-ary quantifier from a multiplicity of unary quantifiers.

- (5) a. Nimeni n-a vãzut nimic. N-word Neg-has seen N-word "Nobody saw nothing./ Nobody saw anything."
 - b. $\forall x \sim, \forall y \sim \rightarrow \forall x, y \sim$

One problem with this approach is that it cannot account for the asymmetry illustrated in (1) and (2). More precisely, this view does not explain why the quantificational force of N-words could vary with respect to the syntactic position of the N-word, nor does it explain the restrictions on the occurrence of the negative marker observed in Italian, Spanish and E Portuguese when the N-word is preverbal.

2.2 *N-words are indefinites*

Under this view, N-words are indefinites that lack quantificational force of their own (Ladusaw 1992, Acquaviva 1993, Giannakidou 1997, Giannakidou & Quer 1997). Given an analysis of indefinites à la Heim 1982 and Diesing 1992, N-words are interpreted as free variables. The negative reading associated with them is the result of them being bound by the negative operator. This implies that the scope of N-words does not derive from movement to a scope position, but from binding by a negative operator which is itself in a scope position. The negative operator is in SpecNegP and it unselectively binds one or more instances of N-words through a mechanism of operator variable coindexing at LF. The difference between NC and non-NC languages arises from the fact that in NC languages the negative operator is overt, whereas in non-NC languages it must be covert.

One problem with this view concerns the criterion for distinguishing between NC and non-NC languages, i.e. the overtness of the negative marker. Languages like Spanish, Italian, and E Portuguese are hard to classify as consistently NC or non-NC under this view, since sometimes the negative operator is overt and sometimes it is covert, as shown in (1) above.

Another problem is that it assimilates N-words to Negative Polarity Items (NPIs). However, N-words and Negative Polarity Items (NPIs) can be distinguished on the basis of the following properties, which, for lack of space we illustrate only with Spanish: first, N-words, but not NPIs, have an inherent negative meaning. When occurring as answers to questions, they have a negative meaning (Zanuttini 1991):

Second, unlike NPIs, N-words can be modified by *almost*. This suggests that they are universal quantifiers (Zanuttini 1991).

2.3 The ambiguity view. Deprez 1997

In this approach, N-words are analysed as indefinites that can have a weak or strong interpretation. When N-words are strong indefinites, they undergo Quantifier Raising (QR) and are interpreted as part of the restrictive clause of a tripartite structure. When N-words are weak, they remain internal to VP and are interpreted within the nuclear scope as cardinality predicates or as variables that can be unselectively bound. This ambiguity is lexically resolved in languages like English, which lexically distinguish between N-words with quantificational properties (i.e. stong) and N-words that behave like weak indefinites.

This approach seems to combine the virtues of the accounts discussed under 2.1 and 2.2 and thus seems to be more promising for explaining the contrast illustrated in (1) and (2). Thus, N-words seem to be strong in preverbal position and weak in a postverbal position. One could thus explain the restrictions on the co-occurrence of N-words with a negative marker by showing that the co-occurrence is blocked only when the N-word is strong. However, this account has nothing to say about the contrast between Italian, Spanish and E Portuguese, on the one hand, and Romanian, on the other hand.

In other words, the challenge for this view would be to explain why preverbal N-words in Romanian can co-occur with the negative marker, in spite of the fact that they are preverbal and therefore strong.

3. Our Proposal

We propose that N-words in the languages under consideration are unambiguously indefinites. In line with the indefinite theory, we assume that N-words introduce a variable which is unselectively bound by a unique negative operator (expressed by the negative marker). However, we depart from the indefinite theory, which assumes that the negative reading associated with N-words is the result of them being bound by the negative operator, and we propose that the negative feature of N-words is a feature intrinsic to the N-words themselves. Crucially, we propose to dissociate the [neg] feature from the quantificational or operator feature ([qu]). We thus allow on the one hand for the possibility of lexical items that are negative but which do not act as syntactic operators (N-words), and on the other hand for items that are quantificational (i.e. act as operators), but which are not negative (the nonnegative quantifiers):

(8) NPI: [-neg] [-qu] Neg Q: [+neg] [+qu] N-words: [+neg] [-qu] Other Qs: [-neg] [+qu]

Distinguishing between the [neg] feature and the [qu] feature allows for a better characterization of the similarities and differences between expressions that participate in NC (N-words, negative quantifiers, and NPIs). N-words are similar to NPIs in that they both lack a quantificational feature, and they both need a licenser. However, N-words differ from NPIs in two respects. First, NPIs can be licensed not only by the negative operator, but also by other non veridical operators. In contrast, N-words can only be licensed by the negative marker. Second, N-words are negative intrinsically, whereas NPIs get their negative reading as a result of the fact that they are bound by a negative operator. In fact, when NPIs are not licensed by the negative operator, but by some other type of operator, they are not interpreted negatively:

- (9) a. Did you see anyone?
 - b. If you say <u>anything else</u>, you'll get into trouble.

On the other hand, N-words share with Negative quantifiers the fact that they have a [neg] feature, but at the same time they differ from the latter in that they lack a [qu] feature, whereas the latter have one.

We also propose that whenever N-words act as negative quantifiers, (i.e. preverbal N-words), this is not because of a [qu] feature intrinsic to the N-word, but is instead the result of the interaction of the N-word with Focus. More precisely, we propose that preverbal N-words are in a Focus position. Since Focus is quantificational, an N-word in Focus acquires quantificational features. The remaining of this section is divided into two parts: we first show that the scope properties of N-words differ from the scope properties of genuine quantifier phrases (QPs). In the second part, we present arguments that preverbal N-words are interpreted as Focus.

3.1 The [qu] feature of N-words

As mentioned above, we propose that N-words have no [qu] features. This is what makes N-words different from negative quantifiers. This is confirmed by the fact that at least object N-words do not seem to be subject to QR, unlike genuine quantifiers. Unlike QPs, which can have narrow or wide scope with respect to another QP in the sentence, the relative scope of N-words is fixed and reflects the relative position of the N-word with respect to a QP at PF. Thus, the scope of a QP in object position can be either narrow or wide, as illustrated in (10), and this contrasts with an N-word in object position, which cannot take scope over a QP in subject position.

(10) (Cel putin/mai mult de) doi profesori au examinat fiecare student.	(R)
(Pelo menos) dois professores examinaram cada aluno.	(EP)
(Almeno) due professori hanno esaminato ogni studente.	(I)
"(at least) two professors examined every student."	
(11) Cel putin doi copii n-au vãzut nici un hot/ pe nimeni.	(R)
Pelo menos duas crianças não viram nenhum dos ladrões/ninguém.	(EP)
Almeno due bambini hanno visto nessuno /nessun ladro.	(I)
"At least two children saw none of the thieves/nobody."	

The sentences in (10) can have two interpretations: either that all the students were examined by the same two professors, or that every student was examined by two professors, but not necessarily the same two professors. In contrast, (11) only allows one reading, i.e. the one in which there are at least two children who didn't see anybody. The other interpretation, under which there is nobody who was seen by at least two children is excluded.

When occupying a preverbal subject position, however, N-words are quantificational (they induce a distributive reading), and take scope over a QP in object position. This is true not only about Spanish, Italian, and E Portuguese, but also about Romanian, as shown by (12):

(12) allows a distributive interpretation under which for each child, he saw either no thief at all, or one thief, or two thieves, but not more than two thieves. This seems to create a problem for our assumption that N-words do not have [qu] features. In the remainder of this section, we show that the relevant asymmetry is the one between preverbal and postverbal N-words, and that the [qu] feature of preverbal N-words is not an intrinsic feature of the N-word, but is acquired by virtue of association of the N-word with Focus.

3.2 Preverbal N-words are in focus

Our proposal is that the asymmetry noted above between N-words in a preverbal position, which have quantificational properties, and N-words in a postverbal position, which do not, is due to the fact that an N-word moved to the preverbal field is in a focus position. Since focus is quantificational, the N-word acquires quantificational properties, too. Before we embark upon providing arguments for this proposal, let us clarify what we mean by focus.

3.2.1 Focus The literature distinguishes two types of focus (Rochemont 1986, Erteschik-Shir 1997, Kiss 1998): (i) information focus, which simply expresses non-presupposed or new material, and which is present in every sentence, and (ii) identificational focus, which restricts a contextually presupposed set to an exhaustive subset for which the predicate holds. In this paper, by focus we strictly mean identificational focus. Semantically, identificational focus represents the value of the variable bound by an abstract operator expressing exhaustive identification. Syntactically, identificational focus acts as an operator, which moves (overtly or covertly) to a scope position in the Spec of a functional projection and binds a variable.

¹ Identificational focus also has a more restricted variant, *contrastive focus*. The latter operates on a closed set of entities whose members are known to the participants of the discourse (Kiss 1998)

- 3.2.2 Focus vs. Topic (Cinque 1990, Rizzi 1997, Zubizaretta 1998, etc) The structure of our argument is the following: (i) in all the languages under consideration, the preverbal position can in principle be either a Topic or a Focus position.² Topics have different properties than fronted focus constituents. Most importantly, Focus is quantificational (i.e. it acts as a syntactic operator), whereas Topic is not; (ii) we show that preverbal N-words pattern with preverbal focus rather than with preverbal topic. The following are the most important Topic/(dislocated) Focus differences³:
 - a) a preverbal Topic is clitic doubled, preverbal Focus is not:⁴
- (13) a. <u>Il tuo libro</u>, <u>lo</u> ho letto. (Rizzi, 1997) (I) the your book, it. CL have.I.1SG read "Your book, I've read."
 - b. <u>IL TUO LIBRO (*lo)</u> ho letto (, non il suo). the your book it.CL have.I.1SG read (not the his) "It's your book that I've read, not his."
- b) Focus is quantificational, Topic is not. Focus constituents are analysed as operators, and this clearly distinguishes them from Topic constituents. In other words, Focus binds a variable, whereas Topic binds a null constant. This

TopicFocusP has the function of introducing contrastive focus. One piece of evidence for this proposal is the distribution of clitics. The only constructions where the presence of the clitic is possible are the ones involving topic elements, to the exclusion of topic-focus elements.

² The situation in E Portuguese is controversial: Costa 2000 considers that E Portuguese doesn't have preposed focalisation, like Italian and that preposed constituents are interpreted as Topics. However, as pointed out by Ambar, 1999, E Portuguese has two available preverbal positions, as in (i).

⁽i) [TopicP [TopicFocusP [...

⁽ii) O teu vestido, a Joana vestiu-o. "Your dress. Joana put it on"

⁽iii) O teu livro, li-o. O TEU LIVRO li (, não dele).

³ To illustrate the general differences between preverbal Focus and Topic, we mostly used Romanian examples, but the same is assumed to hold for the other languages under discussion. ⁴ This is not always the case in Romanian, where fronted definite objects may be doubled even when they are contrastively stressed (focalized). However, as noticed by Cornilescu 2000, when the preverbal constituent is a bare quantifier (BQ), the contrast shows up in Romanian as well: preverbal BQ objects can never be clitic doubled in Romanian, and they are always Focus.

a. Orice as mânca, asa îmi e de foame. anything would.1sG eat, that's how hungry I am

b. *orice l-as mânca, asa îmi e de foame. anything itCL would-1SG eat, that's how hungry I am

128 DANIELA ISAC

can be shown by testing the properties of Focus vs Topic in two contexts: weak crossover contexts, and parasitic gaps contexts. Thus, a Topic never gives rise to weak crossover effects, as shown in (14a), whereas such effects are detectable with Focus, as shown in (14b):

- (14) a. Pe oricare copil_i îl_i rãsfatã mama lui_i. (Cornilescu, 2000)
 PE any childTOP him.CL spoils mother his
 "His mother spoils any child."
 - b. *Pe oricine_i rãsfatã mama lui_i.

 PE anybody.FOC spoils mother his

Moreover, given that parasitic gaps are licenced by variables, Focus licenses parasitic gaps, but Topic does not:

- (15) a. *Scrisorile_i le_i- ai trimis t_i fãrã sã verifici e_i. letters-the.TOP them.CL have.2SG sent without to check "The letters you have sent without checking." (Motapanyane 1994)
 - b. Doar douã scrisori_i am rimis t_i fãrã sã verific e_i . only two letters.FOC have 1SG sent without to check
 - c. Pe celelalte le- am verificat.

 PE others them.CL have.1SG checked

 "I've sent only two letters without checking. The others I've checked."

To sum up on the properties of Focus: focused constituents are analysed as (syntactic) operators, submitted to movement to a preverbal position.

Given the contrast above, we can test whether a preverbal N-word is Topic or Focus by testing whether it can co-occur with clitics and whether it has operator properties. Preverbal N-words cannot be clitic doubled:

- (16) a. *Niente_i non l_i' ha mangiato. (Frascarelli 2000) (I) nothing neg it.CL has eaten "(S)he hasn't eaten anything."
 - b. *Nada_i vestiu-o_i a Joana (EP)
 nothing put-on it.CL Joana
 "Joana didn't put on anything."
 - c. *A nadie_i lo_i vi. (S)
 nobody himCL saw-I
 "I didn't see anybody."

d. *
$$Pe \ nimeni_I \ nu \ l_i$$
- $am \ v \tilde{a} z u t$. (R)
PE N-word neg him. CL have. SG seen
"I have seen nobody."

Moreover, preverbal N-words induce WCO effects, as shown in (17) below, and they license parasitic gaps, as shown in (18).⁵

(18) a. *NIMIC n-am trimis fără să verific*. (R) Nothing not-have. 1SG sent without to check

These properties indicate that preverbal N-words have quantificational properties and that they pattern with Focus, rather than with Topics. Notice that the same properties are also exhibited by postverbal N-words, provided they are under stress. In other words, when under stress, the N-word is quantificational, even if it is postverbal, as shown by the parasitic gap test below (notice that in all (b) examples the N-word is in situ, and not focused):

- (19) a. *N-am trimis NIMIC fãrã sã verific*. (Alboiu 2000) (R) Not-have.1.SG sent NOTHING without to check
 - b. *N-am trimis nimic fãrã sã verific. not-have.1.SG sent nothing without to check "I haven't sent anything without checking."

(20) a. Ho mandato NIENTE senza controllare. (I)

b. ??Ho mandato niente senza controllare.

⁵ For WCO in E Portuguese we got inconclusive judgements from our informants. According to Costa 1998, one cannot test WCO in Portuguese. It is beyond the scope of this paper to clarify this issue. What is of interest for us is whether preverbal N-words in E Portuguese have [qu] properties. We think that they do, as shown by (19).

(21) a. (Eu) Não enviei NADA sem corrigir.

b. * Não enviei nada sem corrigir.

(I) not sent nothing without to check

"I haven't sent anything without checking."

This is not surprising, given that actually all focused constituents under heavy stress have quantificational features, even if they are not dislocated to a preverbal Focus position in the overt syntax (see Alboiu 2000 for Romanian).

Summing up, we have argued that focus is quantificational. Under the assumption that preverbal focus constituents raise to FocP, this amounts to saying that Foc⁰ hosts [qu] features. We will implement this by assuming that the focus feature itself is a [qu] feature. We have also argued that N-words do not have an intrinsic [qu] feature, but that they can become quantificational under emphatic stress, which is associated with a focus interpretation. Focus can be signalled either by stress alone, or by stress and a preverbal syntactic position.

3.3 The [neg] feature of N-words

As mentioned above, we propose that N-words have an intrinsic [neg] feature. This distinguishes N-words from NPIs, which acquire a negative interpretation by virtue of being bound by a negative operator, but do not have an intrinsic [neg] feature. This proposal is in conflict with the view expressed in Giannakidou 2000, who explicitly argues against the existence of a negative feature in N-words. According to Giannakidou, if N-words had a negative feature, one would expect them to (i) contribute negative meaning in the absence of the negative marker, and (ii) give rise to double negative readings under co-occurrence with another N-word, or with a negative marker. However, the examples in (22), (23) show that N-words in NC languages are used in non-negative contexts without contributing negation and (24) shows that they can co-occur without inducing a double negative reading:

(22) Perdimos la esperanza de encontrar ninguna salida. (Laka 1990) (S) lost.1PL the hope to find n- exit "We lost hope of finding our way out".

(23) E venuto nessuno? (Acquaviva1997)

Be.3SG come n-person
"Has anyone come?"

(24) a. Nessuno ha letto niente.
nobody have.3.SG. read nothing
"Nobody read anything."

b. Gianni non ha visto niente.
John not have.3SG seen nothing
"John didn't see anything."

Let us address these issues one by one. The fact that in NC languages N-words seem to be used in non-negative contexts without contributing negation is in our opinion the result of two factors: first, in languages where this happens, the paradigm of N-words happens to be lexically indistinguishable from that of NPIs. In fact, there are NC languages which do distinguish lexically between N-words and NPIs, and interestingly, these languages must use NPIs in non-negative contexts. This is exacty what one would expect under our view, since NPIs do not bear a negative feature, but N-words do:

Second, expressions that do contribute negative meaning in the absence of a negative marker do so in the sense that negation comes to have scope over the entire clause (the clause containing such expressions becomes a negative clause by virtue of including such an expression):

This suggests that the difference between such expressions and the N-words in NC languages may be one of the scope of the negative feature. The scope of the [neg] feature in N-words is confined to the N-phrase itself (phrasal scope), whereas the [neg] feature of the Dutch *niemand*, for example, has scope over the entire clause. The wider scope of the [neg] feature in the Dutch *niemand* can be explained by the fact that such expressions have not only a [neg] feature, but also a [qu] feature, and by assuming that it is the [qu] feature that is actually responsible for raising the [neg] feature to a position from which it has

132 DANIELA ISAC

scope over the entire clause. Such expressions can be construed either as in (27a) or as in (27b):

(27) a.
$$niemand = \lambda P \ \forall x \ [person (x) \rightarrow \neg P(x)]$$
 (Giannakidou 2000)
b. $niemand = \lambda P \ \neg \exists x \ [person (x) \& P(x)]$

Given these construal rules, it is clear that negation alone cannot provide the \forall or \exists quantifier. Rather, the quantifier and the negative feature are distinct and independent.

To sum up, the fact that these expressions can contribute negative meaning in the absence of a negative marker is an outcome of their [qu] feature. Given that quantified expressions undergo QR, all of their features (including the [neg] feature) are raised to a scopal position. In contrast, N-words do not provide negative meaning to the whole clause simply because N-words do not have a [qu] feature that would trigger QR, and not because they lack a [neg] feature.

The fact that N-words do not give rise to double negative readings under co-occurrence with another N-word or with a negative marker can be explained in a similar way. More specifically, a double negative reading does not obtain when two negative features co-occur, but rather when two negative quantificational features do so. If the [neg] feature does not cluster together with a [qu] feature, it will never undergo QR to a scope position and will never come to cancel another [neg] feature in a similar scope position. As proposed above, the scope of the [neg] feature in N-words is confined to the N-phrase itself (phrasal scope). This means that the presence of two N-words in a sentence will not give rise to double negative readings, because the two negative features have their own separate domain, and the two never interact.

Notice, however, that a double negative reading can actually obtain even in these languages. Sentences like (28) below have two possible interpretations: one which is equivalent to a positive sentence, in which N-words undergo simplification, and one under which N-words do not undergo simplification, i.e. a simple negation interpretation:

```
(28) <u>Nimeni nu</u> iubeste pe <u>nimeni</u>. (true in no-love worlds) Nobody not loves PE nobody "Everyboddy loves somebody."
```

Interestingly, the double negative reading appears only if both of the N-words are under stress (Puskás 1998, Vinet 1998, Isac 1998). As shown in the

preceeding section, stressed N-words are quantificational, even if in a post-verbal position.

To sum up, in this section we proposed that the [neg] feature should be dissociated from the [qu] feature. N-words have a [neg] feature but no [qu] feature. The fact that N-words do not provide negative meaning to the whole clause is explained by their lack of a [qu] feature. No [qu] feature means that N-words are not subject to QR and so the scope of N-words never comes to go beyond the N-phrase itself. We also proposed that even though N-words do not have an intrinsic [qu] feature, they can become quantificational under emphatic stress, which is associated with a focus interpretation.

Our analysis can also explain the two properties of N-words which distinguishes them from NPIs, according to Zanuttini 1991. N-words have a negative meaning when they occur as answers to questions because they are under focus. It is well known that one way to test whether something is in focus is to test whether that XP can be the correspondent of a wh-word in the answer to a wh-question. Notice that in (6) above, *nadie* is precisely the answer to the wh-word, which means that it is in Focus. The [qu] properties of *nadie* in (6) thus follow from the fact that it is in Focus, and not from the existence of an inherent negative [qu] feature. As to the possibility of modifying N-words with *almost*, this does not necessarily prove that N-words have quanificational properties, and might simply be a reflexion of the fact that 'no x P(x)' is equivalent logically to 'for any x not P(x)'. Due to this equivalence, N-words share properties with universal quantifiers, but this does not necessarily mean that we can represent N-words as universal quantifiers (Dobrovie Sorin 1999).

4. Co-occurrence with negative markers

In this section we turn to the asymmetry between languages in which a preverbal N-word co-occurs with a negative marker (e.g. Romanian) and languages in which a preverbal N-word cannot co-occur with a negative marker (Italian, Spanish, E Portuguese). We propose the following: (i) in Italian, Spanish and E Portuguese, the FocP, which is located in the CP domain, syncretically hosts both [foc] features and [neg] features. Feature syncretism has been proposed before in the literature, by Laka 1990, Zubizaretta 1998, Frascarelli 2000, and others. Below, we will briefly discuss Frascarelli's 2000 proposal; (ii) in contrast to Italian, Spanish and E Portuguese, in Romanian, the [foc] feature and the [neg] feature are hosted by distinct heads.

4.1 Frascarelli 2000

Frascarelli proposes that FocP in Italian hosts both [foc] features and [neg] features. In her view, the negative marker *non* heads its own maximal projection, and it must move to F in overt syntax in order to check the [neg] feature in F. Alternatively, an N-word can move to SpecFP and check both [neg] and [foc] features simultaneously. The overt movement of the N-word is not necessary, since the [neg] feature in F is weak (the [neg] feature can be checked in situ). What triggers movement of the N-word to the spec of FP is the focus feature. The N-word moves in order to get contrastive focus interpretation. Only one neg element is allowed in FP in overt syntax (either the neg head *non*, or an N-word).

We find Frascarelli's proposal that FocP in Italian hosts both [foc] features and [neg] features interesting, and we adopt it. However, we find the actual implementation of this idea problematic for several reasons. First, it is not clear why only one negative element (either *non* or the N-word) is allowed in FP in overt syntax, if each of them raises to FP for different reasons. In particular, in Frascarelli's view, *non* raises to F⁰ to check a [neg] feature, while the N-word raises to Spec FP to check the [foc] feature. Second, if the [neg] feature on F is weak, as she proposes, how come *non* **must** move to F? What prevents *non* from staying in its base position (i.e. in the head of NegP) when an N-word moves to specFP? If what drives movement of *non* to F is the [neg] feature on F, and if this feature is checked by the N-word, then movement of *non* to F is indeed unmotivated, but there is no reason why *non* should not be able to stay in situ, since there is no intrinsic feature on *non* that drives movement.

4.2 *Our proposal*

We build on Frascarelli's 2000 proposal that FocP in Italian hosts both [foc] features and [neg] features and we claim that this hypothesis could be extended to Spanish and E Portuguese. However, instead of [neg] features, we hold that Foc actually hosts polarity (pol) features. This could explain why not only negative elements can be focused, but also affirmative ones. In addition, we claim that the negative marker *non* Merged in the Focus head, rather than be Merged in the head of the NegP and then move to Focus. When *non* is Merged in Foc⁰, it instantiates both the [pol] feature and the [foc] feature of Foc. This is confirmed by (29) below, which shows that the negative head can be contrastively focused in Italian:

⁶ We will not deal here with affirmative polarity features and we will be concerned exclusively with negative ones.

(29) Se ti credessi lo farei, (Frascarelli 2000)

If you believe.1.SG it do.COND.1.SG

ma io NON credo a quello che dici.

but I not believe to what that tell.2.SG

"If I believed you, I would do it, but I do not believe what you are saying."

Alternatively, the [foc] feature in FocP can be checked by XPs that move in the Spec FP. Such XPs can be affirmative, or negative. When an affirmative XP raises to Spec FP, it will check the [foc] feature in F, and the [pol] feature in F is checked independently, by the negative marker:

- (30) a. A LUIGI non ho detto la verita. (Frascarelli, 2000) TO LUIGI not have 1sG told the truth
 - b. Non ho detto la verita A LUIGI. not have 1.5G told the truth TO LUIGI "I haven't told the truth to LUIGI."

When a negative XP (i.e. an N-word) raises to Spec FP, it will check both the [pol] and the [foc] feature. The [pol] feature will be valued as negative, by agreement with the [neg] feature on the N-word, and this is why the merger of a negative marker bearing [neg] features becomes superfluous, and thus ungramamatical. (31) shows that preverbal N-words are interpreted as contrastive, and thus that the preverbal N-word does check a [foc] feature in F:

- (31) a. Lei cosa ci guadagna? NULLA ci guadagna. (Frascarelli 2000) "What does she get out of it? She gets NOTHING."
 - b. Sto in mani sicure ahah pecché meglio 'e voi NISCIUNO le capisce 'sti cose.

"I am in safe hands because NOBODY understands these things better than you."

There is one contrast that emerges within the set of XPs that can move to SpecFP. This XP can be an N-word, but not an NPI. Clearly, this contrast cannot be explained by the fact that N-words have [neg] features, whereas NPIs do not, since not all XPs that raise to Spec FP have to be negative. The difference cannot be due to the definiteness properties of the N-words vs NPIs either, since both are indefinites. We think that the relevant difference has to do with the particular way in which N-words and NPIs are licensed. More specifically, NPIs are licensed under c-command, and must be in the c-

136 DANIELA ISAC

command domain of their licensor, whereas N-words are licensed by feature checking, and must be in the checking domain of their licensor. Both are licensed by the [pol] feature in F, but NPIs must always be within the c-command domain of F⁰, whereas N-words can also occupy the SpecFP position. SpecFP is in the checking domain of F, but not in its c-command domain. This difference can ultimately be related to the fact that N-words, but not NPIs, have a [neg] feature. Checking the [pol] feature requires a matching feature on the goal. N-words have a matching feature, i.e. [neg], but NPIs do not.

The same type of analysis can be extended to Spanish and E. Portuguese. The E Portuguese case is a little more complex but is amenable to the same type of account. According to Ambar 1999, preverbal focused material in E Portuguese is hosted by a TopicFocus Phrase, which has two types of features: [focust] and [topicf]. We propose that TopicFocusP also hosts [neg] features. The [neg] features can be checked either by the negative marker, or by a dislocated N-word. When the negative marker is merged in the head of TopicFocusP, it simultaneously checks the [neg] feature, as well as the [focust] and the [topicf] features. Consequently, the topic and the focus constituents will be under the scope of the negative marker, and the only type of focus that can obtain is information focus (in situ). Alternatively, the [neg] feature in TopicFocusP can be checked by a dislocated N-word. When an N-word is moved to SpecTopicFocusP, it checks the [neg] feature and the focus feature. However, an N-word, and a bare indefinite in general, cannot be a topic, and thus cannot be assumed to check the topic feature. A contrast thus emerges between preverbal contrastive elements that can be assumed to have both [topicf] and [focust] features, and to check both topic and focus features simultaneously (Ambar 1999), and preverbal bare indefinites (N-words and bare quantifiers), which do not have [topicf] features and cannot check the latter against the TopicFocus head. The question arises as to how this feature is checked? We propose that the [topicf] feature is checked by the verb that raises to the head of TopicFocusP. If this is on the right track, then the proposed contrast between preverbal bare indefinites and preverbal focus constituents that are not bare indefinites should be visible with respect to verb movement. In particular, the prediction is that the verb moves in front of the subject with preverbal N-words (and bare indefinites), but not with the other preverbal focus constituents. The expectation is borne out. In (32) below, that contains a preverbal N-word, the subject must follow the verb, in contrast to (33), that contains a preverbal focused object that is not a bare indefinite, and in which the verb follows the subject:

- (32) Preverbal N-word (Bare indefinite) as Focus: V inversion
 - a. A ninguÈm (*lhe) devolveu a Maria o seu manuscrito. (EP) to nobody returned Maria his manuscript
 - b. *A ninguÈm, a Maria (lhe) devolveu o (seu) manuscrito.
- (33) Preverbal XP (not a Bare indefinite) as Focus: no V inversion
 - a. A tarte...a Maria comeu... (outras coisas nao sei...) (EP)
 The pie... Mary ate... (what more she ate, I don't know)
 - b. *A tarte...comeu a Maria.

To sum up, in Italian, Spanish and E Portuguese the two features in F ([foc] and [pol]) can be checked simultaneously, by the same lexical item, or separately, by independent lexical items. The first option is instantiated by either *non* or an N-word as a checker; the second option is manifested in (30), where the focused constituent *A LUIGI* checks the [foc] feature and the negative marker *non* independently checks the [pol] feature.

The asymmetry between these languages and Romanian regarding the cooccurrence of preverbal N-words with the negative marker is a consequence of the fact that the two relevant features, [pol] and [foc], are located in different syntactic positions in Romanian. In particular, we claim that in Romanian, the [pol] feature and the [foc] feature are hosted by disjoint projections, as in (34):

$$(34) \begin{bmatrix} FP & F^0 & [\dots & Pol^0 & Pol^0 & Pol^0 \end{bmatrix}$$

The negative marker heads PoIP and bears a [neg] feature. Since the [foc] feature is hosted by an independent projection, the negative marker in Romanian does not simultaneously check the [foc] feature, as it does in Spanish, Italian, and E Portuguese. The [foc] feature can thus be checked independently, by a preverbal N-word.

This asymmetry between the location of the [pol] and [foc] features in Romanian vs Spanish, Italian, and E Portuguese is confirmed by the fact that in Romanian, preverbal N-words can be separated from the verb by lexical material, whereas in Italian, Spanish, and E Portuguese it cannot. In this latter group of languages, the preverbal N-word must be adjacent to the verb.

(35) Nimic copilul ãsta n- a mâncat.

nothing child this neg has eaten

"This kid has barely eaten."

- (36) a. ?*A nadie, Maria le devolvió su manuscrito. to nobody Maria DAT.CL returned his manuscript (S)
 - b. *A nadie le devolvió Maria su manuscrito*. to nobody DAT.CL returned Maria his manuscript (Zubizaretta 1998)
- (37) a. *A nessuno Mario ha parlato. (I) to nobody Mario has talked.
 - b. A nessuno ha parlato Mario. 7 to nobody has talked Mario "Mario talked to nobody."

5. Conclusion

In this paper we have shown that the lexical ambiguity of N-words in Italian, Spanish and E Portuguese is only apparent. N-words in these languages are unambiguously indefinites or pure variables. We proposed that N-words have [neg] features, but no [qu] features. This explains why they are not able to render the sentence they occur in negative. The [neg] feature of N-words is confined to the level of the constituent representing the N-word, and it never raises to a scope position from which it could negativize the whole clause. Raising to a scope position is a property of phrases bearing [qu] features.

We accounted for two types of asymmetries: (i) the asymmetry between the preverbal and the postverbal position of N-words, which applies to Italian, Spanish, E Portuguese, and Romanian alike, and (ii) the asymmetry between languages in which a preverbal N-word co-occurs with a negative marker (e.g. Romanian) and languages in which a preverbal N-word cannot co-occur with a negative marker (Italian, Spanish, E Portuguese). In order to account for the first asymmetry, we proposed that the quantificational feature of preverbal Nwords is not to be identified with the negative feature within the N-word, but is instead the result of the interaction of the N-word with Focus. We proposed that preverbal N-words are in a Focus position. Since Focus is quantificational, an N-word in Focus acquires quantificational features. The differences between Italian, Spanish, and E Portuguese, on the one hand, and Romanian, on the other hand, have been accounted for by the particular way in which [pol] features cluster together with [foc] features in these languages. In Spanish, Italian, and E Portuguese, both [pol] and [foc] are features of the same functional head, whereas in Romanian they are features of independent functional heads.

⁷ For examples from E Portuguese, see (32) above.

References

- Acquaviva, Paolo. 1993. *The Logical Form of Negation: a Study of Operator-variable Structures in Syntax*. PhD Diss. Scuola Normale Superiore, Pisa.
- Alboiu, Gabi. 2000. *The Features of Movement in Romanian*. PhD Dissertation. U of Manitoba.
- Ambar, Manuela. 1999. "Aspects of the Syntax of Focus in Portuguese". *The Grammar of Focus*, ed. by Georges Rebuschi and Laurice Tuller, 23-53. Amsterdam: John Benjamins.
- Besten, Hans den. 1986. "Double Negation and the Genesis of Afrikaans". *Substrata versus Universals in Creole Languages*, ed. by Pieter Muysken and Norval Smith, 185-200. Amsterdam: John Benjamins.
- Cinque, Guglielmo. 1990. *Types of A-bar Dependencies*. Cambridge: MIT Press. Cornilescu, Alexandra. 1997. "Some Notes on the Syntax of the Subject". *Revue Roumaine de Linguistique* XLII. 101-147.
- Cornilescu, Alexandra. 2000. "The Double Object Construction in Romanian". *Comparative Studies in Romanian Syntax*, ed. by Virginia Motapanyane, 83-135. Amsterdam: Elsevier.
- Costa, João. 2000. "Focus in situ: Evidence from Portuguese". *Probus* 12: 2.187-228.
- Deprez, Viviane. 1997. "A Non-unified Analysis of Negative Concord". *Negation and Polarity: Syntax and Semantics*, ed. by Danielle Forget et al., 53-75. Amsterdam: John Benjamins.
- Diesing, Moly. 1992. Indefinites. Cambridge, Mass.: MIT Press.
- Dobrovie Sorin, Carmen. 1999. "Settling Discord over Negative Concord". LSRL 29 handout.
- Dowty, David. 1994. "The Role of Negative Polarity and Concord Marking in Language Reasoning". *SALT IV: Proceedings of the Fourth Conference on Semantics and Linguistic Theory*. ed. by Mandy Harvey and Lynn Santelmann, 114-144. Ithaca, NY, Cornell University.
- Erteschik-Shir Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge: CambridgeUniversity Press.
- Frascarelli, Mara. 2000. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht: Kluwer Academic Publishers.
- Giannakidou, Anastasia. 1997. *The Landscape of Polarity Items*. PhD Dissertation. University of Groningen.
- Giannakidou, Anastasia & Josep Quer. 1997. "Long-distance Licensing of Negative Indefinites". *Negation and Polarity: Syntax and Semantics*, ed. by Danielle Forget et al., 95-113. Amsterdam: John Benjamins.
- Giannakidou, Anastasia. 2000. "Negative...concord?". *Natural Language and Linguistic Theory* 18.457-523.

- Haegeman, Liliane & Raffaella Zanuttini. 1991. "Negative Heads and the Neg Criterion". *The Linguistic Review* 8:233-252.
- Haegeman, Liliane. 1995. *The Syntax of Negation*. Cambridge: Cambridge University Press.
- Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases, PhD Diss. U.Mass.
- Higginbotham, Jim. & Robert, May. 1981. "Questions, Quantifiers, and Crossing". *The Linguistic Review* 1.41-79.
- Isac, Daniela. 1998. *Sentence Negation in English and Romanian*. PhD Dissertation. University of Bucharest.
- Kiss, Katalin. E. 1998. "Identificational Focus versus Informational Focus". *Language* 74:2.245-273.
- Ladusaw, William A. 1992. "Expressing Negation". *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory*, ed. by Chris Barker and David Dowty, 237-259 Columbus: Ohio State University.
- Laka, Itziar. 1990. Negation in Syntax: on the Nature of Functional Categories and Projections. Ph.D Dissertation, MIT.
- Motapanyane, Virginia. 1994. "On Preverbal Positions in Romanian". *Canadian Journal of Linguistics* 39:1.15-36.
- Motapanyane, Virginia. 2002. Romanian Left Periphery. Ms. U. of New Brunswick.
- Puskás, Genoveva. 1998. "On the Neg-criterion in Hungarian". *Acta Linguistica Hungarica*, 45:1-2.167-213.
- Rizzi, Luigi. 1982. Issues in Italian Syntax, Dordrecht: Foris.
- Rizzi, Luigi. 1990. Relativized Minimality, Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". *Elements of Grammar: Handbook of Generative Syntax*, ed. by Liliane Haegeman, 281-339. Dordrecht: Kluwer.
- Rochemont, Michael. 1986. *Focus in Generative Grammar*. Amsterdam: John Benjamins.
- Vinet, Marie-Thérèse. 1998. "Contrastieve Focus, N-words, N-words and Variation", *Revue canadienne de linguistique*, 43:1.121-141.
- Wouden, Ton van der & Franz Zwarts. 1993. "A Semantic Analysis of Negative Concord". *Semantics and Linguistic Theory (SALT) 3*, ed. by Utpal Lahiri and Adam Zachary Wyner, 202-219. Ithaca, NY: Cornell.
- Zanuttini, Raffaella. 1991. Syntactic Properties of Sentential Negation. A Comparative Study of Romance Languages. PhD Dissertation. University of Pennsylvania.
- Zubizaretta, Maria-Luisa. 1998. *Prosody, Focus and Word Order*, Cambridge, Mass.: MIT Press.

ROMANCE AND 'SOMETHING ELSE'

DANIELA ISAC & CHARLES REISS UOAM/Concordia University Concordia University

1. Introduction and proposal

In this paper, we examine referentially dependent elements other than pronouns and anaphors, focusing on the syntactic and semantic properties of *X*-else elements (something else, anything else, everywhere else, nobody else) illustrated in (1):

(1) I visited Mary_i and Peter visited [someone else]_{i≠i}

The interpretation of *someone else* in (1) depends on an antecedent - *Mary* - but the referential dependency that is at stake cannot be characterized as coreference, as in the standard referential dependency cases that are subject to Binding Theory. Clearly, the referential index of the antecedent - *Mary* - is not identical to the referential index of *someone else*.

We propose that even though the referential index of the entire *X-else* expression cannot be coindexed with the antecedent, parts of *X-else* can establish such an anaphoric relation with the antecedent. More specifically, we propose that *X-else* can be decomposed into constituent parts, and that a covert anaphoric element is 'contained' or 'hidden' within these parts. Unlike Culicover & Jackendoff (1995), who locate this covert anaphoric element exclusively within the *else* part, as in (2b), we propose that this covert anaphoric element is actually contained within both the *X* part and the *else* part, as in (2c). In our view, *else* operates on a set of elements *x* defined by the element which precedes *else* and it excludes from this set the member which is anaphorically related to the antecedent, picking up the complementary set.

- (2) a. someone else
 - b. someone other than x_i , where x_i is coindexed with the antecedent.
 - c. \exists x person (x) else person (x), where x = {x₁, x₂, x₃,...,x_i,...} and x_i is coindexed with the antecedent.

This proposal is supported by the syntactic analysis that we assign to *X-else*. We propose that *X-else* has the syntactic structure of a partitive construction, as in (3). Evidence for a partitive analysis is identified in several Romance languages (French, Italian, Romanian).

(3) a.	$[Q NP_1]_i$	$[_{PP}\ P_{PART}$	$[NP else [NP_2]_i]$
b.	some-thing	Ø	else e
c.	∃х		else x

Section 2 focuses on the semantics of *X-else*. We propose that the interpretation of *X-else* is compositional and discuss the dependency relation between *X-else* and the antecedent. We compare *else* with similar predicates like *other* and *different*. Section 3 proposes a syntactic analysis of *X-else* as a partitive construction and brings in relevant data from Romance languages. Section 4 includes the conclusions.

2. Semantics of 'X-else'

Before we start, let us be more precise about the kind of constructions that we focus on in this paper. We distinguish between two types of *else*. The first type - 'exclusive *else*' - serves to exclude previously mentioned individuals from the set of possible referents for *X-else*, as illustrated in (1) above. In contrast, the second type - 'additive *else*' - does not exclude the antecedent, but adds a new referent to the one introduced by the antecedent. In (4) below, *else* does not exclude *a hint* from the set of possible referents for *what else*, but adds a new object to the referent of *a hint*. A suitable paraphrase for (4) would be (5):

- (4) I gave you a hint already. [What else] do you want?
- (5) [What else in addition to the hint I gave you] do you want?

This distinction is confirmed by the fact that in Russian, these two uses of *else* correspond to two different lexical items.

```
(6) a. escho shto nibyd' (Russian) else what thing "what else"
b. shto nibyd' drugoje what thing other / in addition
```

Also, in French, the two interpretations of *else* are rendered by two different expressions, as illustrated in (7):

(7) Qui encore qui d'autre (French) quoi encore quoi d'autre "who/what else" (additive) "who/what else" (exclusive)

In this paper, we concentrate mostly on 'exclusive' *else*, and discuss 'additive' *else* only when it contrasts with 'exclusive' *else* in a relevant way.

In this section we focus on the interpretation of *X-else*. In the first part we argue that *X-else* is compositional and thus that its interpretation should follow from the interpretation of its constituent parts, and in the second part we examine the dependency relation between *X-else* and the antecedent. Following Culicover & Jackendoff (1995), we argue that this dependency can be captured by an anaphoric relation of the standard type and that the anaphor is covertly 'contained' within *X-else*. Contra Cullicover & Jackendoff, however, who propose that the covert anaphor is hidden just within *else*, we argue that a covert anaphoric element is contained within each of the constituent parts of *X-else*.

2.1 Compositionality

The compositionality of X-else is supported by the fact that one can tease apart the (semantic) contribution of each element within X-else to the interpretation of the whole complex expression. Distributionally, both X and else can be independent of each other. Else in (8) can appear without X, and X in (9) can occur without else:

- (8) Eat your dinner, or else no TV.
- (9) Peter saw someone, and Mary saw someone else.

Else also appears in combination with various X's: wh-words, or quantifiers:

- (10) a. who / what / when / where / how else (wh-else)
 - b. somebody / everybody / nobody else (Q-else)

When *X* is a quantifier, its scope possibilities are completely parallel to its scope possibilities when it occurs without *else*. Thus, in (11) below, *everyone else* can bind an indefinite or another quantifier, just like *everyone* (Culicover & Jackendoff 1995):

(11) John bought a red balloon, but everyone else bought a teddy bear.

(a teddy bear is in the scope of everyone else)

Similarly, *nobody else* in (12) below can license negative polarity items, just like *nobody* (Culicover & Jackendoff 1995):

(12) Nobody else has ever bought anything like that.

And finally, the difference between *someone else* and *something else* is parallel to the difference between *someone* and *something* (Culicover & Jackendoff 1995).

- (13) a. Bill is a poet, but Susan wishes he were {something / *someone} else.
 - b. Bill wishes to be {something / *someone} like a poet.

On the basis of the above, we will assume that the interpretation of X-else is compositional. The next step is to determine the contribution of each of the two elements -X and else - to the interpretation of X-else. The general strategy will be to identify simple anaphors embedded within one of the 'pieces' that compose X-else. The background assumption is that expressions such as X-else should be treated as complex or extended anaphors and that their anaphoric behavior follows from the existence of simple anaphors which are embedded within them (Keenan 1988, Safir 1991).

- 2.1.1 *The contribution of 'X'*. Let us start from the observation that X in X-else must be either a bare quantifier, or a wh-element, as illustrated in (14a,b):
- (14) a. someone / somebody / something / somewhere / etc. else everyone / everybody / everything / everywhere else anyone / anybody / anything / anywhere else noone / nobody / nothing / nowhere else
 - b. who / what / when / where / how else

What the X's in (14a) and (14b) above share is the fact that they introduce a quantifier and a variable. The quantifier can be an existential, universal or interrogative/wh one, and the variable is restricted by a property such as [+human], [-animate], [+place], [+time], etc.

Both the presence of the quantifier and that of the variable are crucial. The examples in (15) below are unacceptable because X does not introduce a variable:

(15) *the student else *Peter else

The unacceptability of (16), on the other hand, is due to the fact that X does not introduce a quantifier. In (16), X is an indefinite. According to Diesing (1992), indefinites introduce variables that are bound by a quantifier that is contributed elsewhere in the sentence (existential closure, unselective quantifiers, etc.):

(16) * a student else

The ungrammaticality of (17), though, seems to be problematic for our generalization that X must contribute both a variable and a quantifier:

(17) *which else

*two (students) else

*many else

X in (17) seems to satisfy both conditions in our generalization; it introduces both a quantifier and a variable. However, else expressions with the X's in (17) are ill-formed. The examples under (17) contrast with the ones in (14) in that X in (17) is a "non-bare" quantifier, while in (14) it is a "bare" quantifier. Syntactically, bare quantifiers have a reduced structure as compared to non-bare ones. In particular, bare quantifiers are exhaustively dominated by an NP node (Cinque 1986, Dobrovie-Sorin 1994), whereas non-bare quantifiers are DPs. Semantically, this distinction correlates with the D-linking properties of the respective items; (bare quantifiers are D-free, while non-bare ones are Dlinked (Pesetsky 1987), as well as with their ability to undergo Quantifier Raising (QR) and create a quantifier-variable configuration (Cinque 1986). According to Cinque (1986), bare quantifiers undergo QR and create a quantifier-variable configuration, whereas non-bare quantifiers ('quantified NPs' in Cinque's terms) do not undergo QR and do not create a quantifier-variable configuration. Even though the quantifier within a non-bare quantifier also has quantifier features, its scope is restricted to its NP complement. In contrast, a bare quantifier takes scope over the whole sentence. So our generalization regarding the properties of X could be restated as follows: X must contribute a variable and a quantifier that undergoes QR, i.e. a quantifier whose scope is not confined to the DP it is part of.

The question that arises is whether one can find any explanation for this generalization. Why should the quantifier introduced by X have to raise? Our

suggestion is that it must do so because the *else* phrase must raise to a scope position from which it can 'see' the antecedent. However, this is merely a suggestion, and we have no full account of this.

Going back to the semantic contribution of X, our proposal is that the variable introduced by X ranges over a set of elements x, one of which is coindexed with the antecedent. In other words, the covert anaphoric element is located within X and the contribution of X is to define a set that includes among its members an x_i which is coindexed with the antecedent. In (18), *someone* defines the set of all x's that are human, and this set includes among its members an x_i which is anaphorically linked to Mary. To anticipate, we will propose that the contribution of else – the other compositional part of X-else – is to exclude x_i from the set defined by X.

(18) I visited Mary_i and Peter visited [someone else]_{i≠i}

Notice that X always defines a superset with respect to the antecedent(s). The variable introduced by X is restricted by a property (someone defines the set of x's that are human, somewhere defines the set of all x's that are places, etc.; similarly, who defines the set of all x's that are human, what defines the set of all x's that are objects, etc.) and the antecedent is defined by a larger set of properties. Since the number of individuals that are defined by a larger number of properties is more restricted than the number of individuals defined by fewer properties, X is always a superset with respect to the antecedent(s). In (18), X in X-else is defined by the unique property [+human], whereas the antecedent – Mary – is defined by a larger set of properties: [+human], [+female], etc. Notice also that the set of properties that defines the antecedent necessarily includes the property that restricts the members of X; otherwise no anaphoric relation can obtain between a member of X and the antecedent. In (18) the set of properties that defines Mary necessarily includes the property [+human] that defines the members of X.

- (19)??I visited Mary_i and Peter visited [something else]_i / [somewhere else]_i / [somewhat else]_i
- 2.1.2 The contribution of else. Intuitively, else serves to exclude previously mentioned individuals (or, rather, an x_i which is coindexed with the antecedent) from the set of possible referents for X-else. In this sense, else could be taken to perform the set theoretical operation of relative complementation: else picks up the relative 'complement' of the antecedent from the set of possible referents.

Apart from *else*, there are other expressions that perform similar exclusion operations. This semantic effect of exclusion can be syntactically encoded in an explicit way with certain phrases that can occur in construction with *other* and *different*, and which serve to exclude individuals mentioned in the same NP:

(20) Some other student besides Albert will fail the next exam.

I hope that someone other than Cindy brings the salad.

Everyone <u>else besides</u> Jake will drink cheap beer if it's available. (Hand, 1987)

This is <u>different from</u> what you said before. <u>a different</u> assignment <u>than</u> the one I passed out.

In this section we briefly discuss *other* and *different* and show how *else* differs from them

2.1.2.1 'Else' vs 'other'. According to Safir (1991), other should be analysed as a particular type of contrastive predicate, i.e. as an exclusion predicate:

(21) Contrastive predicates:

- a. <u>Similarity predicates:</u> someone unlike / similar to / very dissimilar to / completely different from / such as himself
- b. <u>Comparatives</u>: a woman taller / more generous / less wealthy than herself
- c. <u>Exclusion predicates</u>: noone apart from / but himself, everyone except himself, someone besides / other than / in addition to himself

Even though *else* also serves to exclude the antecedent from the set of possible referents for X-else, there are reasons to reject an analysis of *else* along the lines of *other*. As argued by Safir (1991), *other* is a predicate with two overt arguments (X other than Y). The Y argument in *other* is introduced by than and it expresses the excluded individual, while X expresses the set from which the exclusion must be performed. Thus, in (22a), the set from which the exclusion must be performed is the set of humans and is defined by *someone*, whereas the member that is excluded from this set is Bob. Both the set X from which the

¹ The Y argument can also be introduced by *besides*. However, the reading of *some other student besides Albert* is not exclusive, but additive. Under this additive reading, *other* is similar to *else*, in that additive *else* can also take a Y argument introduced by *besides* (*someone else besides Peter*).

predicate *other* will eventually exclude an individual, and the individual to be excluded are expressed overtly within the NP including *other*.

- (22) a. someone other than Bob
 - b. ???someone else than Bob

In contrast, *else* only has one overt argument (*X-else*). This argument expresses the set from which the antecedent is excluded. Crucially, the individual to be excluded is not overtly expressed within the NP containing *else*, but is part of the previous discourse. This is in sharp contrast with the argument expressing the individual to be excluded in *other* expressions, which cannot be expressed in the previous discourse. In other words, *else* has a discourse anaphoric reading, whereas *other* has no such reading. So there is no overlap between the two. The only exception is *another*, which can have an anaphoric reading, as in (23) below:

- (23) a. Mary met a student and Jane met another one / student.
 - b. Mary met a student and Jane met someone else / *(a) student else

However, even in this example, *another* has different properties than *else*. In particular, even though both *another* and *else* are associated with a unique overt NP internal argument that expresses the set from which the exclusion must be performed (X), the nature of this argument differs with *another* vs *else*. The X argument associated with *another* can be a common noun, *student* in (23a), whereas the corresponding argument with *else* cannot. Actually, the nature of this argument distinguishes not only between *another* and *else*, but generally between *other* and *else*. When it occurs in X-else, X can only be a bare quantifier or a wh-expression, whereas with *other*, it can also be an indefinite DP, or a non-bare quantifier, as illustrated in (24):

(24) someone / something / somewhere / *a student / *three students else someone / something / somewhere / a student / three students other than α

Moreover, X can overtly occur on either side of *other*, but it can only occur on one side of *else*.

(25) one student other than Bob; one other student someone else: *else someone

2.1.2.2 'Else' vs 'different'. Just as else, different involves some kind of comparison between two or more elements. From the point of view of the syntactic realisation of these arguments, one can distinguish several uses of different.

First, the two arguments can be overtly expressed DP internally, as in (26) below.

(26) a different assignment than the one I passed out.

This is clearly something that distinguishes different from else.

Second, one of the arguments can be overtly expressed external to the constituent containing *different*, in the previous discourse.

(27) Lou bought a book of origami. Daniel bought a different book. (Tovena & van Peteghem, 2003)

As mentioned above, *else* always has a discourse anaphoric reading, so this is a context that brings *different* and *else* together.

Thirdly, the two arguments can be 'collapsed' together in the *different*-phrase. Syntactically, only one argument is overt.

(28) The children like different cartoon shows. (Carlson 1987)

Semantically, different usually expresses a comparison between tokens of the same type. This is most obvious in (28) above, where all elements that are being contrasted come from the same set, namely the set described by the different-phrase.² The same is true about (26) and (27) above, in which the term of comparison is overt. Even though the referent of the different-phrase is disjoint from that of the antecedent and of the than phrase, the latter still plays a role in determining the former by 'specifying the shared descriptive content' (Tovena & van Peteghem 2003). In (26), the term of comparison and the item with which it is being contrasted are both tokens of the type 'assignment', while in (27), they are both 'books'. If the antecedent and the different-phrase belong to different types, the result is at best odd, as shown in (29a). The only possible

² Carlson (1987) offers a unified account of all readings of *different*, while Beck (2000) argues that there are two subtypes: (i) exemplifies the reciprocal use of a relational adjective and (ii) a particular use of the comparison operator:

Detmar and Kordula live in different cities.

⁽ii) Every girl read a different book.

interpretation under which (29a) would be acceptable would be one in which the drinking of beer last night was some kind of assignment. Similarly, if the different-phrase and the than-phrase belong to different types, the result is also odd, unless there is some context that forces an interpretation in which a book of origami and a piano are objects of the same type.

- (29) a. ??a different assignment than the beer you drank last night
 - b. ??Lou bought a book of origami. Daniel bought a different piano.

In contrast to *different*, which relates objects of the same type or kind, *else* contrasts objects of different types. *Something else* in (30) cannot mean 'another apple'. *Else* does not exclude the antecedent apple from the set of apples. Rather, it excludes the type or the kind 'apple' from the set of edible types or kinds.³

- (30) John ate an apple and Mary ate something else.
- (31) below also shows that *else* excludes a type, since the antecedent is a predicative indefinite which has a type/kind reading:
- (31) John wants to be a doctor, and Mary wants to be something else.

So it seems that the crucial semantic difference between *else* and *different* is that *else* operates on types whereas *different* operates on tokens, or objects. However, the distinction is not so neat, as *different* does not always express a comparison between tokens of the same kind. Sometimes, the comparison is between two different kinds or types.

(32) Our last car was different from this one.

³ In this we differ from Culicover & Jackendoff (1995). Consider the following example, which in their opinion can be interpreted either as (a) or as (b):

⁽i) John saw a red balloon and Bill saw something else.

a. John saw a red balloon and Bill saw something other than that red balloon.

b. John saw a red balloon and Bill saw something other than a red balloon.

⁽a) claims that one possible interpretation of *something else* is that it expresses distinctness from some previously mentioned individual, while (b) shows that it can also express distinctness of category (or type) from that of some previously mentioned individual.

We disagree with the judgements in (i). In particular, we believe that (i) cannot have the reading in (a), and that (i) cannot be paraphrased with (ii):

⁽ii) John saw a red balloon and Bill saw another red balloon.

(32) could involve a comparison between two objects or tokens of the same type of kind (i.e. the kind 'car'). The kind 'car' is seen as being instantiated or 'realized' (Carlson 1977) by the set of all objects of the kind 'car', and different compares two such objects. Under a second interpretation, (32) could mean that our last car was of a different brand, i.e. type, than this one. In this case, different expresses a comparison between two types or kinds. So in this case, it seems that the semantic distinction between else and different does not hold. Crucially, however, under the second interpretation of (32), where the comparison is between two types, the two types must be subtypes of the same kind. In other words, the kind 'car' is partitioned into several subkinds (corresponding, say, to the various existing brands of cars) and different compares two of these subkinds. In contrast, else never relates two subkinds of the same kind. The antecedent and the referent of else must belong to two different kinds altogether.

To sum up on the differences between *else* and *other/different*: syntactically, *else* differs from *other* and *different* in that only one argument can be overtly expressed within the constituent containing *else*, whereas the constituent containing *other* or *different* can overtly express two arguments, i.e. both the term of comparison and the element subject to comparison. *Else* shares no distributional context with *other*, apart from *another*, but it shares with *different* the 'discourse anaphoric' context. Semantically, *other*, *different*, and *else* express a comparison or a contrast at different levels. *Other* expresses a contrast between objects or tokens of the same type / kind; *different* may express a comparison between objects of the same type / kind or between subkinds of the same kind; and finally *else* always contrasts objects of different types / kinds.

To sum up on the semantic contribution of *else* to the compositional interpretation of *X-else*, *else* performs a set theoretical operation of relative complementation: *else* picks up the relative 'complement' of the antecedent from the domain defined by *X*. Moreover, our discussion above regarding the differences between *else* and *other / different* also showed that *else* is sensitive to types, rather than to tokens. In other words, *else* 'sees' the antecedent as a member of a type / kind, not simply as an individual.

2.1.2.3 *Analysis*. So far, we have proposed that (i) *else* performs an operation of exclusion, i.e. it excludes the antecedent from the domain defined by *X* and it picks up its complement set; and (ii) the antecedent and the referent of *else* must belong to two different kinds or types.

Following Carlson (1977), we represent the notion that an object is an object of some kind by a relation R (for 'realize'). Thus, the set of objects that are of the kind 'cats' is given in (33):

(33)
$$\forall x^0 [R(x^0, c^K)],$$

where x^0 ranges over entities from the domain of objects c^K denotes the kind 'cats'

Our previous discussion showed that *else* does not just exclude the individual object denoted by the antecedent but the whole kind instantiated by the antecedent. In other words, what *else* excludes is (34), i.e. the set of all objects that realize the kind instantiated by the antecedent.

(34) $\forall x^0 R(x^0, ant^K)$, where ant^K is the kind instantiated by the antecedent.

The way in which the kind which is instantiated by the antecedent is expressed is either by using 'a proper name for kinds' (a bare plural in English), as in (35a), or by indicating the property that defines that kind, as in (35b).

- (35) a. John likes books but Mary likes something else.
 - b. John wants a book but Mary wants something else.

Remember that *else* operates on the set defined by X in X-*else*. In *somebody else*, it operates on the set of humans, in *something else* on the set of inanimate things, in *somewhere else*, on the set of places, etc. The set defined by X can be seen as the set of all objects that realize the kind instantiated by X, as in (36). In *somebody else*, for instance, X denotes the set of all objects that realize the kind 'human', in *something else*, X denotes the set of all objects that realize the kind 'inanimate thing', etc.

(36)
$$R(x^0, X^K)$$
 where X^K is the kind instantiated by X in X -else.

Thus, else ultimately contrasts two sets of objects that realize two different kinds.

(37)
$$\lambda x^0[R(x^0, X^K) \& \forall y^0[R(y^0, ant^K) \Rightarrow x^0 \neq y^0]],$$
 where X^K is the kind instantiated by X in X -else and ant^K is the kind instantiated by the antecedent

What (37) says is that *else* denotes the set of individuals x that realize the kind instantiated by X and which are different from the individuals that realize the kind instantiated by the antecedent. Thus, *else* in *somebody else* denotes the set in (38), i.e. the set of individuals that are humans and that are different from any individual that realizes the kind instantiated by the antecedent:

(38)
$$\lambda x^0 [R(x^0, \text{humans}^K) \& \forall y^0 [R(y^0, \text{ant}^K) \Rightarrow x^0 \neq y^0]]$$

The quantifier introduced by X will eventually bind a variable in a set like (37).

Let us see how this works. In (39a), *else* operates on the set defined by X, i.e. the set of objects that realize the kind 'thing', and it excludes from this set the set of objects that realize the kind 'apple' instantiated by the antecedent. This is shown in (39b):

- (39) a. John ate an apple and Mary ate something else.
 - b. $\exists x^0 ([R(x^0, inanimate^K) \& \forall y^0 [R(y^0, apple^K) \Rightarrow x^0 \neq y^0]] \& ate (Mary, x^0))$

What (39b) says is that there is an object of type 'inanimate' which is different from all objects of the type 'apple', such that this object was eaten by Mary. Similarly, in (40a), *else* operates on the set of inanimate objects defined by X and it excludes from it the set of objects that realize the kind 'green apple'. This is shown in (40b):

(40) a. John ate a green apple and Mary ate something else. b. $\exists x^0 ([R(x^0, inanimate^K) \& \forall y^0 [R(y^0, green^K \& apple^K) \Rightarrow x^0 \neq y^0]]$ & ate (Mary, x^0))

Notice that in this case the kind realized by the antecedent is defined by two properties ('greenness' and 'applehood'), rather than just one. This, however, creates no problems for our analysis, since a kind defined by two properties will still be realized by a set of objects, i.e. the objects that lie at the intersection of the set denoted by one of the properties with the set denoted by the other property. The denotation of kinds could thus be generally regarded as a set of properties associated with the respective kind such that whatever realizes that kind has all those properties and whatever has all those properties realizes that kind.

By definition, the objects that realize a kind defined by several properties have all of these properties. In (40a), the set of objects that realize the kind 'green apple' must have both the property of being green and that of being

apples (they represent the intersection of the set of objects that are green with the set of objects that are apples). Thus, in (40) *else* excludes only objects that realize the kind instantianted by the antecedent, i.e. only those objects which are at the same time green and apples, which means that the denotation of *something else* will still include objects that are green but are not apples, as well as objects that are apples but not green.

- (41) a. John ate a green apple and Mary ate something else, i.e. a red apple.
 - b. John ate a green apple and Mary ate something else, i.e. a green banana.

The same type of analysis can be extended to cases in which the antecedent is defined by more than two properties, including proper names. Proper names could thus be analysed as a set of properties associated with the respective individual. The only difference between cases in which the antecedent is a proper name and cases like (40) is that the set of objects that have all the relevant properties is reduced to a unique object. Proper names could thus be regarded as kinds of their own, i.e. as the unique member of the set that represents the intersection of the sets denoted by the property that define the respective bearer of the proper name.

A third type of example that we want to discuss is one in which neither the referent of the antecedent, nor the referent of *X-else*, can be an object level individual.

(42) John wants to be a doctor, and Mary wants to be something else.

In (42), neither *a doctor*, nor *something else* denote an object level individual. This is a well known fact about predicative nominal constituents that occur in a post copular position. *A doctor* in (42) does not denote an individual, but it expresses a property. *A doctor* in (42) denotes the set of all objects that have the property of being a doctor. The same applies to *something else*, which is also used predicatively and which occurs in a similar syntactic position. If this is true, then the only difference between (42) and the examples we discussed above is that in (42) the antecedent, i.e. *a doctor*, cannot be said to refer to an object that realizes the kind 'doctor', but instead refers to the kind 'doctor' itself. The denotation of *X-else* in (42) will thus be similar to its denotation in the examples we have already discussed:

(43)
$$\exists x^0 ([R(x^0, \text{things}) \& \forall y^0 [R(y^0, \text{doctor}^K) \Rightarrow x^0 \neq y^0]])$$

To sum up, in this section we have shown that unlike *other* and *different*, *else* expresses a contrast between objects of different types / kinds, and we have proposed an analysis that relies on the relation of 'realization' between objects and kinds. We have discussed several examples and have shown how our analysis works.

Before we end this section, we would like to discuss an apparently problematic example:

(44) There is *a man* in the waiting room, and *somebody else* is on the phone for you.

Somebody else in (44) can mean another man, in other words, it does not exclude the set of all objects that realize the kind 'man', contra our analysis. Notice, however, that the interpretation of else in (44) is not exclusive, but additive. A suitable paraphrase for (44) would be (45):

(45) There is *a man* in the waiting room, and *in addition to that*, there is a man on the phone for you.

All we have said so far applies to 'exclusive' *else*. As mentioned above, the crucial semantic difference between 'exclusive' *else* and 'additive' *else* is that the former compares types or kinds, whereas the latter contrasts objects.

3. Syntax: partitivity

Our analysis above relies on the assumption that *else* operates on a set of objects defined by X. In other words, we assumed that *else* has access to a set of objects introduced by an element that precedes *else* in linear order. The way we implemented this was to assume something like (46):

(46) X else

 $\exists x \text{ else } x$

In this section we raise the following question: How come the two constituent parts of *X-else* share the same domain or variable? We propose that this analysis is supported by syntactic facts. More specifically, we propose that *X-else* has the syntax of a partitive construction. Evidence for this analysis is found in Romance languages like French, Italian, and Romanian, as well as in other languages, like Dutch or Finnish.

The crosslinguistic data in (47) shows that in French, *X else* constituents must contain an overt partitive P, whereas in Swedish and Dutch *else* is marked for Genitive, i.e. Partitive, Case.⁴

(47) a.	quelqu'un	d'	autre	(French)
	someone	of	else	
	quoi	d	autre	
	what	of	else	
b.	någon		annars	(Swedish)
	iemand		anders	(Dutch)
	someone		else- GENITIVE	

⁴ A reviewer points out that the French *de* in (47a) seems to be the same as the pre-adjectival *de* surfacing in certain constructions without the lexical head noun and since it is not clear that preadjectival *de* in general should be analysed as a partitive *de*, the same kind of doubt could be cast on (47a). Notice, however, that pre-adjectival *de* can be of two types (Jones 1996): (a) pre-adjectival *de* that appears in constructions like (i), where both the adjective and the modified noun are overt. This type only shows up with prenominal adjectives:

- (i) a. De nouveaux livres sont au programme.

 new books are on-the program
 - b. *De livres nouveaux sont au programme.
 - c. *Livres de nouveaux sont au programme.
- (b) pre-adjectival de that appears in constructions like (ii), in which the adjective appears without the lexical head noun:
- (ii) a. Luc voit quelque chose d'intéressant. 'Luc sees something interesting.'
 - b. Paul connaît quelqu'un de très intelligent. 'Paul knows someone very intelligent.'
 - c. Je ne fais rien de spécial. 'I don't do anything special.'

This latter type of de shares properties with partitive de. First, it is always preceded by an indefinite element, which is typical for partitive constructions. Second, the constructions in (ii) could easily be assigned a partitive interpretation. *Quelque-chose d'intéressant* denotes some object x that belongs to the set of interesting objects; *quelqu'un de très intelligent* denotes some person x that belongs to the set of very intelligent persons.

Even though d'autre can appear in both types of constructions, we think that de in $quelqu'un\ d$ 'autre is like the latter type of de, i.e. the one that shares properties with the partitive de.

- (iii) a. Jules a dansé avec d'autres filles. 'Jules danced with other girls.' (Jones 1996)
- b. Luc n'a vu personne d'autre. 'Luc did not see anyone else.' (Jones 1996) Just like de in (ii) above, d'autre is preceded by indefinite expressions, as shown in (iv), and it could easily be assigned a partitive interpretation. Quelqu'un d'autre denotes some person x that belongs to the set of persons that are in the complement set of some presiously mentioned individual.
- (iv) personne / rien / quelqu'un d'autre 'noone / nothing / someone else'

Another argument is the fact that *X* in *X-else* must be indefinite, as observed above. In languages which have a rich case morphology, indefiniteness or 'quantitative indeterminacy' are related to Partitive case (Kiparsky 1998, Giusti 1992). Kiparsky (1998) points out that an object NP with a quantitatively indeterminate denotation in Finnish gets assigned Partitive case. In contrast, if the interpretation is that of a delimited set (of known or unknown cardinality), the NP gets Accusative case:

- (48) a. Anu-lla on loistava-t oppilaa-t (Finnish)
 Anu.ADESS be.3SG brilliant.PL.ACC student.PL.ACC
 "Anu has brilliant students."
 - b. *Anu-lla on loistav-i-a oppila-i-ta*Anu.ADESS be.3SG brilliant.PL.<u>PART</u> student.PL.<u>PART</u>
 "Anu has (some) brilliant students."

(47) and (48) suggest that X takes a partitive PP complement not only in French, where the partitive preposition is overt, but also in Swedish, Dutch, and English, where the partitive preposition is covert:

(49) a. Q
$$[PP P [NP else [NP]]]$$

b. $quelqu'un$ d' $autre^5$ e (French)
c. $n\mathring{a}gon$ $[PP Ø [NP annars.GEN]] e$ (Swedish)
d. $iemand$ $[PP Ø [NP anders.GEN]] e$ (Dutch)
e. someone Ø else e

Before we provide further evidence for this partitive structure, let us notice that the picture in (49) is not complete. Giusti (1992), Cardinaletti & Giusti (1990), Zamparelli (1998), and others, have convincingly shown that a partitive structure contains not only a partitive PP, but also an additional NP which precedes the PP. There are cases in which both NPs are spelled out, as in (50b), and cases in which only one of them is overt, as in (50c,d).

(50) a. Q
$$[_{NP}NP_1][_{PP}P[_{NP} NP_2]$$

⁵ We analyzed *autre* as a prenominal modifier in (49b). However, as pointed out by Petra Sleeman (p.c.), *autre* could also be postnominal, as in (i):

⁽i) un homme autre que son mari 'another man than her husband' Since it is not clear whether in (49) we are dealing with the prenominal or the postnominal autre, an alternative analysis is possible, as in (ii):

b. molti libri dei libri que mi hai dato (Giusti 1992)

multe carti dintre cartile pe care mi le-ai dat many books of-the books which to-me have. 2sG given "many of the books you have given to me"

molti o dei lihri (Italian) c. multe dintre cãrti (Romanian e. livres (French) beaucoup des "many of the books"

d. molti libri Ø e
multe cãrti Ø e
beaucoup de livres Ø e
"many books"

The two NPs in (50a) above must be coindexed, as shown by (51) below.

(51) *molti ragazzi dei libri que mi hai dato (Italian)

*multi bãieti dintre cãrtile pe care mi le-ai dat (Romanian)

many boys of-the books which to-me have. 2SG given

If our proposal above that X-else elements have a partitive syntax is on the right track, then (50a) should also be the analysis for X-else, not only for examples like (50b-d). Since the two NPs in (52a) are coindexed, partitivity can account for why the two constituent parts of X-else share the same domain or variable, as in (52c):

(52) a. $[Q NP_1]_i [PP P [NP else [NP_2]_i]]$ b. some-one \emptyset else e c. $\exists x_i$ else x_1

Let us now provide further evidence for a partitive analysis of *X-else*. In particular, we are looking for evidence for the existence of two NPs in *X-else* constituents, as indicated in (52a). If such evidence could be found, then a partitive analysis would be more credible.

First, some languages can overtly express either NP_1 , as in (53a), or the NP_2 , as in (53b).

(53) a. where else

dove altro (Italian)

b. elsewhere

altrove (Italian)

c. *where elsewhere

*dove altrove (Italian)

Even though English and Italian can express either NP₁ or NP₂, these languages cannot express both NP₁ and NP₂ at the same time (see 53c). However, the syntax we propose in (52) suggests that this should be possible, and that we should expect to find languages in which both NP₁ and NP₂ are overt. Such evidence can be found in Romanian. In Romanian, the wh-paradigm is parallelled by a paradigm which is morphologically derived from the wh-one by means of the suffix *-va*. This second paradigm includes indefinite bare quantifiers, and in view of its transparent morphology, it is reasonable to assume that the indefiniteness is actually the contribution of *-va*. Given the analysis of indefinites assumed by Heim (1982) and Diesing (1992), it is reasonable to assume that *-va* introduces a variable.

(54) wh bare quantifiers (wh+-va)

cine cine-va

who who-va "somebody"

unde unde-va

where where-va "somewhere"

când când-va

when when-va "sometime"

care care-va

which which-va "somebody"

ce ce-va

what what-va "something"

cum cum-va

how how-va "somehow"

X-else expressions in Romanian always include a wh-va word which follows *else*. Now, this wh-va word can either show up by itself, as in (55), or, most interestingly, can occur in addition to an independent wh-word or bare quantifier which precedes *else*, as in (56).

(55) *alt-cine-va; alt-ce-va; alt-unde-va; alt-cum-va;* else-who-*va;* else-what-*va;* else where-*va;* else-how-*va;* "somebody else"; "something else"; "somewhere else"; "somehow else";

alt-când-va else-when-va "sometime else"

(56) a. *cine* alt-cine-va who else-who-va

who else-somebody "who else"

ce alt-ce-va what else-what-va

what else-something "what else"

când alt-când-va when else-when-va

when else-sometime "when else"

 $\begin{array}{ll} \textit{cum} & \textit{alt-cum-va} \\ \underline{\text{how}} & \text{else-}\underline{\text{how}}\text{-}\textit{va} \end{array}$

how else-somehow "how else"

unde altundeva where else-where-va

where else-somewhere "where else"

b. *nimeni altcineva* nobody else-who-*va*

nobody else-somebody "nobody else"

nimic altceva nothing else-what-va

nothing else-something "nothing else"

(56) thus shows that the variable expressed by -va words duplicates the variable implicit in the wh-word or in the bare quantifier. This is exactly the type of evidence we were looking for, i.e. evidence that two NPs are involved in X-else.

To sum up on this section, we have shown that our semantic analysis of *else* as involving access of *else* to the variable introduced by X is supported by a syntactic analysis of X-*else* as a partitive. A partitive involves two NPs which are coindexed, as in (52a), and this supports a semantic analysis in which a replica of the variable introduced by X is available for *else* to operate on.

4. Conclusions

In this paper we have shown that the interpretation of X-else constituents is compositional, and that they contain an anaphoric element of the standard type. This anaphoric element is located both within X and within else at the same time, and else ranges over the same set as that defined by X. This analysis is

supported by the partitive syntax that we proposed for X-else. The operation that else performs on the set defined by the variable introduced by X is one of relative complementation: else excludes the set of objects that realize the type or kind instantiated by the antecedent from the set of objects defined by X.

We have also compared *else* to similar 'exclusion' predicates, such as *other* and *different*. We have shown that *else* differs from *other* and *different* in that only one argument can be overtly expressed within the constituent containing *else*, whereas the constituent containing *other* or *different* can overtly express two arguments, i.e. both the term of comparison and the element subject to comparison. *Else* shares no distributional context with *other* (apart from *another*), but it shares with *different* the 'discourse anaphoric' context. Semantically, *other*, *different*, and *else* express a comparison or a contrast at different levels. *Other* expresses a contrast between objects or tokens of the same kind; *different* may express a comparison between objects of the same kind or between subkinds of the same kind; and finally *else* always contrasts objects of different kinds.

References

- Beck, Sigrid. 2000. "The Semantics of *Different*: Comparative Operator and Relational Adjective". *Linguistics and Philosophy* 23.101-139.
- Cardinaletti, Anna & Giuliana Giusti. 1990. "Partitive *Ne* and the QP Hypothesis". Ms., University of Venice.
- Carlson, Greg. 1977. "A Unified Analysis of the English Bare Plural". Linguistics and Philosophy 3.413-456.
- Cinque, Guglielmo. 1986. "Bare Quantifiers, Quantified NPs and the Notion of Operator at S-Structure". *Rivista di Grammatica Generativa* 11.33-63.
- Culicover, Peter & Ray Jackendoff. 1995. "Something Else for Binding Theory". Linguistic Inquiry 26:2.249-277.
- Diesing, Molly. 1992. Indefinites. Cambridge, Mass.: MIT Press.
- Dobrovie-Sorin, Carmen. 1994. *The Syntax of Romanian*. Berlin: Mouton de Gruyter.
- Giusti, Giuliana. 1992. La Sintassi dei Sintagmi Nominali Quantificati: Uno Studio Comparativo. PhD Dissertation, University of Venice.
- Hand, Michael. 1987. "Other and Else: Restrictions on Quantifier Domains in Game Theoretical Semantics". Notre Dame Journal of Formal Logic 28:3. 423-430.

- Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD Dissertation, U.Mass.
- Jones, Michael Allan. 1996. *Foundations of French Syntax*. Cambridge: Cambridge University Press.
- Keenan, Edward. 1988. "Complex Anaphors and Bind Alpha". *Papers from the* 24th Annual Regional Meeting of the Chicago Linguistic Society ed. by Diane Brentari, Gary Larson & Lynn Macleod, 216-232. Chicago: Chicago Linguistic Society.
- Kiparsky, Paul. 1998. "Partitive Case and Aspect". *The Projection of Arguments: Lexical and Compositional Factors* ed. by Miriam Butt & Wilhelm Geuder, 265-307. Stanford: CSLI Publications.
- Pesetsky, David. 1987. "WH-in-situ: Movement and Unselective Binding". *The Linguistic Representation of (In)definiteness*, ed. by Eric Reuland & A.G.B. ter Meulen, 98-129. Cambridge, Mass.: MIT Press.
- Safir, Ken. 1991. "Implied Non-coreference and the Pattern of Anaphora". Linguistics & Philosophy 15.1-52.
- Tovena, Lucia & Marleen van Peteghem. 2003. "Facets of *Different* in French: différent and autre". Empirical Issues in Formal Syntax and Semantics 4 ed. by Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr & Francis Corblin, 63-79. Paris: Presses Universitaires de Paris-Sorbonne.
- Zamparelli, Roberto. 1998. "A theory of Kinds, Partitives and of/z Possessives". *Possessors, Predicates and Movement in the Determiner Phrase* ed. by Artemis Alexiadou & Chris Wilder, 259-301. Amsterdam & Philadelphia: John Benjamins.

LEARNABILITY ORDER IN THE FRENCH PRONOMINAL SYSTEM*

JACQUELINE VAN KAMPEN Utrecht University

1. The claim

The present paper will reconsider the acquisition of free anaphors in French. Hamann et al. (1996) have pointed out that object clitics do not appear as easily in French child language as subject clitics. When the amount of object clitics is compared with the total amount of object phrases, there is delayed rise in the use of object clitics. The phenomenon is less marked for subject clitics. Further study by Jakubowicz et al. (1998) was directed at SLI children. The acquisition delay for object clitics was clearly reaffirmed. Since subject/object asymmetries have been a concern for some time in the theory of syntax, Hamann et al.'s original observations led to proposals that derive the acquisition delay from the argument licensing of clitics. French subject and object clitics do not form a homogeneous class. Kayne (1975:86f) assumes the subject pronoun cliticized to the head Vfin at at PF, whereas object clitics move from the canonical object position to a position in front of the VP. Jakubowicz et al. (1998) argue that the subject/object asymmetry in acquisition follows from that syntactic difference. The subject clitic would not only express the subject's D-features (person/number/gender), but also the verbs Ifeature 'finite'. This additional I-function would enhance the learnability of the subject and so explain the comparitively delayed rise of object clitics.

^{*} I would like to thank for indispensable discussions and disagreements: Sergio Baauw, Arnold Evers, Aafke Hulk, Celia Jakubowicz, Alain Rouveret, Marlies van der Velde and two anonymous reviewers, as well as critical audiences in Paris and Groningen.

¹ The free anaphors are (full/weak/clitic) pronouns as opposed to bound anaphors (reflexives).

² The identification of the subject clitic with the I-marking of finiteness makes sense in the specific analysis of Jakubowicz et al. They have no clitic movement, but consider finite verb and associated clitics as the result of a merge operation. French would be a pro-drop language, since its verbal paradigm needs support of additional person agreement (cf. Hulk 1986, Nash & Rouveret 1997). For some critical notes on the merger analysis of Jakubowicz et al. and a defence of a movement analysis, see Rouveret (to appear).

Interesting though this proposal is, my own research on the rise of free anaphors in Dutch child language (Van Kampen 1997, 2001) suggests an alternative explanation. It turns out for Dutch that the acquisition graph for all free anaphors coincides with the acquisition graph for systematic D-marking on argument structure (determiners before nouns). The distinction between +/- clitic pronoun or between +/- subject anaphor is irrelevant. I will argue below that almost the same holds for French. The use of free anaphors depends on the previous acquisition of D-marking. There is no acquisition difference between object and subject clitics. The early rise of subject clitics is apparent only. The early subject clitics that spoil the game are *shadow pronouns* supported by a local subject adjunct (*papa_i*, *ili fume une pipe*). The crucial distinction is not subject clitic versus object clitic, but shadow pronoun versus free anaphor. This leads me to the claim that French exhibit the learnability order in (1).

(1) Systematic I-marking and shadow pronouns *precede* Systematic D-marking and free anaphors.

Let me add a few terminological points. I-marking is the marking of a predicate by a factor <+I>. This factor generalizes over a variety of devices {copula, auxiliary, modal, finite morphology}. D-marking is the marking of arguments by a factor <+D>. This factor also generalizes over a variety of devices {article, demonstrative, possessor, quantifier}.

Throughout, I make a sharp distinction between the situational context of a sentence versus its (linguistic) discourse context. The situational context is in principle available to the child, without additional grammatical devices. The discourse context is in my view not accessible to the child until there is the systematic application of D- and I-marking (cf. Schaeffer 1997). These are deictic devices of anaphoric marking. They may relate the sentence to previous utterances.

1.1 Previous studies

The present paper makes use of several distinctions from previous generative acquisition studies. In the present perspective they construe the learnability hierarchy in (1).

³ The term 'shadow pronoun' was used by Perlmutter (1972) and is due to the Arabian grammatical tradition. The term 'resumptive pronoun' is nowadays generally used for the relation between a pronoun and an antecedent in relativization and wh-movement. Some people like to use the term 'shadow pronoun' for the pronoun that is bound by a dislocated topic (Ray Mestrie p.c.).

- (i) Situation-bound child language. Pronouns in child language are at first situation-bound only. Pronouns that refer to arguments in previous utterances appear much later (Lyons 1977/1979, Atkinson 1979, Hickmann 1982).
- (ii) Clitics (French subject/object clitics) refer in a way different from full pronouns. They do not refer deictically and gesture-sustainable in a situation-bound way. They refer to highly presupposed structure that becomes available due to previous discourse. The discourse context-depending referring property of clitics has been used in Avrutin (1994) and Baauw & Cuetos (to appear) for their analysis of the acquisition of binding. I will argue here that this context dependency of clitics explains their position in the learnability hierarchy.
- (iii) Full pronouns Dutch introduces the situation-bound 3^{rd} person preferably by the gesture-sustainable demonstrative pronoun die. The use of the 3^{rd} person pronoun hij/hem ('he/him') is more discourse-oriented. Haegeman (1996) notices that Dutch children at first make an almost unexceptional use of the (situation-bound) demonstrative. I will show by longitudinal graphs that the acquisition of the 3^{rd} person pronouns hij/hem coincides with the acquisition of D-marking.
- (iv) *D-marking on nouns* (determiners in French and Dutch) and *free anaphors* (clitics in French, strong/weak pronouns in Dutch) represent the same syntactic category <+D> as argued by Postal (1966). Within that line, Heim (1982) has proposed that the <+D> elements are the anchoring points of discourse grammar. She lists the D-marked arguments in a stack of elements that will further count as presupposed in the immediate discourse. The acquisition of determiners and 3rd person pronouns can be seen as the acquisition of the Heim-stack. The parallel acquisition has been argued for by Hoekstra & Hyams (1995) and Schaeffer (1997). My graphs show that the device of discourse-bound 'free' anaphors does not get into operation, unless a reference stack as postulated by Heim (1982) has been build up by systematic D-marking.
- (v) French subject dislocations. In adult French, subject DPs are predominantly dislocated topics doubled by a subject clitic (Lambrecht 1981, Ashby 1988). Ferdinand (1996) and De Cat (2002) argue that French child language uses the dislocated topic + subject clitic from early on. The subject clitic is not a free anaphor, it is locally bound to the topic noun, as a 'shadow pronoun'. It is the dislocated topic that is the referent in these constructions, not the subject clitic (cf. Hickmann & Hendriks 1999: ftn. 6).

1.2 Dutch data that reveal an acquisition order

The Dutch child data reveal four marvellously ordered acquisition facts of subjects and objects. They are listed in (2) and (3). The corresponding longitudinal graphs reflect a rise to the adult level of I-marking and D-marking.

(2) Parallel acquisition of:

- a. <+fin> marked predication
- b. subjects with inherent (non D-marked) reference

(3) Parallel acquisition of:

- a. systematic D-marking of arguments
- b. discourse anaphora

Phenomenon (2a)/(2b) cannot be a mere coincidence. They represent the acquisition of the 'Extended (I-) Projection Principle' (EPP: Rothstein 1983), the subject obligation for standard sentences. The EPP requires a subject in Spec,I for grammatically marked predicates IP. Phenomena (3a)/(3b) are no coincidence either. They can be seen as the acquisition of discourse grammar as based on a stack of referential points <D> (Heim 1982, cf. Hoekstra & Hyams 1995).

The acquisition of the EPP in Dutch child language (2) precedes the acquisition of discourse reference (3). The longitudinal graph for I-marking precedes the one for D-marking.⁴ See the graphs in section 2.1.

(4) Acquisition order in Dutch child language (Van Kampen 2001) I-marking (2) precedes D-marking (3)

Although both the EPP and the free anaphors are based on ϕ -chains, i.e. on the D-features of person/number/gender/definiteness, the EPP is a matter of *sentence grammar*, whereas the free anaphors are a matter of *discourse grammar*.

Arguments need to be licensed in syntax. The licensing devices (determiners and/or morphological case) are language specific and have to be acquired. The aim of the grammatical licensing is that each argument relates a theta role to a reference marking. See for this relation Williams (1994), Rouveret (2002) and for its acquisitional relevance Van Kampen (1997:130f).

⁴ The present analysis considers I-marking (<+/-finite> oppositions) rather than T(ense)-marking (<+/- past> oppositions). The Tense opposition appears at a later stage in acquisition (Evers & Van Kampen 2001).

The acquisition order in (4) implies that the subject in early child language will lack D-marking. It will be based on nouns and pronouns with situation-bound reference, like proper names and demonstratives. This will be supported empirically in section 4.2.

The initial evidence for the basic acquisition order in (4) follows from my research of Dutch (Van Kampen 1997, 2001). It now raises the questions in (5).

- (5) a. Do the correlations (2a/b) and (3a/b) and the acquisition order between them in (4) hold for French child language as well?
 - b. How can we explain the delay of object pronominalization in French?

I will answer these questions in the following way.

- (6) a. I-marking precedes D-marking holds for French child language as well.
 - b. (i) The acquisition of free anaphors in French makes no distinction between subjects and objects. (ii) Free anaphors appear right after systematic D-marking (iii) Subject cli tics in child French that appear before systematic D-marking are *shadow pronouns* doubled by a topic.

In order to get empirical support for the statements in (6), I have to partly reinterpret the acquisition data in Hamann et al. (1996), Jakubowicz et al. (1998). I will proceed as follows.

The next section will review the acquisition of Dutch I-marking and D-marking and the simultaneous acquisition of D-marking and free anaphors. Section 3 will consider the almost simultaneous acquisition of systematic D-marking and free anaphors in French child language. Section 4 will argue that systematic I-marking precedes systematic D-marking in French as well. Section 5 will conclude that the I-marking of predicates and the D-marking of arguments is the watershed between proto-grammar and real grammar. It interprets the acquisition of I-marking and D-marking as adding deictic markers that enable the sentence to relate to the discourse rather than to the immediate situation.

All French child data are derived from the Grégoire files (Champaud corpus) and all Dutch data from the Sarah files (Kampen corpus), both available in CHILDES.

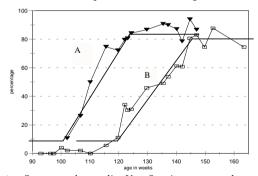
2. Pronoun acquisition in Dutch

2.1 Systematic I-marking and reference marking of subjects

The reference marking of subjects for the EPP in Dutch child language is at first a matter of nouns and pronouns with situation-bound deictic reference, like demonstratives and proper names for persons, animals, places and toys (cf. (2b)). The appearance of these subjects coincides with the rise of I-marked predicates ((2a/b)). However, apparent cases of null subjects in finite sentences in the language of the child seem to contradict this point of view. The same holds for early French child language. I will deal with this issue in section 4.

Simple counts for Dutch show that the (2a/b) I-marking of the subject precedes (3a/b) D-marking of argument structure (cf. (4)). The acquisition order (I-marking precedes D-marking) was argued for in Van Kampen (2001) by means of the graphs in (7). Graph A establishes the rise of Sarah's use of <+I> {copula, auxiliary, modal, finite morphology} in predications. Since Dutch is a V-2nd language, the <+fin> verb appears in C^o-position. Graph B establishes the rise of Sarah's use of <+D> {article, demonstrative, possessor, quantifier} before nouns.⁵ In Dutch, the use of a D^o is obligatory with singular count nouns and definite plural nouns only. I counted the +/- oppositions in these contexts.⁶

(7) Sarah (Dutch): Acquisition of I-marking and D-marking (Van Kampen 2001)



Graph A: Sentences that realize $V < +fin > in \ge two$ -word utterances Graph B: Noun Phrases that realize D^o in obligatory contexts in $\ge two$ -word utterances

⁵ The graphs for D-marking and I-marking are based on ≥ two-word utterances only. One-word utterances may invite adult over-interpretation (Evers & Van Kampen 2001).

⁶ In French, the use of an explicit D^o is obligatory with all nouns, except for certain predicative uses (*il est matelot*). Moreover, the French D-system has phi-oppositions of <+/- gender>, <+/- number>. The difference between the French and Dutch D-system may affect starting point and speed of the acquisition graph, but that is irrelevant for the present argumentation.

The acquisition lines in (7) are parallel (20 weeks). The straight lines have been added to the graphs. They underline the fact that the graphs represent a kind of parameter setting (Evers & Van Kampen 2001). The graphs in (7) nicely show that D-marking follows I-marking In Dutch.

This leaves us with the last point, the simultaneous acquisition of discourse anaphora (3b) and the D-marking of argument structure (3a). Section 2.2 for Dutch and section 3 for French child language present simple quantifications that show the dependency.⁷

2.2 Systematic D-marking and discourse anaphora

Before she uses D-marking systematically, Sarah uses demonstrative pronouns (*die/dat/deze/dit* 'that/this'), but no 3rd pers. pronouns (strong pronouns *hij/zij/hem/haar/het* or weak ones *ie/ze/'m/d'r/'t* 'he/she/him/her/it'). The primary selection of the demonstrative was observed by Haegeman (1996) for the Dutch child Hein.

The early demonstrative *die* is situation-bound.⁸ After I-marking, the early pronoun *die* appears almost exclusively in the Spec,C position of Dutch. This can be confidently asserted, since Dutch is a V-2nd language. The I-marking of root predicates by <+fin> coincides with the C-marking of root predications by some <+fin>. No matter how one prefers to analyze this fickle coincidence, the finite verb of root sentences appears in the C^o-position. The Spec,C position is thereby identified. The early pronoun *die* is a demonstrative ('that one'). It has been characterized by Van Kampen (1997: 92f) as an *A-bar pronoun*. The function of the demonstrative A-bar pronouns versus the personal A-pronouns can be characterized as 'referent highlighted' versus 'referent presupposed'. The highlighted referent of the A-bar pronoun in child language is a perceptually salient and gesture-sustainable topic in the speech situation.⁹ The demonstrative A-bar pronoun opposes to the personal pronoun *hij/hem*

⁷ They generalize a result already pointed out for object clitics in Hamann et al. (1996) and Jakubowicz et al. (1998).

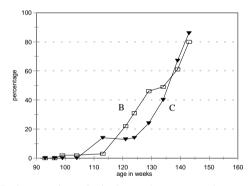
⁸ Dutch highly prefers the use of a general demonstrative *die* for pronominaized topics. It applies to <-animate> as well as to <+animate> objects and subjects. *Die* replaces *dat/dit/deze*, which are specified for (and/or) gender, proximity, number. The gender variant *dat* is disregarded by Sarah. She uses the demonstrative *die* as a situation-bound topic (in non-copula contexts) in more than 90% of the cases.

⁹ Later, in the adult language, the A-bar d-pronoun may also refer to an antecedent in the previously uttered discourse. It announces that its referent constitutes a change of topic w.r.t. the previous sentence. The value 'presupposed' of the A-pronouns means that its referent is known like the <+D, +def> marked nouns, either because they are sufficiently situational salient, or because they have been identified in the previous discourse.

('he/him'). The personal pronoun does not appear naturally in the Spec,C position. It is an A-pronoun and its referent is presupposed.

The child's selection of the A-bar demonstrative pronoun die fits the picture of early child language as situation-bound. The same holds for the child's avoidance of the A-pronouns. These tend to need discourse grammar and discourse grammar is still absent as long as the grammatical device of anaphoric D-marking has not been acquired. By constructing a graph for the rise of 3rd person pronouns as a percentage of the amount of referential expressions, I could measure Sarah's growing ability in the reference anchoring of arguments. The rise of free anaphors with A-pronoun status coincides with systematic D-marking in a striking way (Van Kampen 2001). The ratio DP<+pro> / DP<+/-pro> reaches the ratio of her adult conversation partner in less than a half year. The same period and the same speed is shown for Sarah's systematic D-marking on nouns, the ratio <+D [— NP]> / <+/-D> [-NP]. The parallel graphs B and C in (8) below reveal how Sarah's use of 3rd person pronouns reaches the level of her adult conversation partner at the same week that she acquires systematic D-marking of nouns. From that point on, she consistently applies D-marking for 80%-90%.

(8) Sarah (Dutch): Acquisition of D-marking and pronouns (3rd person)



Each measuring point in the graphs represents 2 consecutive files.

Graph B: Noun Phrases that realize D° in obligatory contexts in \geq two-word utterances

Graph C: Ratio of 3rd person pronouns w.r.t. nouns measured as a percentage of the ratio in the speech of the mother within the same files.

The parallel acquisition of systematic D-marking and discourse anaphora in Dutch suggests the universal points in (9).

- (9) a. Systematic D-marking of argument structure selects and temporarily stacks the passing referential expressions, as proposed by Heim (1982).
 - b. Items in the stack can be referred to by discourse anaphora.
 - c. The device of discourse-bound 'free' anaphors does not get into operation, until such a reference stack has been build up by systematic D-marking.

Systematic D-marking takes place due to D°-labeled articles or due to Case°-labeled case-endings (Van Kampen 1997:130f). Systematic D-marking is the grammatical key to discourse grammar. If so, one expects empirical evidence for an analysis of the French pronominal system along the lines already mentioned under (6). Discourse anaphors are based on systematic D-marking and there is no acquisitional distinction between subject and object in this respect.

3. Child French and D-marking

The points in (6) above imply a far reaching parallel between acquisition hierarchies in Dutch and French. I-marking of the predicate would precede Dmarking of the arguments in both languages (6a). Further, (6b) (i) there is no acquisition difference between subject free anaphors and object free anaphors; (ii) free anaphors appear after systematic D-marking; and (iii) statements to the contrary rely on shadow pronouns, a characteristic of French in general and highly preferred in child French. The present section will discuss this triple claim (6b). The next section will argue for claim (6a) that I-marking precedes D-marking for child French. All empirical arguments below are based upon figures derived from the files of Grégoire. 10 I will first establish Grégoires acquisition point for systematic D-marking, i.e. the point in time from whereon consistently satisfies the adult D-marking for some Subsequently, I will relate systematic D-marking to the acquisitional order for the various types of anaphors {shadow pronouns, A-bar pronouns, subject clitics, object clitics}. The clitic free anaphors (subject clitics as well as object clitics) should appear after systematic D-marking, since they rely on discourse presupposition. The anaphors that appear well before systematic D-marking should be anaphors that belong to early child language. They are gesture

¹⁰ When I wrote this paper, the CHILDES database offered no more than two longitudinal corpora for French (Grégoire and Philippe). The Grégoire corpus was the only one that covered the period for the rise of I- and D-marking. Philippe is already too old at the first session. The same holds for the group 1 children in Jacubowicz & Rigaut (1997). They have a mean age of 2;4.10. Fortunately, a new corpus of 3 French speaking children has been added to CHILDES (York corpus, Plunkett 2002) recently.

sustainable and they can be used directly in the speech situation. These are the A-bar pronouns <+C, +pro, +/-wh>, like the Dutch *wie*, *die* (cf. section 2.2) or the French *qui*, *ça*.¹¹ A third group of anaphors are the shadow pronouns, see the examples in (10) below. They appear in the French topic construction. The topic is dislocated and the argument position is filled in by the shadow pronoun (Lambrecht 1981). The shadow pronoun is in an A-position and bound by a DP-topic in the nearest A-bar position. See for such locality considerations (De Cat 2002: 92).

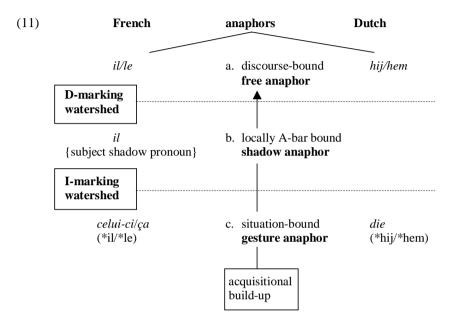
- (10) a. ce_i voisin, il_i est guéri de son rhumatisme "This neighbor, he is recovered from his rheumatism" il_i danse le sirtaki sur la table, ce_i voisin "He dances the sirtaki on the table"
 - b. ce_i voisin, je l_i'ai vu danser le sirtaki sur la table
 "This neighbor, I have seen him dancing on the table" je l_i'ai vu danser, ce_i voisin
 "I have seen him dancing, this neighbor"

Adult French allows topic-dislocation of the subject (10a), as well as topicdislocation of the object (10b). However, child French overwhelmingly prefers the A-bar dislocated topic for subjects only. The following may hold. The root sentences for early child language are mostly without further discourse relations. Such sentences need an explicit topic and this is overwhelmingly done by doubling the subject. Hence, the French child may initially take doubling as the hallmark of subjecthood. The shadow pronouns are not discourse dependent, but sentence dependent. They may appear and do appear before systematic D-marking. Root sentences within discourse will in general not re-affirm the running topic. They do not need the dislocated subject topic. Discourse opens the way for a running topic and a single subject clitic (il dort). Discourse grammar also opens the way for object clitics, both topic-bound (shadow variant) and single (discourse presupposed). The dislocated object topic (je le vois, l'ours) marks a change of the running discourse topic. Hence, it functions in discourse. As a matter of fact, Gregoire's dislocated object topics appear later than his single object clitics (single le > topic-bound le). This is the reverse acquisition order of subjects (topic-bound il > single il). This will be shown in the tables (12) and (13) below.

_

¹¹ Before I-marking, <+C>/A-bar means 'non-argumental position'. The structural consequences will come in after I-marking. See Van Kampen (1997:chap.4/7).

Let this suffice for the three types of anaphors relevant in the present context. Their acquisitional hierarchy for French and Dutch is diagrammed in (11).



Quantitative evidence in the Grégoire files confirms the acquisition scheme in (11). Consider the table in (12). In column (12)d the gray area marks the period immediately following the acquisition of systematic D-marking. Systematic D-marking in (12)d coincides with a dramatic drop in the use of topic-bound *il*, see (12c). There is a fall from 89% to 37/35% of the topic-bound shadow pronoun *il*. The percentages in (12c) give the proportions between subject clitics as shadow pronouns and subject clitics as free anaphors. One might even argue that the rare, and only, occurrences of *elle* before 2;3.0 is stressable and not necessarily a clitic and is better brought into column (12a). Be this as it may, the quantitative jump of D-marking at 2;5.1 in (12)d shows that the distinction *il*/free anaphor and *il*/shadow pronoun is justified for acquisition.

(12) *Grégoire* (French): Rise of D-marking and subject clitics (3rd person)

	a. single	b. single	c. subject clitic	d.
	subj. noun	subj. clitic	+ topic noun	D-
			(in % w.r.t	marking
age	N prop N	il elle	single clitic)	
1;9.18-28	0 0	0 1	8 89%	7%
1;10.20	0 0	0 2	7 78%	6%
1;11.22	0 0	0 2	7 78%	3%
2;0.5	1 1	0/2* 0	19 95%	14%
2;1.25	1 1	0 0	3 **	53%
2;3.0	0 2	2 4	8 61%	60%
2;5.1	0 1	19 0	11 37%	97%
2;5.13-27	4 6	66 28	51 35%	100%

D-marking acquired

The other indication of the relation between discourse anaphors and systematic D-marking is the first appearance of object clitics. Objects are not used as topics in child French as long as the 'topic noun + shadow pronoun' appears to be the hallmark of subject-hood. See column (12a) for the marginal presence of subject nouns in Spec,I. Object clitics are unambiguous free anaphors in child French. They cannot be used to refer directly in the speech-situation, and they have no function yet as shadow pronoun. This leads to the expectation that object clitics will appear just after and not before the systematic D-marking of argument structure. Consider now the figures for object-clitics in (13).

(13) *Grégoire*: Rise of D-marking and appearance of object clitics (3rd person)

age	in weeks	a. D-marking		ct cliic son <i>le/la</i> + topic noun
1;9.18-28	89-91	7%	0	0
1;10.20	94	6%	0	0
1;11.22	98	3%	0	0
2;0.5	105	14%	0	0
2;1.25	112	53%	0	0
2;3.0	117	60%	0	0
2;5.1	125	97%	9	0
2;5.13-2;5.27	127-129	100%	10	4

D-marking acquired

 \downarrow

The object clitic appears in the Grégoire files for the first time at 2;5.1, when D-marking is acquired (cf. Van der Velde et al. 2002: fig.1). At that moment, clitics may function as free anaphors. That is the way Grégoire begins to use

^{*} ∂ fait? (imit. of 'qu'est-ce qu'il fait?') ** too small numbers

them, as 'bare' discourse anaphors not doubled by a topic noun. See the gray area in (13b).

There is a difference between the acquisition of free anaphors in Dutch and French. The rise of 3rd person free anaphors in Sarah's conversations is simultaneous to the rise of D-marking, see (7). Grégoire's acquisition of clitic arguments, whether subject or object, is not simultaneous with D-marking, but follows systematic D-marking, see (12)/(13). The acquisition difference between the French clitic free anaphors and the Dutch non-clitic free anaphors is a side issue in the present context. It is plausibly related to the further grammaticalization of clitics in French, as compared to the Dutch non-clitic pronouns. The Dutch positions for subject and object are not sensitive to the +/- anaphor status. That is different for French. Both subject and object clitic imply the acquisition of a different construction in addition to the pronominalization.

Systematic D-marking of the argument structure is a watershed in language acquisition anyway.

- (14) a. Theta-marking gets tied up with sets of referential slots.
 - b. D-marked arguments become accessible as antecedents of discourse anaphors.

It is attractive to speculate that systematic D-marking and discourse anaphors mark the point where language becomes "human" in the sense of being discourse-related and hence situation-free.

4. Child French and I-marking

4.1 I-marking precedes D-marking

I now return to the claim (6a) I-marking precedes D-marking. This acquisition order is shown in (15). Grégoire applies systematic I-marking half a year earlier than systematic D-marking. I-marking is almost instantaneous (>80% at 1;10.20).

age in weeks		a. I-marking	b. D-marking
1;9.18-28	89-91	67%	7%
1;10.20	94	80%	6%
1;11.22	98	83%	3%
2;0.5	105	79%	14 %
2;1.25	112	82%	53%
2;3.0	117	88%	60%
2:5.1	125	98%	97%

(15) Grégoire: (French): Acquisition of I-marking and D-marking

The time gap between the acquisition of I-marking and D-marking has also been observed for Dutch, cf. (7). It can be maintained that the acquisition of <+fin/I°> marked predication manifests the acquisition of the EPP, if the obligatory presence of the subject is relativized as in (16).

b. mode-implied subjects

c. discourse D-marked subjects

The <+D> marked subjects (16c) are the regular case in adult language, but not in early child language before systematic D-marking. The situation-bound subjects (16a) (proper names and demonstratives) and the mode-implied subjects (16b) characterize early child language. They appear only marginally in adult language.

4.2 Situation-bound subjects

Before the acquisition of D-marking, child language realizes the EPP in situation-bound contexts when the subject is the topic of the sentence ((16a)). Child Dutch mainly applies its d-pronoun *die* and nouns as quasi proper names.

French child language at first realized the <+fin> subject by the demonstrative ca (18a) or by a construction of a dislocated topic + shadow pronoun (18b). The shadow pronouns appear before systematic D-marking.

It must be noticed that in root infinitive subjects need not appear at all, not even situation-bound. Root infinitives may have no more than topic announcers (cf. Hoekstra and Hyams 1995).

- (19) a. rangE tout seul, Grégoire (tidied/tidy up all alone, G.) (G. 2;1.25) (as far as G. is concerned, he 'tidy' up everything by himself)
 - b. *die niet lachen* (that (one) not laugh) (S. 2;0.17) (as far as that one is concerned, he is not laughing)

The stress pattern in type (19b) betrays the topic status. Its subject status is dubious and probably an adult overinterpretation. The topic status in child French is clear from the fact that (the rare) non-finite sentences in early child French appear almost exclusively with a right-dislocated topic as in (19a). Root infinitives like *je mettre lunettes* do not, or hardly, appear (see Ferdinand 1996:167 for some exceptions). The pronouns *je*, *tu*, *il* do not appear before I-marking, since they fit the Spec,I only (Pierce 1989, Hoekstra & Hyams 1995). The present perspective is that subjects become obligatory by the EPP after the grammaticalization of predication by I-marking (Van Kampen 1997:36). This departs from a far more common view that simply postulates the EPP and derives the obligatory presence of subjects, lexicalized or empty, as an a priori (Sano & Hyams 1994, among others). That view fails to predict the longitudinal quantifications that have been presented here.

4.3 *Mode-implied subjects*

Before the acquisition of discourse grammar, subjects are sometimes absent in I-marked predicates (Rizzi 1994, among others). I have identified the subject-less utterances as representations of situation-bound operator modes (16b) (Van Kampen 1997: 105f). The best example in the adult language is the imperative. The situation-bound modes have a <+fin/I°> verbal form, but they lack a subject. I propose, following De Haan (1987), that these early <+fin> forms are modal operators. They express modes that imply the presence of a specific person. Unlike the real subjects, mode-implied subject cannot vary in person, i.e. *veux* is inherently 1st person. Moreover, they are situation-bound. The modes are listed in (20) (see Palmer 1986 for intentional (dynamic) versus deontic modality in adult language).

(20) Modes in early child language

a.	wish/ability of the child	intentional mode	(for 1^{st} p.)
b.	command by the child	deontic mode	(for 2^{nd} p.)
c.	decision about naming/characterizing	constative mode	(for 3^{rd} p.)

None of these modes have a syntactically expressed subject. The subject je (veux) as in (21a) below and il (est) as in (21c) is pragmatically implied rather than being present syntactically. The mode-implied subject gaps appear not only in the early speech of Grégoire, but also in early child Dutch and English. They are probably a characteristic of early child language in general. This crucially differs from previous analyses, like Rizzi (1994) who argues that null subjects in child language arise in discourse-bound contexts. My claim is that the discourse-bound contexts are not available for the child before the acquisition of D-marking.

(21) Mode-implied subj	English	<u>French</u>	<u>Dutch</u>
a. 1 st pers.	wanna bear	veux partir	wil beer/kan doen
b. 2 nd pers.	take doll	mets voiture	doe ogen dicht
c. 3 rd pers.	is X/goes there	est tombE/est ours	is beer/moet zo

One may speak about operators and 'mode-implied' subjects, because the <+fin> forms in these constructions (*est/veux/mets* 'is/want/put') are highly frequent and serve as major pragmatic oppositions in standard situations. It is a matter of fixed operators. Except for the constative mode, there are no substantial examples in child French of 3rd person 'drop' like in (22), at least not for Grégoire.¹²

(22) a. *lit un livre	(reads a book)
b. *marche dans la rue	(walks in the street)

The strong influence of the subject implying <+fin, mode> is also demonstrated by the fact that the 1^{st} and 2^{nd} person pronouns je/tu do not, or hardly, appear in the Grégoire files until D-marking has become systematic. 1^{st} and 2^{nd} person were at first taken care of by the subject implying modes of (21). Hamann et al. (1996:table 4) show the same acquisition delay for je/tu,

¹² Parallel examples like (22) do appear in child Dutch. These are examples of discourse-implied null topics. Discourse-implied null topics concern null 3rd person d-pronouns only. Topic-drop is restricted to Spec,C and appears after the acquisition of V-2nd. See Van Kampen (1997:chap.4) for topic-drop in adult and child Dutch.

attributed here to the subject-implying modes of child language. The need for 1st and 2nd person pronouns follows from the growing predominance of the EPP-type of predication.

5. Conclusion

The following conclusions have now been reached.

- (23) a. I-marking precedes D-marking in both child Dutch and child French.
 - b. D-marking opens the way to discourse-oriented language. This again holds for child Dutch as well as for child French
 - c. The French *il* appears before systematic D-marking, but as a shadow pronoun not as a free anaphor
 - d. The EPP in child Dutch as well as in child French is realized before systematic D-marking. To that end, early child Dutch makes use of proper names and demonstratives, and early child French makes use of qa and a dislocated topic + shadow pronoun.

The learning steps in (23) suggest a general acquisition strategy that turns a proto-grammar for situation-bound structures into a truly human grammar for discourse-oriented structures that are systematically situation-free.

References

Ashby, William J. 1988. "The Syntax, Pragmatics, and Sociolinguistics of Left- and Right-dislocations in French". *Lingua* 75.203-229.

Atkinson, Martin 1979. "Prerequisites for Reference". *Developmental Pragmatics* ed. by Elinor Ochs & Bambi B. Schieffelin, 229-249. London: Academic Press.

Avrutin, Sergey 1994. *Psycholinguistic Investigations in the Theory of Reference*. PhD. dissertation MIT.

Baauw, Sergio & Fernandez Cuetos. to appear. "The Interpretation of Pronouns in Spanish Language Acquisition and Breakdown". *Language Acquisition*.

Cat, Cecile de. 2002. French Dislocation. PhD. dissertation York University.

Evers, Arnold & Jacqueline van Kampen. 2001. "E-language, I-language and the Order of Parameter Setting". UiL OTS Working Papers 00105-S-S, Utrecht University.

Ferdinand, Astrid 1996. The Development of Functional Categories. The acquisition of the subject in French. PhD dissertation Leiden University.

Haan, Ger J. de. 1987. "A Theory-bound Approach to the Acquisition of Verb Placement in Dutch". Formal Parameters of Generative Grammar ed. by

- Ger J. de Haan & Wim Zonneveld, vol. III, 15-30. Yearbook 1987, Utrecht: UiL-OTS.
- Haegeman, Liliane 1993. "The Morphology and Distribution of Object Clitics in West-Flemish". *Studia Linguistica* 47.57-94.
- Haegeman, Liliane 1996. "Root Infinitives, Clitics and Truncated Structures". *Generative Perspectives on Language Acquisition* ed. by Harald Clahsen, 271-307. Amsterdam & Philadelphia: John Benjamins.
- Hamann, Cornelia, Luigi Rizzi & Uli H. Frauenfelder. 1996 "On the Acquisition of Subject and Object Clitics in French". *Generative Perspectives on Language Acquisition* ed. by Harald Clahsen, 309-334. Amsterdam & Philadelphia, Penn.: John Benjamins.
- Heim, Irene R. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD. dissertation University of Massachusetts, Amherst.
- Hickmann, Maya 1982. The Development of Narrative Skills: Pragmatic and meta-pragmatic aspects of discourse cohesion. PhD. dissertation University of Chicago.
- Hickmann, Maya & Henriette Hendriks. 1999. "Cohesion and Anaphora in Children's Narratives". *Journal of Child Language* 26.419-452.
- Hoekstra, Teun & Nina Hyams. 1995. "The Syntax and Pragmatics of Early Clause Structure". ms. Leiden University & UCLA.
- Hoekstra, Teun & Nina Hyams. 1998. "Aspects of Root Infinitives". *Lingua* 106.81-112.
- Hulk, Aafke 1986. "Subjet clitics and the pro-drop parameter". Formal Parameters of Generative Grammar, papers of Going Romance ed. by Peter Coopmans, Ivonne Bordelois & Bill Dotson Smith, 107-121. Dordrecht: ICG printing.
- Jakubowicz, Celia & Catherine Rigaut. 1997. "L'Acquisition des Clitiques Nominatifs en Français". *Les Pronoms: Morphologie, Syntaxe et Typologie* ed. by Anne Zribi-Hertz, 57-99. Paris: Presses Universitaires de Vincennes.
- Jakubowicz, Celia, Lea Nash, Catherine Rigaut & Christophe Gérard. 1998. "Determiners and Clitic Pronouns in French-speaking Children with SLI". *Language Acquisition* 7.113-160.
- Kampen, Jacqueline van. 1997. First Steps in Wh-movement. Delft: Eburon.
- Kampen, Jacqueline van. 2001. "Pragmatic Bootstraps for Category Assignment: from Non-discourse towards Discourse Oriented Language". *Proceedings of ELA* (Lyon, December 5-8) ed. by Sophie Kern. CD-rom.
- Kayne, Richard 1975. French Syntax. Cambridge, Mass.: MIT Press.
- Lambrecht, Knud 1981. *Topic, Anti-topic and Verb Agreement in Non-standard French.* (= Pragmatics & Beyond, 6). Amsterdam: Benjamins.

- Lyons, John 1977. Semantics. Cambridge, Mass.: Cambridge University Press.
- Lyons, John 1979. "Deixis and Anaphora". *The Development of Conversation and Discourse* ed. by Terry Myers, 88-103. Edinburgh: University Press.
- Nash, Lea & Alain Rouveret. 1997. "Proxy Categories in Phrase Structure". *NELS* 27 ed. by Kiyomi Kusumoto, 287-304. Amherst, Mass.: GLSA.
- Palmer, Frank R. 1986. *Mood and Modality*. Cambridge, Mass.: Cambridge University Press.
- Perlmutter, David 1972. "On the Relevance of Shadow Pronouns to French Relatives". Papers from the 5th Annual Meeting of the Chicago Linguistic Society. Chicago, Ill.: CLS.
- Pierce, Amy 1989. On the Emergence of Syntax. PhD. dissertation MIT.
- Plunkett, Bernadette 2002. "Null Subjects in Child French Interrogatives: a View from the York Corpus". Romanistische Korpuslinguistik: Korpora und gesprochene Sprache/Romance Corpus Linguistics: Corpora and Spoken Language (=ScriptOralia; 126) ed. by Claus D. Pusch & Wolfgang Raible, 441-452. Tübingen: Narr.
- Postal, Paul 1966. "On So-called Pronouns in English". *Monograph Series on Language and Linguistics 19* ed. by Francis P. Dinneen, 177-206. Washington, D.C. Georgetown University Press.
- Rizzi, Luigi 1994. "Early Null Subjects and Root Null Subjects". *Language Acquisition Studies in Generative Grammar* ed. by Teun Hoekstra & Bonnie D. Schwartz, 151-176. Amsterdam & Philadelphia: John Benjamins.
- Rothstein, Susan 1983. The Syntactic Forms of Predication. PhD. diss. MIT.
- Rouveret, Alain to appear. "Why are Complement Clitics Difficult to Master?" *Essays on Syntax, Morphology and Phonology in SLI* ed. by Celia Jakubowicz, Lea Nash & Ken Wexler. Cambridge Mass.: MIT Press.
- Sano, Tetsuya and Nina Hyams.1994. "Agreement, Finiteness and the Development of Null Arguments". *NELS 24* ed. by Mercè Gonzàlez, 543-558. Amherst, Mass.: GLSA.
- Schaeffer, Jeanette 1997. *Direct Object Scrambling in Dutch and Italian Child Language*. PhD. Dissertation UCLA Los Angeles.
- Velde, Marlies van der, Celia Jakubowicz & Catherine Rigaut. 2002. "The Acquisition of Determiners and Pronominal Clitics by Three French-speaking Children". *Proceedings of GALA 1999* ed. by Ingeborg Lasser, 115-132. Frankfurt: Peter Lang.
- Williams, Edwin 1994. *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.

NOMINAL ELLIPSIS AND MORPHOLOGICAL STRUCTURE IN SPANISH*

LAURA M. KORNFELD & ANDRÉS L. SAAB CONICET - Universidad Nacional del Comahue

1. Introduction

In this paper we account for nominal ellipsis in Spanish, a language in which this phenomenon is far more productive than in other Romance languages, such as French or Italian.

We present here a new set of data that poses difficulties to previous analyses, especially the ones postulating the existence of empty categories or deep anaphora in these cases (such as Brucart 1987). We adopt in general the approach of PF-deletion (see Sag 1976 and Chomsky & Lasnik 1993), but refine some ideas related to morphological identity in the light of our own data. All this seems to support the general framework of Distributed Morphology (DM) (Halle & Marantz 1993). We show that nominal ellipsis is strongly constrained by operations and properties of Morphological Structure (MS), such as morphological identity, late insertion, adjacency and competition.

The paper is organized in this way: first, we will present the basic data with respect to Spanish nominal ellipsis, focusing on the different constraints that appear when there is an article or a definite determiner before the elided noun. We will also present evidence against the approach of nominal ellipsis as a deep anaphor or empty category. Second, adopting the framework of Distributed Morphology, we will account for a set of problematic data for previous explanations of the phenomenon in terms of ellipsis as PF-deletion (Depiante & Masullo 2001). Third, we will focus on a gap in the Spanish nominal ellipsis paradigm: the fact that the article cannot be followed by a preposition different from de. We will especially discuss Raposo's (1999) proposal that de has peculiar properties that differ from other Spanish

^{*} We would like to thank Marcela Depiante, for inspiration, discussion and help of all sorts; Inés Kuguel, for correcting the first translation of this paper; Jairo Nunes and João Costa, for their Portuguese data; Jorge Hankamer, for stimulating discussion, and, finally, the two anonymous reviewers and the audience of Going Romance, for many useful comments.

prepositions. Finally, we will show that the DM notion of competition can explain the well-known differences in the properties of nominal ellipsis between two groups of Romance languages (Spanish, Portuguese and Catalan vs. Italian and French).

2. The basic data

It has been frequently noted that one characteristic feature of Spanish is nominal ellipsis. The standard cases include the possibility of eliding the noun before an AP (1a), a PP (1b) or a relative CP (1c) (examples adapted from Brucart's 1987 thesis):

- (1) a. *Utiliza el coche antiguo para trabajar y el* e *nuevo para* uses-he the car old for working and the *e* new for *el fin de semana*.

 week-ends
 - "He uses the old car for working and the new one for week-ends."
 - b. El hijo de Luis y el e de Antonio se han hecho muy amigos. the son of Luis and the e of Antonio SE have done very friends "Luis' son and Antonio's have become good friends."
 - c. La casa que visitaste ayer y la e que viste hoy The house that visited-you yesterday and the e that saw-you today son de Alberto.

are of Alberto

"The house that you visited yesterday and the one that you saw today belong to Alberto."

Although it is far more productive than in other Romance languages, Spanish nominal ellipsis has a number of non-obvious constraints. Note that the article or definite determiner contrasts with other Spanish determiners, that freely accept nominal ellipsis even if there is no complement to the noun (see 2), as well as the occurrence of a preposition different from *de* heading the modifier PP (see 3):

- (2) a. *Buscaba el e. looked for-I the.M.SG e
 - b. Buscaba uno/este/alguno /tres e. looked for-I one/this/some.M.SG/three e "I looked for one/this/some/three."

- (3) a. Compraré el / este / alguno / uno / tres e de matemáticas. will buy-I the / this/ some / one.M.SG / three e of maths "I will buy the / this / some / one / three about maths."
 - b. Compraré *el / este / alguno / un / tres e con lazo / will buy-I the / this / some / one.M.SG / three e with lace / a cuadros / en cartón / sin grabados. to squares / in cardboard / without prints
 "I will buy *the / this / some / one / three one(s) with lace / with squares / in cardboard / without prints." (Brucart 1987)

Besides, the gap always has to be a constituent: No (cf. examples in 1) or some intermediate projection or No, such as *mesa con rueditas* in (4):

(4) María consiguió la mesa con rueditas grande y Pedro la e Mary got the table with wheels.DIM big and Peter the.F.SG e chica. small "Mary got the big table with little wheels and Peter the small one."

Brucart (1987:228), in one of the most detailed studies about Spanish ellipsis, rules out the idea that nominal ellipsis can be due to "transformations of elision of syntactic categories", and supposes that it consists of an empty category generated in the base. In Hankamer & Sag's (1976) terms, Brucart would postulate that nominal ellipsis is an instance of deep anaphora.

We only sketch out here two arguments against this view of nominal ellipsis. First, nominal ellipsis is different from *one*-substitution in English with respect to the possibility of selecting complements. It is a well-known fact that, in English, *one* cannot take *of*-complements, whereas in Spanish *de*-complements can occur freely with nominal ellipsis, as the equivalent sentences in (5) show:

- (5) a. *The students of physics are taller than the ones of chemistry. (Panagiotidis 2003)
 - b. Los estudiantes de física son más altos que los de química.

Panagiotidis (2003) attributes the ban on *of*-complements to the lack of conceptual content of *one*, a pronominal noun in his theory. That is, *one* cannot assign a θ -role to its complement. If this analysis is on the right track, then we should take exactly the contrary assumption for Spanish: the possibility of *de*-

complements is due to the null N having semantic content and, thus, assigning a θ -role to its complement. It seems plausible, then, that the null N is precisely *estudiantes*, whose semantic features would be present into the second N° during the derivation before Spell-Out.

Second, although it is not easy to construct this kind of evidence, it seems possible to prove that there could be internal structure in the cases in which a N' is elided, as some kind of 'missing antecedent' phenomenon:

(6) Nunca comí una torta con higos, pero Juan probó la e que never eat-I a cake with figes, but Juan tasted the.F.SG e that hizo Pedro y dice que le parecieron deliciosos. did Pedro and says-he that to-him seemed-they delicious "I have never eaten a cake with figes, but Juan tasted the one that Pedro did and he says that they seemed delicious to him."

The example above suggests that the empty category has internal structure. Then, it constitutes strong evidence against a deep anaphora analysis and in favor of a transformational analysis of nominal ellipsis (see Raposo 1999 and section 4 for more arguments).

3. Nominal ellipsis and Distributed Morphology

We adopt here the hypothesis of ellipsis as PF-deletion and intend to refine previous proposals of this kind. Particularly, we consider Depiante & Masullo (2001), who observe that the elided noun sometimes does not coincide with its antecedent in number (7a), although it must do so in gender (7b):

- (7) a. *Juan visitó a sus tíos y Pedro prometió visitar al* e *de él.* Juan visited his.PL uncles and Pedro promised visit the.M.SG *e* of his
 - b. *Juan visitó a su tío y Pedro prometió visitar a la e de él.

 Juan visited his.SG uncle and Pedro promised visit the.F.SG e of his
 - c. Juan visitó a su tío y Pedro prometió visitar a los e de él. Juan visited his.SG uncle and Pedro promised visit the.M.PL e of his "Juan visited his uncle(s) and Pedro promised to visit his."

Depiante & Masullo utilize this difference in the behavior of the two morphological categories with respect to nominal ellipsis to propose that, in Spanish, gender is an intrinsic feature of the noun, while number is syntactic and, in fact, heads a functional projection intermediate between DP and NP. That is to say, nouns enter into the numeration bare in number, but inflected in

gender. Evidence of this is the fact that a singular noun with a plural antecedent (as *tio* in 7a) can be elided: given the condition of strict morphological identity (Lasnik 1995 and others), this could only happen if at some point in the derivation both nominal elements are identical. In this way, the differences between gender and number virtually parallel the distinction done by Lasnik (1995) between English auxiliary verbs (inflected already in the numeration) and main verbs (non inflected in the numeration): while gender features would only be checked in the syntax, number features would be acquired throughout the derivation.

Two observations, however, can be done to Depiante & Masullo's proposal. First, it does not seem so easy to determine at which point in the derivation the identity between the antecedent and the elided noun would happen, considering that, according to these authors, ellipsis is PF-deletion. A greater problem is represented by sentences such as (7c), where – contrary to (7a) – the antecedent is singular and the elided noun is plural: in this case, it does not seem clear how the plural morpheme –*s* can be deleted, since this morpheme constitutes an independent element of the numeration and heads its own projection in the syntax. In other words, it is necessary to say 'something more' about the antecedent of the plural morpheme in (7c).

A way of solving these difficulties is to adopt the principles of Distributed Morphology (DM) (Halle & Marantz 1993). We assume, then, the existence of a Morphological Structure (MS) post Spell-Out, in which the insertion of lexical items is produced, as well as the operations of agreement and concord. In this way, syntax and LF only operate with morphosyntactic and semantic features, without any phonological information. This is an important point that distinguishes our analysis from Depiante & Masullo's, who assume that phonological features are already specified in the numeration. In (8), a scheme of the sentence just post Spell-Out is represented, in a very simplified way (recall that all lexical items are actually bundles of formal features at this point in the derivation):

(8) Juan visit Pas Pos-3SG tío y Pedro promet Pas visitar D tío Pl de D-3SG

When insertion of lexical items takes place, two possibilities arise. If all the items are inserted, after the relevant operations of fusion and phonological "adjustment" we obtain the non elided sentence, *Juan visitó a su tío y Pedro prometió visitar a los tíos de él*. However, it is also possible to skip the insertion of the phonological features of *tíos*, generating the sentence with ellipsis *Juan visitó a su tío y Pedro prometió visitar a los de él*. The lexical

item *tío* is not inserted because it is identical to its antecedent in the first part of the sentence, while the plural node is afterwards associated with the determiner. We consider ellipsis, then, as a phenomenon of non insertion of phonological features into terminal nodes, under strict identity of lexical and formal features.

On the other hand, sentence (7b) violates the condition of strict identity, because the masculine form of the antecedent noun (tio) does not coincide with the feminine form of the elided noun (tia). As Depiante & Masullo observe, gender is lexical in the sense that the gender mark is inserted in the lexical X^o slot as part of the noun stem. Then, the phonological features of the noun could only not be inserted in N^o if there was a proper antecedent, such as tia(s): MS cannot separate the stem ti- from the gender mark -a, since they constitute a single item. In other words, $X^o=tio+fem$ can never be the same word as $X^o=tio+masc$. Therefore, the identity with the masculine antecedent could not be computed in any way and the non-insertion of the phonological matrix of tia accounts for the ungrammaticality of the sentence.

In this way, our proposal maintains the view of ellipsis as PF-deletion in spirit, but adds some elements that increase its plausibility.

A strong conceptual argument in favor of late insertion of phonological features (especially in a PF-deletion view of ellipsis) is constituted by a number of phenomena of phonologically conditioned allomorphy. A good example is the Spanish alternation of el as allomorph of the feminine article la, before a stressed /a/s sound:

(9) El aula chica está ocupada, pero <u>la</u> /*el e grande no. the.(M)SG room small is occupied but the.(F)SG/*the.(M)SG e big not "The small room is occupied but the big one is not."

Without the late insertion assumption, it would be necessary for (9) to postulate that some morphophonological rules have the property of altering the original phonological matrices of the items (that is to say, they could transform la into

¹ We assume that gender does not head its own projection (against Picallo 1991). This seems to be obvious for the "standard" cases of arbitrary gender (where gender does not participate in syntax at all and is determined only in MS), but could also be applied in the cases where gender has some semantic interpretation related to 'sex'. We cannot develop a detailed argumentation here, but interesting lexical irregularities (e.g., in parental terms: *tío-tía*, *hermano-hermana* and *suegro-suegra* exhibit gender marks, but *yerno-nuera*, *padre-madre* are pairs with different genders and different roots) seem to suggest that the gender feature is always computed with the N°.

el or el into la). Rules of this kind would weaken the elegance of the theory too much.

In our analysis, on the contrary, the occurrence of el and la in (9) would follow directly, since, under the assumption of late insertion, syntax does not operate with phonological matrices but only with bundles of abstract morphosyntactic features. So, the insertion of the phonological features of aula, with stressed initial /a/, determines the choice of el for the first slot with the features [D, definite, feminine, singular]; at the same time, the non insertion of the phonological features of the second aula determines the choice of la as the 'winner' allomorph for the other slot with the same features. Note, finally, that, given (9), it seems natural to assume that the insertion of phonological features occurs in two different stages: first, the lexical items and later the functional ones. The insertion of (allomorphs of) functional items is sensitive, then, to the previous occurrence of lexical items.²

On the other hand, it is important to observe that the determiner must express variable inflectional features in order to license nominal ellipsis. That is to say, an invariable form of D does not license ellipsis. A good example is provided by the differences in the distribution of two Spanish wh-elements, *qué* and *cuál*:

² As proposed by Kornfeld (2003), this idea could be extended to several kinds of phenomena in many languages. For example, the alternation of prepositions with verbs of movement in French (e.g., aller en Angleterre / au Japon / à Paris) seems to be sensitive not only to gender but also to some very specific semantic features of the inserted noun. The fact that so much (phonological and semantic) information about lexical nodes has to be available in the moment of inserting functional items suggests that, ultimately, lexical morphemes could be absolutely specified already in syntax, against some DM versions (e.g., Harley & Noyer 1999).

³ The role played by the determiner could be even more relevant, in fact. Actually, certain data show that Spanish nominal ellipsis is strongly constrained when the elided noun is bare: if the antecedent and the elided noun are both plural, the sentence is marginal (cf. example i below), but, if there is no identity of number, the sentence is ungrammatical (cf. ii), specially when the structures are not strictly parallel (the examples are considerably better if coordinated structures and the same verbs appear, as in Leonetti 1999b:819):

 ⁽i) ?? Como en la verdulería solo quedaban pimientos rojos, no tengo e amarillos.
 as at the greengrocer's only are left peppers red not have-I e yellow.PL

⁽ii) *Como en la verdulería solo quedaba un pimiento rojo, no tengo e amarillos. as at the greengrocer's only is left a/one pepper red not have-Ie yellow.PL "As at the greengrocer's only red peppers /a red pepper are / is left, I don't have yellow ones."

In our view, the ungrammaticality of (ii) is derived from the fact that the plural morpheme of *pimientos* could not be deleted, because it has no antecedent, given the singular number in the first noun and the lack of a determiner for the second noun.

(10) ¿Qué/cuál libro de Borges te gusta y *qué/cuál e de Bioy es what / which book of Borges you.CL like and *what / which e of Bioy is mejor?

best?

"What / which book of Borges pleases you and *what / which one of Bioy is the best?"

Note that the contrast between *qué* and *cuál* consists, precisely, in the opposition '(morphologically) variable/invariable': *qué* is compatible with masculine, feminine, and plural nouns (*qué chico/a/s*), whereas *cuál* is a singular form compatible with feminine and masculine nouns (*cuál chico/a*) and has a plural form *cuáles*.⁴ In Spanish, then, nominal ellipsis is possible only if it is licensed by a determiner that expresses inflectionally at least number. Note that this remark resembles Lobeck's (1995) about English NP ellipsis, although her analysis differs considerably from ours.

4. Adjacency and Morphological Structure

It has been observed that in Spanish (and other Romance languages), nominal ellipsis is possible with noun complements (cf. 11a), but not with adjuncts (cf. 11b) (just the opposite of *one*-substitution in English):

- (11) a. Vi a los estudiantes de física, pero no a los e de lingüística. saw-I the students of physics but not the.M.PL e of linguistics "I saw the physics students but not the linguistics ones."
 - b. *Vi a los estudiantes con gafas, pero no a los e sin gafas. saw-I the students with glasses but not the M.PL e without glasses "I saw the students with glasses but not the ones without glasses." (Raposo 1999)

However, Raposo (1999) convincingly shows that the contrast in (11) cannot be reduced to the asymmetry between complements and adjuncts, since not all complements can license nominal ellipsis, as we can see in the Portuguese

⁴ As a reviewer points out to us, this analysis can be extended to the quantifier *cada* "every" which, like *qué*, is also invariable. The difference is that the former is not compatible with plural nouns (**cada chicos* "every boys"). Thus, the prediction is that *cada* cannot license ellipsis. This is borne out:

⁽i) *cada estudiante de lingüística y cada e de física every student of linguistics and every e of physics

example (12a). At the same time, adjuncts introduced by de are clearly grammatical as in (12b):

- (12) a. *A crença em Deus é mais forte que a e no Diabo. the belief in God is stronger than the F.SG e in Devil "The belief in God is stronger than the one in the Devil."
 - b. Las personas de tarjeta azul pasan; las e de tarjeta roja quedan the persons of card blue pass; the.F.PL e of card red stay aquí.

here

"The persons with a blue card can pass; the ones with a red card must stay here."

The contrast in (11) reflects, then, the opposition between de and the other prepositions with respect to ellipsis in Spanish and Portuguese. Raposo establishes the following generalization:

(13) At PF, the definite determiner cannot immediately precede a full preposition within an anaphoric DP.

Raposo proposes that Spanish and Portuguese definite determiners are clitics (or, more specifically, proclitics), since their lexical phonological matrix is weak (i.e., it lacks lexical stress). As clitics, the determiners need to be attached to a full word, in order to satisfy their morphophonological requirements. Raposo assumes the theory of phases proposed by Chomsky (1998) and argues that full PPs are barriers for cliticization. Recall that in Chomsky's (1998) system, phases are defined by their propositional character. Raposo proposes to extend this idea to the NP: a nominal phase would include at least a predicate N and the set of projections selected by it. Besides, the phase must contain the full set of extended projections, that is, D and (in the cases in which DP is the complement of a P) also P.

Then, the difference between (11a) and (11b) is explained because, in the latter, *sin* heads a PP that constitutes a phase in itself (*sin gafas*), so it is not in the same phase as the noun *estudiantes*. In the former, *de*, being a dummy case marker or the reflex of inherent genitive case, is morphologically determined by the matrix noun and late inserted in the morphological component. In other words, *de* does not belong to the phase containing the noun *lingüística*, but it is lately inserted in *los estudiantes*. As a consequence, nominal ellipsis is not possible in a sentence such as (11b) because the determiner cannot be

cliticized, since the PP *sin gafas* corresponds to another phase than the DP *los estudiantes*. On the contrary, in a sentence such as (11a), the cliticization rule can be applied, because *de* is lately inserted in the DP *los estudiantes*, so it is in the same phase as the determiner.

Thus, the analysis in terms of phases predicts that definite determiners must be adjoined to an element of their own phases. The only preposition that satisfies this requirement in Spanish and Portuguese is de.⁵

Although we agree with Raposo in that the contrast between de and the other prepositions is a question of PF level, we find that his analysis cannot explain more complex paradigms. Note that all the data presented by him crucially involve the use of de as a dummy preposition. However it is not difficult to imagine examples in Spanish in which de does not function in this way, i.e., examples with de in which the elided noun does not determine its occurrence within the phase of the determiner. And, in spite of this, for at least some speakers, sentence (14) is perfectly grammatical:

(14) El chico que me señalaste y el e del que te hablé son the boy that I.DAT showed-you and the e of which you.DAT talked-I are amigos.

friends

"The boy that you showed me and the one about which I talked to you are friends."

In (14), *de* cannot simply be a mark of case lately inserted in a position adjacent to the noun, its original position being the head of the complement of *hablar*. In fact, the preposition is within a phase different from the determiners, i.e., the CP containing it (recall that in Chomsky 1998 CPs constitute phases). Raposo's analysis predicts, contrary to the facts, that (14) must be ill-formed.

On the other hand, explaining the contrasts in (11) by the defective lexical character of *de* would not be correct. In (15), for instance, the preposition has a clear semantic value of 'source':

⁵ Jairo Nunes and João Costa (p.c.) do not agree with Raposo's Portuguese data and argue that ellipsis is also possible with a preposition other than *de*:

⁽i) O caderno com a capa amarela é meu, o e com a capa vermelha é da Maria. the folder with the cover yellow is mine the e with the cover red is of-the Maria "The folder with the yellow cover is mine, the one with the red cover belongs to Maria." (BP. Jairo Nunes)

⁽ii) Não vi os estudantes com óculos, mas sim os e com cabelo louro. (EP, João Costa) not saw-I the students with glasses, but AFF the e with hair blond "I didn't see the students with glasses, but I did see the blond-haired ones."

(15) El adjunto que se mueve no es el e del que se extrajo the adjunct that SE moves not is the e of which SE extracted-it el elemento-qu.

the element-wh

"The adjunct that moves is not the one from which the wh-element is extracted."

In sum, Raposo's analysis undergenerates, since it predicts that some grammatical sentences (such as 14 or 15) are out. It seems, then, that Raposo's generalization in (13) must be slightly modified, at least for Spanish:

(16) At PF, the definite determiner cannot immediately precede a preposition other than *de*.

So, independently of the distinction between full and empty prepositions and of the structural position or domain in which the prepositions are, the morphological component only allows cliticization of the determiner with the preposition *de*. The paradigms presented here, then, seem to suggest that MS is sensitive to linearity or adjacency (in the sense of Bobaljik 1994), and not only to structure. It is evident that *de* must have some exceptional morphophonological property that distinguishes it from the rest of the prepositions, but it seems that this property cannot be explained in terms of phases.

Although we will not offer here a detailed explanation, we have found that the exceptional property of *de* consists in its own phonological weakness. It has been noted that *de* is a weakly stressed preposition, even more weakly stressed than other prepositions, since it does not allow focus: ¿CON or SIN? ¿PARA? ¿POR?, but *¿DE?.⁶ In this way, we prefer an exclusively phonological account, based not only on the phonological weakness of the article (as Raposo does), but also on the phonological properties of the preposition. *De* would be, then, a proclitic that must adjoin to the following element and the determiner, in turn, also adjoins to the whole complex.⁷

⁶ Note that even the preposition a can be focused, unlike de:

⁽i) A: Me voy a la oficina. "I will go to the office."

B: ¿A (qué)? "TO (what)?"

⁽ii) A: Voy a hacer una torta. "I will make a cake."

B: ¿DE *(qué)? "OF *(what)?"

⁷ A reviewer points out to us that this analysis is problematic in cases in which the article contracts with a preceding preposition, because once contraction takes place, then there is no necessity of applying the cliticization rule. This is not borne out:

5. Cross-linguistic variation among Romance languages

We assume here a traditional hypothesis in Spanish grammar due to Bello (1847): determiners and pronouns are elements in complementary distribution. For the most part of the cases registered in paradigms (2) and (3), the fact that there is a unique underlying form seems to be obvious, since determiner and pronoun are homophones. The only difference, in some pairs, is the stronger stress on the pronoun or the occurrence of a final o for the masculine (un/uno, algún/alguno, ningún/ninguno) that, as Brucart (1987:242) observes, are rather "superficial phonetic factors". So, it is relatively easy to suppose that they are allomorphs in complementary distribution; that is, un and uno compete for the same slot D ([D, indefinite]), and the fact that one or another form appears is determined by the morphophonological context (i.e., if the noun is elided, uno will appear; if not, un will be the 'winner' form).

With respect to the article or definite determiner, it has been frequently noted that there exist differences between two groups of Romance languages. So, the Spanish examples analyzed here are similar to the Portuguese (see 17) and the Catalan (18) ones, and must be distinguished from the French (19) and the Italian (20) ones:

(17) a. Há uma nova ordem de cousas e eles amam as e there is a new order of things and they love the.F.PL e perversas.

perverse

"There is a new order of things and they love the perverse ones."

However, we think that contraction does not eliminate the clitic characteristic of the determiner. In other words, the contracted complex is also like a clitic, so the cliticization rule must apply.

 ⁽i) Ayer fui al *(supermercado) con descuento.
 yesterday went-I to-the *(supermarket) with discount
 "Yesterday, I went to the discount supermarket."

⁽ii) Estamos hablando del *(hombre) sin gafas.
are-we talking of-the *(man) without glasses
"We are talking about the man without glasses."

⁸ We consider that it would be possible to assume some kind of competition also in the case of articles and pronouns of 3^d person, a traditional idea in Spanish grammar, starting with Bello and supported by the homophony between ℓ (pronoun) and ℓ (determiner). The winner would be determined in each case by the syntactic context: so, if the D is followed only by a nominal ℓ , the pronoun (=the stronger form) must appear (we are assuming here the proposal of Panagiotidis 2003 about considering pronouns as D+a null nominal complement), whereas the article (=the weaker form) can be inserted if any kind of complement (nominal or not) occurs.

b. *A perda da alma exterior implica a* e *da existência* the loss of-the soul exterior implies the.F.SG *e* of-the existence *inteira*.

entire

"The loss of the exterior soul implies the loss of the entire existence."

- c. Quero que pinte um amor como o e que tive em Tel Aviv. want-I that paint-she a love as the M.SG e that had-I in Tel Aviv "I want her to paint a love as the one that I had in Tel Aviv."
- (18) a. La dona de Pigmalió és una de les e més irritants de la the woman of Pigmalion is one of the.F.PL e most irritating of the literatura.

literature

"The woman of Pigmalion is one of the most irritating ones in the literature."

b. Es uma dona sense nom: el e de Galatea se li va is-she a woman without name the.M.SG. e of Galatea SE her Pas donar més tard.

give later

"She is a woman without name: later, people have given her the one of Galatea."

c. Aquests discursos, els e que es deriven de la confiança these discourses, the.M.PL e that SE derive from the faith en l'home...

in the man...

"These discourses, the ones that derive from the faith in the man..."

- (19) a. Je connais beaucoup de filles: elle est la e plus belle.

 I know many girls she is the F.SG. e more beautiful "I know many girls: she is the most beautiful one."
 - a'. *Quant aux photos*, *j'ai regardé* **les/celles* e *apportées* about the photos I have seen *the.PL/these.F *e* brought *par mon père*.

by my father

"Concerning the photos, I have seen the ones brought by my father."

b. Les chiens de la ferme et *les/ceux e du douar...
the dogs of the farm and *the.PL/these.M e of-the douar...
"the dogs from the farm and the ones from the douar..."

*Je disais les données que je savais, inventant *les/celles* the details that I knew, inventing *the.PL/those e I told que je ne savais pas.

that I NE knew not

"I told the details that I knew, inventing the ones that I didn't know."

- (20) a. piccolo e the.M.SG little "the small one"
 - a'. *il/quello e costoso *the/this.M e expensive "the expensive one"
 - *la/auella e di Carlo h. *the/this.F e of Carlo "Carlo's (one)"
 - c. *i/quelli e che mi piacciono *the/these.M e that me please "the ones that I like" (Italian examples from Leonetti 1999a:61)

As we can deduce from these examples, there are a number of contexts in which the article can appear in Spanish, Portuguese and Catalan, but not in French or Italian, where it must be replaced by a pronoun, a stronger form (cf. 19a'-c and 20a'-c). This distinction can be explained because each group of languages has a different distribution of the elements in competition for D nodes with the definite feature. In French and Italian, two D elements (le/celui. il/quello) share the feature 'definite'. Note that *celui* and *quello* in (19) and (20) do not have a real demonstrative meaning (as in other contexts), but rather they are semantically much more neutral than their equivalents in Spanish, Portuguese (este) or Catalan (aquest), as Leonetti (1999b:818) has pointed out. The factor that defines which element (le or celui; il or quello) wins the competition is simply the insertion or non insertion of the phonological features of the noun. As we have said before, the insertion of functional items (in this case, D) is sensitive to the previous insertion of 'true' lexical items (cf. footnote 2). Therefore, if the features of the noun are inserted, the article le/il "wins" the competition (and, in fact, French *celui livre); otherwise, the pronoun celui/auello 'wins' (*le aue je dis). On the other hand, in Spanish, Portuguese and Catalan, the article does not compete with the demonstrative pronoun (Sp./Pt. este, Cat. aquest), because both elements differ in meaning and must be inserted in different slots: [D, definite], for the article; [D, definite, dem], for the demonstrative pronoun.

Although this view about the complementary distribution between articles and a particular kind of pronouns in French and Italian is based on purely synchronic factors, it could be confirmed by the diachrony of Romance languages, since both pairs (*le/celui*; *il/quello*) originate in Latin *ille*.

6. Summary

In this paper, we have analyzed nominal ellipsis in the Romance languages as a process of non-insertion of phonological matrices in the Morphological Structure under strict morphological identity. Our proposal allows us to account in a unified way for a great number of phenomena not explained in previous versions of ellipsis as PF-deletion (Sag 1976, Chomsky & Lasnik 1993 and Depiante & Masullo 2001, for example), such as the sloppy identity concerning number (example 7c), the cases of *la/el* allomorphy before stressed /a/ (9) or the asymmetry between variable and invariable determiners (10). We have also shown that our view of ellipsis has empirical and conceptual advantages over those approaches that take nominal ellipsis as an instance of deep anaphora (Brucart 1987).

A theoretical consequence of our analysis is related to the explanation of phenomena of cross-linguistic variation. If our account of the differences in the behavior of nominal ellipsis in the Romance languages is on the right track, then we are giving a new meaning to the idea that cross-linguistic variation should be related to morphological properties of languages (Chomsky 1993). Finally, the explanatory adequacy of our proposal and its possible extension to other elliptical phenomena (such as VP ellipsis) is a question that only future research can determine.

References

Bello, Andrés. 1847. *Gramática de la Lengua Castellana*. Madrid: Arco Libros.

Bobaljik, Jonathan. 1994. "What does Adjacency do?". MIT Working Papers in Linguistics 22.1-32.

Brucart, José. 1987. La Elisión Sintáctica en Español. Barcelona: Bellaterra.

Chomsky, Noam. 1993. "A Minimalist Program for Linguistic Theory". Chomsky 1995. 167-217.

Chosmky, Noam. 1995. The Minimalist Program. Cambridge: MIT Press.

Chomsky, Noam. 1998. "Minimalist Inquiries: the Framework". *MIT Occasional Papers in Linguistics* 15.1-56.

—— & Howard Lasnik. 1993. "The Theory of Principles and Parameters". Chomsky 1995. 13-127.

Depiante, Marcela & Pascual Masullo. 2001. "Género y Número en la Elipsis

- Nominal: Consecuencias para la Hipótesis Lexicalista". Paper presented at *I Encuentro de Gramática Generativa*, Gral. Roca, 22-24 November 2001.
- Halle, Morris & Alec Marantz. 1993. "Distributed Morphology and the Pieces of Inflection". *The View from Building 20* ed. by Ken Hale & Samuel Keyser, 111-176. Cambridge: MIT Press.
- Hankamer, Jorge & Ivan Sag. 1976. "Deep and Surface Anaphora". *Linguistic Inquiry* 7.391-426.
- Harley, Heidi & Rolf Noyer. 1999. "State-of-the-Article: Distributed Morphology". *GLOT* 4.3-9.
- Kornfeld, Laura. 2003. "Late Insertion, Allomorphy and the Distinction between Lexical and Functional Items", ms., UBA.
- Lasnik, Howard. 1995. "Verbal Morphology: *Syntactic Structures* meets the Minimalist Program". Lasnik 1999. 97-119.
- ——— 1999. Minimalist Analysis. Oxford: Blackwell.
- Leonetti, Manuel. 1999a. Los Determinantes. Madrid: Arco Libros.
- ——— 1999b. "El Artículo". *Gramática Descriptiva de la Lengua Española* ed. by Ignacio Bosque & Violeta Demonte, 787-890. Madrid: Espasa.
- Lobeck, Anne. 1995. *Ellipsis: Functional Heads, Licensing and Identification*. New York: Oxford University Press.
- Panagiotidis, Phoevos. 2003. "One, Empty Nouns, and θ -assignment". Linguistic Inquiry 34.281-292.
- Picallo, M. Carme. 1991. "Nominals and Nominalizations in Catalan". Probus 3.279-316.
- Raposo, Eduardo. 1999. "Towards a Minimalist Account of Nominal Anaphora in Spanish and English", ms., UCSB.
- Sag, Ivan. 1976. Deletion and Logical Form. Doctoral dissertation, MIT.

THE DEVELOPMENT OF INALIENABLE POSSESSION IN ENGLISH AND SPANISH

ANA T. PÉREZ-LEROUX, CRISTINA SCHMITT & ALAN MUNN University of Toronto Michigan State University

1. Introduction

Within generative approaches to language acquisition, the theoretical construct of parameter has evolved towards a smaller, narrow lexical notion of parametrization. Consequently, the field's perspective on the issue of continuity (i.e., how much children's stages are like possible adult grammars), and on our understanding of the task of learning is changing. One important domain of language variation (and thus a central part of the developmental task) pertains to how languages partition semantic spaces - i.e., how the relevant morphosyntactic categories are mapped onto the relevant semantic features. We assume that there is a universal vocabulary of interpretable features available to children's language acquisition device. A second assumption about learning at the syntax/semantic interface is that forms with comparable morphosyntactic distribution lexically compete for a given semantic space (forms avoid being given identical senses, as in Pinker 1984).

The goal of this work is to discuss the acquisition of structures that have seemingly comparable syntactic distributions across languages but map into different semantic distributions. Specifically, we examine the case of the definite determiner in English and Spanish. To converge on the proper representation of the definite, a child would have to generalize its use only to target semantic contexts. Learning in this case has different properties from acquiring purely syntactic facts of a language (such as the word order and the inventory of grammatical categories of the target language). Semantic competition with other determiners predicts non-target interpretations of definites in young children. In this paper, we present two studies of children's interpretation of definite determiners with the purpose of investigating this possibility on a domain of parametric difference in the semantic properties of definite determiners, namely, with inalienable possession construal.

2. Inalienable Construal

Inalienable construal (IC) is a sub-case of a possession relation in circumstances where the possessed element is a part of the possessor element rather than accidentally or legally related to it. In some languages, the inalienable construal is represented internal to the possessed NP, and in others it involves structures relating two nominals that do not themselves form a constituent: one denotes a body part (or some other inalienable possession), and the one denotes the possessor of the body part. A well-known property of the Romance languages is the use of definites for inalienable construal in cases where English would use a (DP internal) possessive determiner (Guéron 1985, 1999; Vergnaud & Zubizarreta 1992). In some cases the external possessor is taken to be the clausal subject (1), and in others it is construed with a dative/benefactive argument, as in (2) and (3), depending on verb type: gestural motion verbs vs. others.

- (1) a. Juan levantó la mano.
 - b. *John raised the hand.
 - John raised his hand.
- (2) a. Juan *(se) lavó la cara.
 - b. *John washed the face.
 - c. John washed his face.
- (3) a. Los niños *(le) dieron la mano a María.
 - b. *The children gave the hand to Maria.
 - The children shook Maria's hand.

Beyond lexical verb class, there are other constraints (see Guéron 1999 for a full review). One important constraint relates to number of definite DP denoting the body part. The body part must be singular, independent of the plurality of the possessor DP, unless the body part is one which is naturally plural with respect to an individual. Thus, in (4), *las manos* 'the hands' can be inalienably construed because Juan and Maria have more than one hand each, but *las cabezas* 'the heads' cannot since, individually, they do not have more than one head, and a singular is required for IC:

(4) Juan y María levantaron las manos/la cabeza/*las cabezas. "Juan and Maria raised the hands/the head/the heads."

2.1 Previous analyses of IC as a parametric difference: a distributional view

Two main traditions in the comparative analysis of definites and inalienable possession exist. One likens the definite determiner to a pronominal anaphor (Guéron 1985) and the other likens it to an expletive determiner and treats the body part DP as a kind, much like the definite generic in French shown in (5a) (Vergnaud & Zubizarreta 1992).¹

(5) a. Le tigre vit en Afrique.

b. The tiger lives in Africa.

A full discussion of the semantics of definites with a kind interpretation needs to build into the semantics other contributors of interpretation, in particular that of number: we abstract from those possibilities here. Dayal (2002) observes that no language has a specific determiner for generic/kind readings, and that the definite determiner is an optimal candidate for reference to kinds when a bare form is not an option. In Chierchia's (1998) analysis of Romance, the definite plural is allowed to subsume reference to kinds when bare forms are not available. So, it is plausible to assume that the child is allowed the kind reading of definites by universal grammar.

Under a parametric view of the differences between English and Romance shown in (1)-(3), whatever the analysis, the child has to make a categorical decision about the target language. In contrast, a distributional view places the locus of variation not in categorical distinctions as expressed by different parameter settings but, rather, in different distributional patterns of form/meaning pairings.

What would one expect in a distributional view of the mapping between syntax and semantics? First, one would expect language internal evidence of the residual extension of the semantic opposition, as well as evidence that children make semantic distributional errors. The first prediction is about "parametric" divergence within a given language, similar perhaps in spirit as the one discussed by Roeper (1999): the expectation is that a given pairing of syntactic forms and the relevant meaning will be instantiated in both languages but that the distributional extension of this pairing will be different.

What has to be made clear is that learning, in these cases, is no longer seen as a problem of setting a parameter for a language (Spanish vs. English), but as

¹ Vergnaud and Zubizarreta (1992:643ff) claim that English definite kinds such as (5b) are not true kinds, but rather prototypes, based on some differences in the use of the definite kind in English compared to French. We ignore these differences here, and assume that the English definite kind is also a kind.

the problem of restricting the extension across which a given syntacticsemantic association is instantiated. There is no one-to-one correspondence between form and sense, but a complex topological space where other forms may block or bleed reference away from the pairing under consideration.

This dual behaviour can be seen in English. In principle, as shown by (1)-(3), English definites do not allow IC. However, IC appears productively in PP domains such as in (6). We argue that cases of IC in locative PPs, contrary to some arguments in the literature, constitute true inalienable construal and are not artifacts. Evidence for this comes from the fact that number seems as restricted in English here as in the Romance IC construction: (6) parallels the facts in (4):²

- (6) a. We hit them on the head/*the heads.
 - b. We hit them in the eye/eyes.

The present study seeks to find evidence for distributional effects in development: in both languages, the initial stages of the semantic representation of the definite determiner should include IC despite the unavailability of the construal in the given context for the target grammar. The English child is predicted to have stages of development where the interpretations of the definite with inalienable possession nouns run contrary to their experience (e.g., the→IC in D.O. contexts). The Spanish child will retain the option of IC for the definite but the English child will have to restrict it. The Spanish child will have to learn that IC may be restricted by number, or by lexical class of verb, but otherwise remains fine with definite determiners. Because the distributional approach does not depend on a parameter, we do not expect full clustering of properties in the intermediate stages. One such property could be the number restriction.

The idea of semantic competition is not new. A reasonable interpretation of it includes the idea that the most language-specific form will always win over the default sense. Chierchia's 1998 proposes precisely this in regards to the senses available for bare nouns. A bare noun can obtain certain interpretations via type-shifting operations that become unavailable when there is an overt determiner capable of contributing that sense to the semantic composition. To that effect, Chierchia proposes a blocking principle that would prevent type-shifting in those domains (1998:360). We adopt this notion

² The data in (6) also support the idea that kind readings are implicated in IC: English does not allow plural definites to be kinds in the usual case as in *The whales are mammals*, yet plural body parts can be definite kinds as in (6b).

of competition but argue that it needs to be refined. We wish to avoid direct competition between forms, as this would lead to a situation where the form with most content features would always win, or the language specific form would always win over the universal default. If this were the case, languages such as French and Spanish should not exist: the possessive determiner always has more information (person/gender/number features, as well as the possibility of an independent referential index), so IC should never be subsumed under the definite if a possessive is available in the grammar.

With respect to possession, one could say that a body part noun is inherently anaphoric: it seeks a body to attach to. But if the semantics of the body part terms were the only relevant determining factor in the IC interpretation, IC should not be subject to any constraints either in Romance or English. However, it is a fact that there are constraints in both languages although they are different constraints. Because the constraints are different, it is easy to take a diverging constraint in a particular environment as paradigmatic of the difference between the two languages and treat 'the other cases' as residual quirks of the grammar of a particular item. However, for the child it is not clear what the paradigmatic cases are, especially in cases where there is a need for a delicate triangulation of form, interpretation and syntactic and semantic context. We would like to suggest that they have to proceed cautiously guided perhaps by universal properties (the fact that definites articles can be associated to kinds, for example), and distributional properties of the competing items.

If, as we are assuming, definites in both languages allow IC, the problem for the child is not one of determining whether the target language has or does not have IC with definites, but rather is the problem of determining its distribution. There has been a tendency in the literature on IC to focus on differences between the definite determiners. However, in contrast to the definite determiner, for which evidence of such differences is slim, it is just as likely that differences in the possessive pronoun system give rise to the crosslinguistic variation in IC. There are far easier to detect properties of the possessive pronouns and possessive constructions in both languages than what has been in generally acknowledged in the IC literature. First, English has a pre-nominal genitive, while Spanish does not. Second, the definite and the possessor are in complementary distribution in English from the syntactic point of view but not in Spanish. Third, the possessive system in English does not agree with the possessed noun but just with the possessor, unlike the Spanish possessive. Finally, it is possible to say los ojos de María "the eyes of Maria" in Spanish but not in English. These differences between the two languages will guarantee a different treatment of the possessive pronouns by the English speaking child and the Spanish speaking child. We would like to suggest that the answer for the different behaviour of the definite in both languages in IC may be much more a function of the differences between the possessives pronouns and possessive structures than differences within the definite determiners. Of course the issue remains as to what accounts for the fact that the definite determiner wins in more cases of IC in Spanish than in English, although the definite does not win in every single case of IC interpretation (e.g., body parts in subject position in Romance may require a pronominal element). Why isn't the definite or the possessive bled away by the possessive in each language or vice versa? Under our approach, one is forced to further restrict the notion of competition to competition of forms with respect to particular senses and in particular syntactic domains.

3. The development of definites in IC contexts

3.1 Previous acquisition research

Various studies have examined the referential properties of definite determiners in child language, focusing on the definite/indefinite distinction (Maratsos 1994, Schafer & de Villiers 1999) and specificity (de Villiers & Roeper 1995; Matthewson, Bryant & Roeper 2001, Schaeffer & Matthewson 2000). Only two studies, Ramos (1999) and Baauw (2000), specifically examined the interpretation of definites with inalienable possession.

Ramos' (1999) study of DP development in SLI (Specific Language Impairment) uncovered a pattern of overgeneralization of inalienable construal in English speaking children. She developed an act-out task using the character of Ms. Potato Head, an assemblable toy with interchangeable body-parts. In the story, Ms. Potato-head is surprised to find a box filled with personal items (hat, bags) and and detached body-parts (hands, noses, etc.) that look just like hers. The story introduces each item, and the child is asked to act-out a target sentence with a definite or possessed object: *she just touched the/her nose*. Error rates for the normally developing children in her study were as high as 30% for the younger (MLU matched) group (aged 3;8-4;5) and 23% percent for the older (Age matched) group (4;7-5;7). Most incorrect responses with the definite objects consisted of the acceptance of the inalienable reading (Ramos, p.c., April 2000). Her findings suggest that normally developing English-speaking children allow IC of definite determiners.

Baauw (2000) argued for morphosyntactic underspecification of the D position in the course of grammatical development, on the basis of experiments in Dutch and Spanish. Using a truth-value judgment task, he examined the comprehension of definites in two groups of Dutch children aged 4;1 to 8;0. In

one story context, three boys either touched their own noses (the distributive/inalienable context) or touched the nose of an extra character (specific, non-distributive context). This provided an alienable referent from the perspective of the subject *the boys*. The definite should be acceptable under the single character/non-distributed context, but rejected in the distributive scenario. When presented with sentences like (7), in distributive contexts, Dutch children willingly accepted non-target inalienable interpretations (i.e., boys touching each their own noses). Correct rejection of the inalienable interpretation of the definite was low for the younger group (about 30%) and higher but still not target for the older children and controls (about 70%).

(7) The three boys touched the nose.

The Spanish children (aged 4;3 to 7;1) were tested with definites in the same distributive and referential contexts as in Dutch with an additional manipulation for lexical class of verb (gestural vs. non-gestural verbs). When presented with sentences as (8a), some of the older children, and many adults rejected the grammatical referential interpretation more than half the time. Concurrently, children were perfectly willing to accept the same sentence in distributive contexts (about 2/3 of the responses). Although adults had higher rates of correct responses, the difference was not significant.

- (8) a. Los niños volvieron la cabeza. (√inalienable/√referential) "The children turned the head."
 - b. Los niños tocaron la oreja. (*inalienable/√referential) "The children touched the ear."

With respect to the lexical verb restriction in (8b) children performed at adult-like levels from the outset. However, the adult controls unexpectedly allowed the inalienable interpretation as much as half of the trials.

3.2 Acquisition predictions and goals of the present studies

The evidence from these two studies on inalienable possession suggest a non-target pattern of overextension of inalienable construal in languages where this interpretation is more restricted, like English and Dutch. However, both authors acknowledge some methodological difficulties. Ramos points out that some children had problems with the fact that her alienable and inalienable items were identical. Baauw's data show overextensions of the definite in the adult control groups. One goal of our study was to seek additional empirical

confirmation for their findings. Furthermore, we sought to examine the distribution of these interpretations in definites in comparison with possessor determiners, and to search for evidence for the potential role of number in constraining the overgeneralization. In the next section we present two studies, one in English, evaluating the contribution of number and pragmatics in children's IC with definite determiners, and a replication in Spanish that additionally included a condition examining the role of lexical verb class (gestural verbs vs. reflexive verbs) in restricting or triggering IC. Given the distributional perspective on the acquisition of the semantics of determiners presented in section 2, and of the existing differences between the two languages, we present the following predictions:

- a) IC with definites should always be possible for children, even when it is an error such as in the case of English.
- b) If this developmental stage (i.e., the possible English error) is the result of something like a parameter, then IC should show a pattern of sensitivity to the number restriction.
- c) However, if the distributional view is correct, then number sensitivity need not be a part of the definite-inalienable error.
- d) There should be a gradual emergence of lexical restrictions leading to target adult distribution (i.e. lexical verb class in Spanish).
- e) Other developmental instances of definite as kinds should appear in English.

We test a)-d) in two act-out studies in English and Spanish. Prediction e) is examined in separate work (Pérez-Leroux, Munn, Schmitt & DeIrish 2002).

4. English study

4.1 *Methods*

We developed an act-out task in a scenario where several characters were playing with Mr. Potato Head and Cootie Bugs body parts (ears, noses, heads, arms, legs and mouths). The dolls representing the characters had soft, easy-to-move limbs (3 Groovy Girls dolls and a bear and frog dolls comparable in size and materials). The storyline indicated that Frog brought his toys (toy body parts, the alienable objects) so everyone could play, as shown in (9). Since no mention was made of inalienable body parts, this presentation biased towards the alienable reading. Children were then asked to act-out a sentence with a body part noun in direct object position.

(9) This time Frog got a leg, some noses, an arm and two heads. He gave Bear an arm. He gave Joey a head. He gave Suzy a leg. He gave Mary another head. The noses, he didn't give to anyone.

Prompt: Mary moved the head.

In articulating these scenarios, the assignment of toy body parts was manipulated to control for number (plural or singular objects and subjects), for possession (whether a toy body part belonged (i.e., 'was assigned') to the relevant characters), and uniqueness (whether it was the only one of its kind in the game.) The presentation of the study was segmented into 3 different toy assignment blocks with 2 tokens for each of the 5 conditions, counterbalanced for determiners, resulting in a total of 30 trials. Unrelated activities were interspersed between blocks. We present here only conditions 1, 4 and 5, which were matched in the Spanish study. Conditions 2 and 3 manipulated contextual elements to test for the effect of felicity conditions in children's overgeneralization. The effect of context was not significant and is discussed elsewhere (Pérez-Leroux, Schmitt & Munn 2002).

Condition 1 had singular objects and singular subjects, in a scenario where use of the definite article was felicitous. That is, the toy body-part was unique (i.e. the only alienable leg in the game), and had been mentioned in the set-up. Target adult responses are included in brackets below.

(10) Condition 1: Singular-Singular Unique
Suzy put the leg on the table.

Suzy put her leg on the table.

(alienable)
(inalienable preferred)

Conditions 4 and 5 had a conjoined plural subject to test the effect of grammatical number on children's willingness to give definites an inalienable interpretation. In Condition 4, the subject and object matched in plurality, and the relevant toy-body parts belonged to two of the characters, so the possessor was ambiguous. In Condition 5 the subject was plural but the object was singular, and only one of the characters in the conjoined subject received the relevant toy object. Character pairs were always distinct in gender, to disambiguate the referent of the possessor determiner (in (12), *his*=Joey's).

(11) Condition 4: Plural subject, plural assigned object

Joey and Mary covered the heads. (alienable distributed)

Joey and Mary covered their heads. (ambiguous)

(12) Condition 5: Plural subject, Singular unique, assigned object

Joey and Mary kissed <u>the</u> arm. (alienable unique) Joey and Mary kissed <u>his</u> arm. (single toy or b.p.)

Condition 5 was constructed to be structurally parallel to the number effect of definite DPs in IC in Romance. English definites and possessors should allow an interpretation where two characters act on a single toy part. A response where two characters acted on one character's body part was considered possible but not expected for the possessive pronoun. However, neither determiner should allow an inalienable distributed interpretation such as the one allowed in the Romance IC construction.

4.2 Participants

Seventeen English-speaking monolingual children recruited in several daycares in Toronto participated in the study. The younger children (N=9) ranged in age from 3;11 to 5;2, mean 4;4. The older children (N=8) ranged from 5;5 to 6;6, mean 5;11. The adult controls were college students recruited at Michigan State University (N=12).

4.3 Results

Responses were videotaped and analyzed. For the analysis, we included only the proportion of anaphoric (SELF) response, representing inalienable readings. For the plural conditions, this was defined as having the target event with each participant character acting over his own target body part, with $X \rightarrow X$'s head, $Y \rightarrow Y$'s head, etc. Most of the other responses consisted in children acting on the assigned toys. In the plural conditions there was a small amount of reciprocal answers (1%-4%), and acting with other individuals's body parts (~5%). Overall non-response or unrelated answers was small for all conditions (3%-5%).

An overall repeated measure ANOVA on SELF responses showed significant results for all factors and interactions. Group was significant at $F_{2,26}$ =5.775, p=.008; determiner was highly significant at $F_{4,104}$ =20.200, p<.0001. All interactions were highly significant except conditions x determiner x group which were significant at $F_{8,104}$ =2.281, p<.027. A Fischer PLD posthoc analysis showed adults to be significantly different from both older and younger children (critical differences of -0.167 and -0.288, p=.002 and p<.0001, respectively). Older and younger children were significantly different with a critical difference of -.121, p=.046.

Figure 1, showing proportions of SELF responses to both determiner types, demonstrates that younger children allow a considerable proportion of inalienable errors with the definite determiner. Error rates are lower for the older children and non-existent in the adult group. All groups treated the possessor as ambiguous, with adults exhibiting a preference for the inalienable.

Conditions 4 and 5 both show higher rates of inalienable construal for definite than in the singular condition. Adults performed as predicted, allowing both interpretation of the possessor in condition 4, and uniformly disallowing distributed SELF interpretation in condition 5, as shown in Figs. 2 and 3.

The adult unwillingness to accept a distributed interpretation for either determiner contrasts with the children performance, where this interpretation was possible in both cases, despite the mismatch in number between subject and object. The statistical comparison of children's SELF responses to definites in condition 4 and condition 5 did not show significant differences. I.e., children's willingness to construe the definite inalienably is not sensitive to number matching or mismatching between subject and object as predicted by the pattern of the Romance languages (difference =.108, t_{16} =1.037, p=.315). The Spanish pattern, though not identical, shows a similar insensitivity.

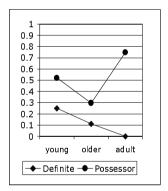


Figure 1: Proportion of SELF (inalienable interpretations) in Condition 1, Singular-Singular.

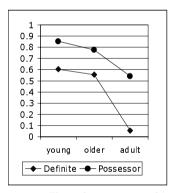


Figure 2: Proportion of SELF (inalienable, distributed interpretations) in Condition 4, Plural-Plural.

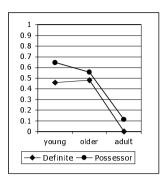


Figure 3: Proportion of SELF (inalienable, distributed interpretations) in Condition 5, Plural-Singular

4.4 Discussion

This study shows that even after methodological improvement over previous studies, English-speaking children still appear to have inalienable errors in their interpretation of definite determiners in direct object position, in a syntactic context not allowed in the target grammar. This suggests the productive availability of inalienable construal, but leaves open the question of the source of these interpretations. The fact that our data show appropriate rates of inalienable possession interpretation with possessive pronouns (i.e., sensitive to features of the context) suggests that the developmental differences observed should not be attributable to child-specific, task-specific patterns, but to developmental differences in form to sense mappings of the definite determiner. Furthermore, the analysis of the pragmatic variables of possession and uniqueness in Pérez-Leroux, Schmitt & Munn (2002), (Conditions 2 and 3 not shown here) suggested that the child's divergence does not have a pragmatic but a grammatical source.

5. Spanish Study

5.1 Methods

The Spanish study was based on the same scenarios as in the English study conditions 1, 4 and 5 reported above. The scenario was similarly set up by the Frog character bringing toy body parts for his friend to play, and with the child asked to act-out sentences with a human (plural or singular) subject, a verb and a direct object with a (ppdefinite+body part noun) structure. The manipulations regarding number on object and subject in the three conditions

remained unchanged (we continue to use the labels Condition 1, 4 and 5 to refer to singular-singular, plural-plural and plural-singular, respectively):

(13) a. Condition 1: Oso movió la mano.

"Oso moved the hand."

(ambiguous)

Condition 4: Mari y Suzy levantaron las cabezas. b. "Marie and Suzy raised the head."

(alienable because plural)

Condition 5: Pepito y Mari sacudieron el brazo. c. "Pepito and Mari shook the arm."

(inalienable preference)

As in the English study, conditions 1 and 5 had a paired body part (legs, arms and ears) and condition 4 had a singleton body-part (heads, mouths and noses). Other grammatical factors differed across languages. In English, the possessive determiner served as the comparison base for the definite. In Spanish, because the definite is ambiguous in the target grammar (preferences for inalienable construal in Baauw's adult group notwithstanding), other control structures were necessary. We wanted to include both items that favoured inalienable construal (such as the possessive) and one that excluded it (such as the demonstrative). Based on one author's intuitions about the emphatic nature of the pronominal possessive determiner in Dominican Spanish we included as controls for the definite a structure with a definite followed by the body part noun, followed by either a demonstrative or a possessive pronoun:

(14) a. El niño movió la mano (plain definite)

El niño movió la mano suya

(possessive)

El niño movió la mano esa

(demonstrative)

"The boy moved the hand/his hand/that hand."

The possibility of IC was secured by using gestural verbs (levantar 'raise'/ subir 'lift'/ bajar 'lower'/ mover 'move'/ sacudir 'shake', etc.) in the Spanish items. In addition, we constructed a direct comparison to Condition 1 with a gestural verb (15a) that had a regular verb without the clitic (15b). This structure requires a reflexive clitic to express inalienable construal, and was expected to elicit only the alienable reading.

(15) a. Gestural verb:

Oso levantó la pierna. "Oso raised the leg."

(inalienable preferred)

b. Dative verb:

Oso rascó la pierna "Oso scratched the leg."

(alienable only)

Each of the condition/determiner combinations had 3 trials, for a total of 30 trials divided into blocks of 10. Verbs/body parts and determiners were alternated across condition to avoid specific imbalance due to lexical effects.

5.2 Participants

Twenty Spanish-speaking children were recruited in a preschool in Santo Domingo, Dominican Republic, servicing primarily upper-middle class, professional families. The control group consisted of 8 adolescents and adults recruited from comparable location and social class. The younger group of children (N=8) ranged in age from 3;2 to 4;5 (mean age 3;11), and the older group (N=12) ranged in age from 5;0 to 6;7 (mean age 5;10).

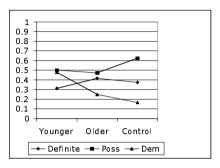
5.3 Results

The proportion of SELF responses in the Spanish study was submitted for statistical analysis separately for each condition. Specific analyses were planned for the definite in the verb condition, and to compare for the effect of number across conditions 4 and 5.

In condition 1, a repeated measures ANOVA on the proportion of SELF response to all groups reveals a non-significant effect of group, significant effect of determiner and no significant interaction ($F_{2,25}$ =.122, p=.885; $F_{2,50}$ = 4.529, p=.015; and $F_{4,50}$ =1.623, p=.183, respectively). The control group performed as predicted, with mid-range proportion of SELF responses to the definite determiner, a bias for the SELF in the case of the possessive, and a small proportion of SELF responses to the demonstrative. The younger children show no discrimination among determiners and the older children show intermediate patterns.

The comparison between condition 1 definite and the verb condition across all groups shows a non-significant effect of group, an almost significant effect of verb and a significant interaction ($F_{2.25}$ =.977, p=.39; $F_{1, 50}$ = 3.230, p=.084; and $F_{2.25}$ =5.645, p=.009, respectively). In Figure 5 we see how controls and older children attend to the lexical verb restriction in their interpretations of the definite (with error rates of 8% and 19%, respectively), but entertain the inalienable interpretation when the verb is a gestural motion verb (37% and

41%, respectively). In contrast, younger children do not discriminate between verb classes, and give more SELF interpretations to the wrong kind of verbs.



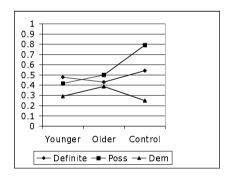
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

Younger Older Control

Dative verb —— Gestural Verb(Cond 1)

Figure 4: Proportion of SELF (inalienable) to all three DP conditions (definite, demonstrative, possessive) in Condition 1, Singular-Singular.

Figure 5: Comparison of the proportion of SELF (inalienable) in definite Condition 1 (gestural verb) and the Verb Condition (reflexive verb without clitic).



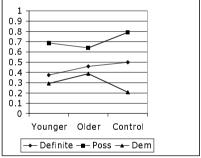


Figure 6: Proportion of SELF (inalienable, distributed) responses to all three DP conditions (definite, demonstrative, possessive) in Condition 4, Plural-Plural.

Figure 7: Proportion of SELF (inalienable, distributed) responses to all three DP conditions (definite, demonstrative, possessive) in Condition 5, Plural-Singular

The overall data of condition 4 are presented in Figure 6. The patterns of differentiation between determiners are not as clear as in condition 1, but the statistical analysis suggests a consistent behaviour: a significant effect of determiners, and a non-significant effect of group and interaction ($F_{2,25}$ =.488, p=.619; $F_{2,50}$ = 4.633, p=.014; and $F_{4,50}$ =1.426, p=.239, respectively).

The data from condition 5 presented in Figure 7 are comparable, with clearer discrimination patterns across determiners, and less marked differences between groups. The statistical analysis shows a non-significant effect of group, a highly significant main effect of verb and a non-significant interaction ($F_{2.25}$ =.058, p=.943; $F_{2.50}$ = 26.751, p<.0001; and $F_{4.50}$ =1.739, p=.156, respectively).

A planned comparison between the definites in the plural-plural and plural singular conditions shows no significant main effects or interaction. This result shows that neither the Spanish-speaking children nor the adult controls were sensitive to the effect of number in the object. The source of this result could be dialectal: Dominican Spanish generally has final [s] deletion around 30%, but in some contexts, the plural marker –s can be deleted as much as 56% (Alba 1982, Gonzalez Tapia 1994). Since plural marking is often not recoverable from phonological cues, it may become inert as a cue for online processing of sentences, masking the operation of the number constraint.

5.4 Discussion

The Spanish experiment shows development for verb type. There is a significant change with the clitic verbs and also in their use of the different types of determiners (this is clear for condition 1). The results also show that Dominican Spanish adults accept plural definites of a singleton body part with IC. Spanish children, unlike the English children, do not give substantially higher numbers of IC answers in plural conditions compared to the singular conditions. That the Spanish children do not behave identically to the English children shows that the increased use of IC in the plural cases for the English children is unlikely to be a task effect brought on by the possible increased complexity of acting out with multiple participants. Were this the case we would expect the Spanish children to be similarly biased, and they are not.

6. Conclusions

Our results replicate the findings in Ramos (1999) and Baauw (2000) in showing that English and Dutch children allowed IC of the definite determiner in non-target language contexts. We found that both English and Spanish children allow IC of the definite determiner, although not in the same proportion in the singular cases. The results support a distributional view of the acquisition of IC in which IC is not parametric in the traditional sense, but that IC is always possible since definites can denote kinds. The fact that the Spanish children allow IC more than English children in the singular case is expected since the input supports the interpretation in the former but not the latter case. Our study also shows that English children are able to differentiate possessives

from definite determiners and that Spanish children (at least for condition 1) are capable of distinguishing the three structures (the simple definite, the def+poss and the def+demonstrative).

Number was not a constraining factor in children's IC with definites, suggesting that the English error was not a form of parameter missetting. However, the fact that number did not seem to play its predicted role in the Spanish adult data, perhaps due to the variability of plural marking in Dominican Spanish leaves questions open about the nature of the constraint. But importantly, the difference between the Spanish and English children with respect to the proportion of SELF responses in the plural conditions, shows that the effect in the English children is not task related.

Finally, the data in Spanish show, as did Baauw's (2000) study, the emerging effect of lexical constraints on verb type.

Overall, the type of gradual development we see in both languages is consistent with the distributional competition view we have outlined in this paper. What is not clear at this point is the exact nature of the set containing the competing items, and the principles that cause it to be constructed and that allow certain items to be chosen as winners. We leave this very large question for future research but we suggest that a careful investigation of the properties of the possessive pronouns in both languages may shed some light on the differences in behaviour between the definites in both languages.

References

Alba, Orlando. 1982. "Función del Acento en el Proceso de Elisión de la /s/ en la República Dominicana". *El Español del Caribe: Ponencias del VI Simposio de Dialectología*, ed. by Orlando Alba. Santiago: Departamento de Publicaciones de la Universidad Católica Madre y Maestra.

Baauw, Sergio. 2000. Grammatical Features and the Acquisition of Reference. Utrecht University. Ph.D. Dissertation.

Chierchia, Gennaro. 1998. "Reference to Kinds across Languages". *Natural Language Semantics* 6. 339-405.

Dayal, Veneeta. 2002. Number Marking and (In) definites in Kind Terms. Ms., Rutgers University.

de Villiers, Jill & Thomas Roeper. 1995. "Barriers, Binding and Acquisition of the NP/DP Distinction". *Language Acquisition* 4. 73-104.

Guéron, Jacqueline. 1985. "Inalienable Possession, PRO-inclusion and Lexical Chains". *Grammatical representation*. ed. by Jacqueline Guéron, Hans Georg Obenauer & Jean-Yves Pollock. Dordrecht: Foris.

Guéron, Jacqueline. 1999. Inalienable Possession. Ms., University of Paris.

- Gonzalez Tapia, Carlisle. 1994. *El Habla Culta Dominicana*. Santo Domingo, Editoria Universitaria de la Universidad Autónoma de Santo Domingo.
- Maratsos, Michael. 1976. The Use of Definite and Indefinite Reference in Young Children. Cambridge: Cambridge University Press.
- Matthewson, Lisa, Tim Bryant & Thomas Roeper. 2001. "A Salish Stage in the Acquisition of English Determiners: Unfamiliar "Definites". *Proceedings of SULA: The Semantics of Under-Represented Languages in the Americas*. University of Massachusetts Occasional Papers in Linguistics 25.
- Pérez-Leroux, Ana T., Alan Munn & Cristina Schmitt 2002. "Syntactic Features and Discourse Factors in Children's Interpretation of Definite Determiners in Inalienable Possessions". *Proceedings of the Canadian Linguistic Association*, 245-258.
- Pérez-Leroux, Ana T., Alan Munn, Cristina Schmitt & M. DeIrish. 2004. "Learning Definite Determiners: Genericity and Definiteness in English and Spanish". *Proceedings of Boston University Conference on Language Development*, Cascadilla Proceedings Project.
- Pinker, Steven. 1984. *Language Learnability and Language Development*. Cambridge: Harvard University Press.
- Ramos, Eliane. 1999. The Syntax of NPs in SLI. University of Massachusetts. Ph. D. dissertation.
- Roeper, Thomas. 1999. "Universal Bilinguism". *Bilinguism: Language and Cognition* 2 (3). 169-186.
- Schaeffer, Jeannette & Lisa Matthewson 2000. "Grammar and Pragmatics in the Acquisition of Article systems". *UCLA Working Papers in Linguistics: Language Development and Breakdown 1* ed. by Jill Gilkerson, Misha Becker & Nina Hyams. 1-39. University of California, Los Angeles.
- Schafer, Robin & Jill de Villiers. 1999. "Imagining Articles: What *a* and *the* can Tell us about the Emergence of DP." *Proceedings of BUCLD* 24 ed. by S. Catherine Howell, Sarah A. Fish & Thea Keith-Lucas, 609-620, Sommerville: Cascadilla Press.
- Vergnaud, Jean-Roger & María Luisa Zubizarreta. 1992. "The Definite Determiner in French and English". *Linguistic Inquiry* 23. 595-652.

SEGMENTAL AND FEATURAL STRATEGIES TO AVOID ADJACENT SIBILANT SEGMENTS IN BALEARIC CATALAN * AN OPTIMALITY-THEORETICAL ACCOUNT

CLÀUDIA PONS MOLL Universitat de Barcelona

1. Introduction

The purpose of this paper is to provide an Optimality-theoretical account of the different strategies triggered to avoid adjacent sibilant segments in Balearic Catalan (BC). In BC, adjacent sibilant segments are systematically avoided. The processes triggered due to this fact differ according to the dialectal variety, the domain of application, and the featural configuration of the consonants involved.

2. Data

As it can be seen in (1a), in Majorcan Catalan (MaC), a sequence of an alveolar sibilant consonant followed by a sibilant segment is resolved through a process of manner dissimilation, which gives a sequence of a stop followed by an affricate as a result. When the first segment of the cluster is palatal, a process of gliding with independent motivations applies (1b). The consonant affected by the process of dissimilation is always the one placed in coda position.

(1) Majorcan Catalan

a. tros sencer /tros##sənsər/ [ˌtrɔt.tsən.'sə] '(the) whole piece' (cf. tros ['trɔs]) tros gelat /trɔs##ʒɛl+a+d/ [ˌtrɔd.d͡ʒə.'lat] 'frozen piece' (cf. tros ['trɔs]) b. mateix suc/məteʃ##suk/ [mə.tej.'suk] '(the) same juice' (cf. mateix [mə.'teʃ]) mateix joc /məteʃ##zɔg/ [mə.tej.'zɔk] '(the) same game' (cf. mateix [mə.'teʃ]) mateix tros /məteʃ##trɔs/ [mə.tej.'trɔs] '(the) same piece' (cf. mateix [mə.'teʃ])

^{*} This work is supported by the grants BFF2001-3798, from the *Ministerio de Ciencia y Tecnología*, 2001/SGR/00004, from the *Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya*, 1999FI, from the Generalitat de Catalunya and

In Minorcan Catalan (MiC), we find the process of manner dissimilation in the same contexts as in MaC, and when the first consonant is palatal as well (2a, b). As shown in (2c), in the varieties of MiC spoken in Ciutadella and Ferreries, a process of gliding or split is triggered when the first consonant of the cluster is palatal:

(2) Minorcan Catalan

a. All varieties

tros sencer	\reaner##sent\	[se ['] .nest.tsan	'(the) whole piece'
tros gelat	/tros##3el+a+d/	[ˌtrɔd.d͡ʒə.ˈlat]	'frozen piece'

b. Minorcan Catalan (general)

mateix joc /məteʃ##ʒəg/ [mə.ted.ˈd͡ʒək] '(the) same game'

mateix xalet /məteʃ##ʃəlet/ [mə.tet.t͡ʃə.ˈlet] '(the) same house'

mateix suc /məteʃ##suk/ [mə.tet.t͡suk] '(the) same juice'

c. *Minorcan Catalan (Ciutadella & Ferreries)*mateix joc /məte\f##309/ [mə.tej. 'zok]~[mə.tej. 'd͡ʒɔk] '(the) same game'

mateix sol /məte\f##suk/ [mə.tej. 'suk]~[mə.tej. 'tsuk] '(the) same juice'

In Eivissan Catalan (EC), a process of deletion/fusion applies, as it does occur in the rest of Catalan varieties. The resultant consonant of the process coincides with the consonant placed in second position (3a, b). If the first consonant is palatal and the second is alveolar, however, a strict process of fusion applies. The result of this process is a retracted alveolar sibilant consonant, which reflects a case of segmental coalescence (3c).

(3) Eivissan Catalan

a.	tros sencer	/tros##sənsər/	[ˌtrɔ.sən.ˈsə]	'(the) whole piece'
	tros gelat	/tros##ʒɛl+a+d/	[ˌtrɔ.ʒə.ˈlat]	'frozen piece'
b.	mateix joc	/məteʃ##ʒɔg/	[mə.te. ˈʒɔk]	'(the) same game'
c.	mateix suc	/məte∫##suk/	[mə.te.ˈsuk]	'(the) same juice'

a grant from the *Institut Menorquí d'Estudis*. I would like to express my gratitude to Maria-Rosa Lloret, who has provided me some fruitful ideas and suggestions on this subject. For valuable comments on this paper, I am also grateful to Pere Grimalt, Joan Mascaró, John J. McCarthy, Max W. Wheeler, and to the audience of Going Romance 2002. I am also indebted to two anonymous reviewers for their shrewd comments.

¹ For an analysis framed on autosegmental phonology of the strategies triggered when a palatal segment is involved, see Palmada (1994a), and for an analysis of these strategies within the framework of OT, see Pons (2003b).

Except for some unproductive cases, all these dialectal varieties show a process of epenthesis in lexical heteromorphemic sibilant clusters (4).²

(4) Balearic Catalan

```
cuses /kuz+z/ ['ku.zəs] '(you) sew' (cf. sents ['sens] '(you) hear') felices /fəlis+z/ [fə.'li.səs] 'happy (fem. plur.)' (cf. útils ['u.tils] 'useful'(pl.))
```

As it can be seen in the previous examples, the strategies triggered to avoid the adjacency of sibilant segments can differ according to the level of application, the dialectal variety, and the consonants involved. As for the level of application, we have seen that in the lexical level a process of epenthesis applies in all the dialectal varieties, whereas in the postlexical level other strategies are triggered. In this level, the type of processes triggered depends on the dialectal variety and the type of consonants involved. In MaC and MiC, we generally find a process of manner dissimilation. In contrast, EC shows a process of deletion/fusion. The nature of the consonants is also relevant for the resolution of the process: when the first consonant of the cluster is a palatal segment, processes like gliding, split, and fusion take place. This fact can be related to the resistance of palatal segments to losing their original configuration. Finally, it must be said that the syllabic position of the consonants is also a factor to be considered: the consonant affected by all these processes is the first one, that is, the consonant placed in coda position. This fact corroborates the perceptual prominence of the segments associated to the onset position, which is cross-linguistically recurrent.

3. Previous analysis

The dissimilation, the fusion or deletion, and the epenthesis processes we find in Catalan have been generally analyzed as different strategies to avoid the adjacency of identical or similar segments. This avoidance is found across a significant number of languages, and the strategies triggered are similar. Different authors (McCarthy 1986, Yip 1988) have attributed this behavior to the Obligatory Contour Principle (OCP), which bans the adjacency of identical elements. According to these authors, the OCP can act as a rule blocker and as a rule trigger. Processes like assimilation, dissimilation, epenthesis, degemination and so on are said to be triggered by the activity of this principle.

² For expository reasons, these data concerning lexical clusters will not be analyzed in the present paper. The analysis of these lexical sequences can be found in Bonet, Mascaró & Lloret (2003) and Pons (in preparation).

(5) Obligatory Contour Principle (OCP)

"At the melodic level, adjacent identical elements are prohibited." (McCarthy 1986)

The variety of processes that we find in Catalan has been largely explored within the framework of autosegmental phonology by Palmada (1994a, b). In these studies, it is argued that the principle responsible of this behavior is the OCP. Palmada (1994a) argues that the adjacent segments affected by this principle are those specified as [+continuant] and CORONAL, that is, adjacent sibilant segments according to her proposal. The process of manner dissimilation of MaC and MiC is understood as the result of a process of deletion of the discordant features (i.e., [+cont] and COR), and the subsequent introduction of the unmarked manner and place features (i.e., [-cont] and COR), which are assigned by two default rules. This would be a typical case of the emergence of the unmarked. The process of reduction that occurs in EC is also interpreted as the deletion of the discordant features and the subsequent deletion of the empty segmental position. Finally, the epenthesis process that applies in all varieties in lexical heteromorphemic sibilant clusters is interpreted as a process of segmental insertion of the unmarked vowel in Eastern Catalan, which prevents the adjacency of the sibilant segments and ensures the preservation of their segmental and featural properties. In addition to some specific problems we will not refer to for expository reasons, the analysis proposed by Palmada, although it is very rich in its description of the processes, fails to account for a) the causes that motivate the triggering of one strategy or another according to the dialect or domain of application, b) the consonant affected by each process, and c) the emerging coalescence effects in these dialectal varieties. In this framework, these facts can only be accounted for by resorting to stipulative arguments which do not have plausible motivation. Indeed, within this interpretation, there is no mechanism that explains why in MaC & MiC a process of dissimilation applies, and why in EC a process of reduction is triggered: just the presence or the absence of certain rules. Similarly, as the author claims, the only way to justify the consonant affected by each process is establishing "that the mechanism of verification of the OCP acts from left to right". Some further questions arise: To what extent can it be claimed that segments specified as CORONAL and [-continuant] are the unmarked in Catalan, if we take into account its range of processes (cf. consonant epenthesis due to syllabic reasons in different Catalan varieties: /tem+re/ [təm. bre]; lleó /səon/ [sə. vo])? Could it be the case that the resultant consonant of the dissimilation process of MaC and MiC is the result of the economy of language as well as a consequence of the contextual environment,

rather than a real instance of the "emergence of the unmarked"? Apart from that, the relation between the OCP principle and the rules triggered to satisfy it remains unclear.

In this paper we are going to prove that these kinds of processes and their motivations are better analyzed within the OT framework, in its version of Correspondence Theory (McCarthy & Prince 1995). The analysis we propose is that the manner dissimilation, the deletion/fusion, and the epenthetic processes that apply in BC are different strategies motivated by the activity of the *[sib][sib] constraint (7a), which bans adjacent sibilant segments.³ This constraint has an articulatory and a perceptual motivation: as pointed out in Wheeler (2003), following Kirchner (1998), sibilant segments require a greater articulatory effort than others consonants do; thus, the avoidance of adjacent sibilant segments. On the other hand, the continuancy of this kind of consonants obscurs its length perception; the reduction process could be interpreted as a consequence of that and the dissimilation process, a prevention strategy (see Boersma 1998 for further discussion on this subject).

4. Manner dissimilation in Majorcan and Minorcan Catalan

The process of manner dissimilation of MaC and MiC can be understood, in this framework, as a strategy to satisfy the *[sib][sib] markedness constraint through the minimal alteration of the consonants of the cluster. This minimal alteration explains the preservation of the segments involved as well as the maximal preservation of the manner and place features associated to these segments. The first fact is expressed through the faithfulness constraint MAX-IO, according to which every segment of the input must have a correspondent in the output. The second fact is partially explained by means of specific versions of the faithfulness constraint IDENT(F), which advocate for the preservation of the input featural specification. As we will see, the relevant constraint in our case is IDENT(-sonorant). This constraint ensures the obstruent configuration of the resultant consonant. This minimal change, however, is conditioned by contextual factors that are expressed in terms of the AGREE(place) markedness constraint, which is highly ranked in Majorcan and Minorcan Catalan (where regressive place assimilation is really common). The

³ In Bonet & Lloret (2002), where other data are analysed, the constraint adduced to explain the avoidance of adjacent sibilant segments is OCP-Sibilant, with the same effects. Wheeler (p.c.) suggests that the constraint responsible for this behavior could be *GEMINATESIBILANT. We disregard this constraint because the processes of dissimilation, fusion and deletion also affect heterorganic adjacent sibilants segments (see examples in section 2), so that another constraint should be invoked to explain these cases.

effect of this constraint favors the coronal nature of the resultant consonant, given that the second consonant of the cluster always has this specification.

In (6) we reproduce a provisional ranking for MaC and MiC, where some basic constraints have been considered (7). The tableau in (8) shows the effects of this ranking on a sequence such as *pos sal* '(I) put salt'.

(6) Provisional ranking for Majorcan and Minorcan Catalan *[sib][sib], ALIGN-Words, ALIGN-Prefix >> MAX-IO >> DEP-IO

(7) Basic constraints

- a. *[sib][sib]: Adjacent sibilant segments are prohibited (see Bonet & Lloret 2002).
- b. ALIGN-Words (ALIGN-W): The right edge of the word must be aligned with the left edge of another word (See McCarthy & Prince 1993; Dols 2000, Bonet & Lloret 2002 for Catalan).
- c. ALIGN-Prefix (ALIGN-Pref): The right edge of the prefix must be aligned with the left edge of the stem (See McCarthy & Prince 1993; Dols 2000, Bonet & Lloret 2002 for Catalan).
- MAX-IO (MAX): Every element in S1 has a correspondent in S2 (McCarthy & Prince 1995).
- e. DEP-IO (DEP): Every element in S2 has a correspondent in S1 (McCarthy & Prince 1995)

(8) pos sal /pɔz##sal/ [pɔt. tsal]4

/pɔz##sal/	*[sib][sib]	ALIGN-W	MAX	DEP
a. [pos. sal]	*!	1 1 1		
b. [ˌpɔ.z <u>ə</u> .ˈsal]		*!		*
c. [pɔ.ˈsal]		! ! !	*!	
☞d. [pɔt.ˈsal]		1		

As can be seen in (8), the most faithful candidate, (8a), with two adjacent sibilants, fatally violates the *[sib][sib] markedness constraint. The candidate (8b), with epenthesis, is discarded by the constraint ALIGN-W, according to which the right edge of a word must be aligned with the left edge of another

⁴ For expository reasons, we disregard, for the moment, the actual candidate in MaC and MiC with an affricate in onset position ([pot. tsal]). This candidate will be introduced later on.

word.⁵ This constraint is highly-ranked in most Catalan varieties, since epenthesis never takes place postlexically. The faithfulness constraint MAX, which bans the deletion of a segment present in the input and which is ranked above DEP (see, for example, Bonet & Lloret 1996 for a justification of this ranking in Catalan) rules out the candidate (8c), with deletion of the first consonant. The candidate selected as the optimal is (8d), that is, the candidate that shows manner dissimilation.

The ranking proposed so far, however, is not complete enough to obtain the desired results, since any candidate with a non-sibilant segment in coda position could be selected as the optimal. In order to obtain the actual candidate as the winner, it is necessary to introduce some new constraints. On the one hand, we must consider the activity of the faithfulness constraints that regulate featural changes. These constraints are responsible for the minimal change adduced before. The constraints relevant for the purposes of our data are IDENT(sibilant), according to which correspondent segments must have the same specification for the feature [sibilant] (9a), and IDENT(-sonorant), according to which an input [-sonorant] segment must also be [-sonorant] in the output (9b). In BC, IDENT(sib) is ranked higher than IDENT(-sont), since sibilant segments usually do not lose their featural configuration for markedness reasons, but other obstruents do.

Another constraint is necessary to achieve the actual candidate, i.e., AGREE(place). As stated in (9c), AGREE(place) requires that adjacent segments have the same place of articulation; as said before, the effect of this constraint favors the coronal nature of the resultant consonant, since the second segment of the cluster has always a coronal specification. This constraint, on the other

⁵ A reviewer points me out that if we consider that the epenthetic vowel is part of the base, ALIGN-Words would not be violated, so that another constraint such ALIGN-Stem should be invoked. In fact, this is not true because the dissimilation process also affects adjacent sibilant segments where the first consonant is not part of the stem (cf. *coses sabudes /kɔz+ə+z##sab+u+d+ə+z/ [ˌkɔ.zət.x̄sə.'βu.ðəs]* 'known things', where the first *s* is the plural morph), so that a candidate such as [ˌkɔ.zə.zə.səˈβu.ðəs] would vacuously satisfy ALIGN-Stem. This constraint, though, could be adduced to discard a candidate with epenthesis as a part of the base, although it could not be high-ranked because in Catalan there is final epenthesis due to syllabic reasons (*centre /sentr/ [ˈsentra]; centre petit /sentr##patit/ [ˌsen.tra.pə.'tit]*, where the epenthetic vowel can be considered part of the stem). See McCarthy (2003) for an extensive discussion of these aspects.

⁶ We consider candidates with non-obstruent consonants in coda position because some languages show drastic featural changes to avoid specific configurations, what has been called 'overkill'; in some varieties spoken in Ghana, for instance, /r...r/ sequences are resolved through a process of dissimilation that turns the first /r/ into [t]. See Struijke & Lacy (2000) for extensive discussion on this subject.

hand, impedes that the process of dissimilation gives a labiodental fricative ([f], [v]) as a result; in fact, this type of consonants share more features with sibilants (stridency and continuancy) than stops do. In BC, where regressive place assimilation is really common, this constraint occupies a high position, concretely above IDENT(-sont) and below IDENT(sib). It should be noticed, on the other hand, that the potential effects of IDENT(place) are inhibited by AGREE(place), which, in Balearic Catalan, is ranked above.

In order to obtain the actual candidate, it is crucial that MAX dominates IDENT(sib) because otherwise the candidate with deletion, which vacuously satisfies IDENT(sib), would be selected as the optimal. Moreover, the ranking of *[sib][sib] above IDENT(sib) is also crucial, because it guarantees the non-sibilant character of the consonant in coda position. As can be seen in (11), the inclusion of these new constraints ensures the selection of the candidate with a stop in coda position

(9) Required faithfulness and markedness constraints

- a. IDENT(sibilant) (IDENT(sib)): Correspondent segments must have the same specification for [sibilant] (See McCarthy & Prince 1995).
- IDENT(-sonorant) (IDENT(-sont)): Correspondent segments must have the same specification for [-sonorant] (See Pater 2001, McCarthy & Prince 1995).
- AGREE(place) (AGREE(pl)): Adjacent consonants must share the place of articulation.

(10) New provisional ranking for MaC and MiC *[sib][sib], ALIGN-W, ALIGN-Pref >> MAX >> **IDENT(sib)** >> **AGREE(pl)** >> **IDENT(-sont)** >> DEP

(11) pos sal /pɔz##sal/ [pɔt. tsal]

/pɔz##sal/	*[sib][sib]	ALIGN-W	Max	IDENT	AGREE	IDENT	DEP
				(sib)	(pl)	(-sont)	
☞a. [pɔt.ˈsal]				*			
b. [pɔp.ˈsal]				*	*!		
c. [pɔr.ˈsal]				*		*!	
d. [pɔl.ˈsal]				*		*!	
e. [pɔm.ˈsal]				*	*!	*	
f. [pɔf.ˈsal]				*	*!		

An additional constraint is required to discard a candidate with change of manner of articulation of the second consonant of the cluster (*[pos.'tal]). The impossibility of such realizations is justified by the positional faithfulness constraint IDENTONSET(F), according to which the consonant in onset position must preserve the place, the manner, and the voice features of its correspondent in the input. The relevant constraint in this case is IDENTONSET(sibilant), which requires that the segment placed in onset position maintains the sibilant specification of its correspondent in the input and that allows the process of affrication of the segment placed in onset position.

(12) Required positional faithfulness constraint

IDENTONSET(sibilant) (IDENTONS(sib)): The segment placed in onset position must preserve the sibilant specification of its correspondent in the input (See Beckman 1998, Lombardi 2001).

(13) New provisional ranking for MaC and MiC

IDENTONS(sib), *[sib][sib], ALIGN-W, ALIGN-Pref >> MAX >> IDENT(sib) >> AGREE(pl) >> IDENT(-sont) >> DEP

Before concentrating on the analysis of the behavior of EC, we should consider another set of candidates, which we have left apart because of their complexity. These are the candidates in (14), among which we can find the actual candidate in MaC and MiC, (14c), and the candidate until now considered the actual, (14e). We can also find a candidate with fusion of both consonants to an alveolar sibilant consonant, (14a); a candidate with fusion to an affricate alveolar consonant, (14b); and a candidate with change of manner of articulation of the first consonant and syllabification of $/s_1/$ and $/s_2/$ in onset position (14d).

(14) Other candidates

a. [pɔ.ˈs ₁₂ al]	1) Fusion of $/s_1/$ and $/s_2/$ to $[s_{1,2}]$; 2) Preservation of manner of articulation of $/s_1/$ and $/s_2/$; 3) Syllabification of $/s_1/$ and $/s_2/$ in onset position
b. [pɔ. ts ₁₂ al]	<i>1</i>) Fusion of $/s_1/$ and $/s_2/$ to $(\overline{ts}_{1,2})$; 2) Change of manner of articulation of $/s_1/$ and $/s_2/$; 3) Syllabification of $/s_1/$ and $/s_2/$ in onset position
c. $[p i t_1 \cdot \widehat{ts}_2 al]$	1) Change of manner of articulation of $/s_1/$ and $/s_2/$; 2) Syllabification of $/s_1/$ in coda position and $/s_2/$ in onset position (actual candidate)
d. [pɔ. ˈt ₁ s ₂ al]	1) Change of manner of articulation of $/s_1/;$ 2) Syllabification of $/s_1/$ and $/s_2/$ in onset position
e. [pɔt _{1.} s ₂ al]	1) Change of manner articulation of $/s_1/$; 2) Syllabification of $/s_1/$ in coda position and $/s_2/$ in onset position

As seen in (15), the candidates with fusion, (15b, c), are discarded thanks to the activity of ALIGN-W. The candidate with change of manner of articulation of the first segment and syllabification of the cluster in onset position, (15e), is discarded because of the activity of IDENTONS(sib). However, two candidates fare even: the candidate until now considered the optimal (15a) and the actual candidate (15d).

1	5١	** ** **	~~1	/mag##cal/	[pot. tsal]
	.))	DOS	Sal	/ D.JZ##Sa.I/	TDDL ISALI

/pɔzı##s₂al/	IDENTONS (sib)	*[sib][sib]	ALIGN-W	Max	IDENT (sib)
6 ^{**} a. [pɔt₁. s₂al]		1 1 1 1	1 1 1 1		*
b. [pɔ. s ₁₂ al]		1	*!		
c. [pɔ. ts ₁₂ al]			*!		
☞ d. [pɔt₁. t͡s₂al]		1	1		*
e. [pɔ. t ₁ s ₂ al]	*!	! ! !	1 1 1		*

In order to achieve the actual output, (15d), it is necessary to introduce a new constraint, SYLLABLECONTACT, which demands that heterosyllabic adjacent clusters show the same or a decreasing degree of sonority.

(16) Required (syllabic) markedness constraint

SYLLABLE CONTACT LAW (SYLLCONT): Adjacent heterosyllabic segments must show the same or a decreasing degree of sonority (Vennemann 1988, Clements 1990; Bonet & Mascaró 1995, Jiménez 1997, 1999 and Pons 2003a for Catalan)

The syllabic transition in the candidate [pot.'sal] (15a) shows an increasing degree of sonority, so that it is discarded thanks to the constraint SYLLCONT. As justified in Pons (in preparation), where SYLLCONT is invoked to explain regressive manner assimilation, this constraint is ranked below IDENT(sib) and above AGREE(place).

(17) New ranking for MaC and MiC

IDENTONS(sib), *[sib][sib], ALIGN-W, ALIGN-Pref >> MAX >> IDENT(sib) >> SYLLCONT >> AGREE(place) >> IDENT(-sont) >> DEP

5. Deletion and fusion in Eivissan Catalan

As said before, in EC sibilant contacts are resolved through a process which gives a single sibilant as a result. Apart from the clusters integrated by a palatal and an alveolar segment, there is no further evidence whether in

Eivissan a process of reduction (which consists on the elision of the first consonant) or a process of fusion of the two segments to one applies. Given the fact that we do not have more empiric evidence than that related to the contacts of a palatal sibilant followed by an alveolar sibilant, we can hypothesize that it can be triggered either a process of fusion or a process of deletion of the first segment. In Optimality Theory, this circumstance can be expressed through the lack of dominance between two constraints that are in conflict with respect to the deletion and fusion processes. This fact motivates the selection of two candidates as optimal: the candidate with fusion and the candidate with deletion. Before we move on, we should see first some basic differences with respect to the ranking proposed for MaC and MiC. As shown in (18), the constraints IDENTONSET(sibilant) and *[sib][sib] are ranked at the top of the hierarchy. The first one guarantees that the change of manner of articulation of the second consonant is not a possible strategy to avoid adjacent sibilant segments in this variety; the second one avoids the presence of two adjacent sibilants. A basic difference between EC and MaC and MiC is the position of IDENT(sib). This constraint must be placed as high as possible in EC in order that the dissimilation process does not apply. The optionality between fusion and deletion can be explained through the lack of dominance between MAX, which disfavors the candidate with deletion, and the alignment constraints, ALIGN-W and ALING-Pref, which rule out the candidates with fusion. In the next tableau, it can be observed that the lack of dominance between both constraints ensures a tie between the candidate with deletion (19c) and the candidate with fusion (19e).

(18) IDENTONS(sib), *[sib][sib] >> **IDENT(sib**) >> **ALIGN-W**, **ALIGN-Pref**, **MAX** >> DEP

(19) *pos sal* /pɔz##sal/ [pɔ.ˈsal]

/pɔz ₁ ##s ₂ al/	IDENTONS	*[sib][sib]	IDENT	ALIGN-W	Max	DEP
	(sib)		(sib)			
a. [pɔs ₁ . s ₂ al]		*!				
b. [pɔ.z ₁ <u>ə</u> . s ₂ al]				*		*!
☞ c. [pɔ. s₂al]					*	
d. [pɔt ₁ . s ₂ al]		1 1 1	*!			
☞ e. [pɔ. s _{1,2} al]		1		*		
f. [pɔs ₁ . t ₂ al]	*!	1	*			

Things become difficult, however, when the second segment of the cluster does not show the same place of articulation as the first one. We should explain why in the sequences of an alveolar sibilant followed by a palatal sibilant this last consonant is the one preserved. In the tableau (20), we see that the wrong candidates are selected as the optimal, and this is because we do not have any constraint that regulates the featural specification of the fused segment, nor any that determines which consonant is deleted in case of deletion. Apart from the actual candidates, (20b), with deletion of the first consonant, and (20d), with fusion of the two segments to a palatal segment, candidates such as (20c), with deletion of the second consonant, (20e), with fusion to an alveolar segment, or (20f), with fusion to a retracted alveolar, are also selected as the optimal.

$/des_1##\int_2 ifr + a + r/$	IDENTONS (sib)	*[sib][sib]	IDENT (sib)	ALIGN- Pref	Max	DEP
a. [dəs₁.∫₂i. ˈfɾa]		*!				
ுb. [də.∫₂i.ˈfɾa]		! ! !			*	
6 *c. [də.z₁i. fra]					*	
				*		
6 **e. [də.s _{1,2} i. ¹fra]				*		
● *f. [də. <u>s</u> _{1,2} i. ˈfɾa]				*		

An important aspect to be considered before analyzing these data is that when dealing with fusion and deletion, the reference to the onset position is not an available strategy to regulate the featural changes. Positional faithfulness constraints such as IDENTONS(F) prescribe that when one segment is placed in onset position it must show the same featural specification as its correspondent in the input. Given the lack of underlying syllabification, these kinds of constraints are not sufficient when there is deletion or fusion. For example, a candidate such as (20c), with deletion of the second consonant and syllabification of the first consonant in onset position, satisfies this positional faithfulness constraint, since it shows the same featural specification as its correspondent in the input. In Optimality Theory, there are different ways to regulate the consonant which must be deleted and the featural specification that the candidate with fusion must acquire. One of them is appealing to the OUTPUT-OUTPUT faithfulness constrains (Benua 1997), as proposed in Bonet & Lloret (2002) for Central Catalan, and adapted to EC in Pons (2003a). Another strategy, pointed out to me by an anonymous reviewer, could be resorting to targeted constraints (Wilson 2001). This kind of approach, however, could deal

only with cases of deletion, but not with cases with fusion. Finally, the fact that is always the second consonant of the cluster the one imposed could be explained by means of a positional faithfulness constraint that targets the initial segment of a root-domain (21a), as pointed out to me by another anonymous reviewer. This constraint rules out, as it can be seen in (22), the candidate with deletion of the second consonant.

As we have seen before, however, in this variety a process of fusion is also possible, so that it is necessary to introduce some more constraints to explain the palatal character that the candidate with fusion acquires. One of these constraints is IDENT(-ant), according to which an [-anterior] input segment must also be [-anterior] in the output (21b). IDENT(-ant) captures the universal resistance of palatal segments to losing their original configuration in spite of being placed in weak syllabic position, such as coda. In fact, a constraint like IDENT(pal) would do the same job. As shown in the tableau (23), this constraint rules out a candidate like (23f), with fusion to an alveolar segment. Finally, the constraint IDENT(place)-INITIALROOTC, ranked at the bottom of the hierarchy, is responsible for the avoidance of a candidate with fusion of the two consonants to a retracted alveolar sibilant segment (23g).

(21) Required I-O and positional faithfulness constraints

- a. MAX-INITIALROOTC (MAX-INRC): The initial segment of a root must be preserved in the output.
- b. IDENT(-anterior) (IDENT(-ant)): Correspondent segments must have the same specification for [-anterior] (See Pater 2001, McCarthy & Prince 1995).
- IDENT(place)-INITIALROOTC (IDENT(pl)-INRC): The featural specification
 of the initial segment of the root must be preserved in the output.

(22) New provisional ranking for Eivissan Catalan

$$\label{local_cont} \begin{split} &\text{IdentOns(sib), Max-InRC, *[sib][sib]} >> &\text{Ident(sib), } \textbf{Ident(-ant)} >> &\text{Align-Pref, Max} >> &\text{Dep} >> &\text{Ident(pl)-InRC} \end{split}$$

d. $[də.z_1i. fra]$ • e. $[də.\int_{1,2}i. fra]$ f. $[də.s_{1,2}i. fra]$

g. [də.s₁ 2i. fra]

(23) desxifrar /dəs##fiir+a+i/ [də.fi. ira]									
Max-	*[sib][sib]	ID	Id	ALIGN-	Max	DEP	IDENT		
INRC	i ! !	(sib)	(-ant)	Pref	! !		(pl)-		
	! !						INRC		
	*!		:						
	! ! !			*	! !	*!			
	1 1 1				*				
	Max-	MAX- *[sib][sib] INRC	MAX- *[sib][sib] ID INRC (sib)	MAX- *[sib][sib] ID ID (sib) (-ant)	MAX- *[sib][sib] ID ID ALIGN- INRC (sib) (-ant) Pref	MAX- *[sib][sib] ID ID ALIGN- MAX INRC *!	MAX- *[sib][sib] ID (sib) ID ALIGN- MAX DEP *!		

*

(23) desxifrar /dəs##ʃifr+a+r/ [də.ʃi.ˈfra]

6. Dissimilation and gliding in Minorcan Catalan

As introduced before, the clusters integrated by a palatal segment followed by a sibilant behave differently. In most MiC dialects, a process of dissimilation applies. Therefore, it can be adduced the same ranking proposed to account for the clusters with an alveolar segment in first position. In some varieties of MiC, those spoken in Ciutadella and Ferreries, a process of gliding is triggered. This process can be understood, similarly to the process of fusion of EC, as a strategy to avoid the adjacency of sibilant segments without losing the palatal configuration of the first segment of the cluster. IDENT(-ant) is the constraint responsible for this behavior, and it must be ranked below AGREE(place), because in these varieties the segments specified as [-anterior], such as the velar stop, assimilate to the following consonant, losing their [-anterior] specification.

(24) Ranking for Ferreries & Ciutadella IDENTONS(sib), *[sib][sib], ALIGN-W, ALIGN-Pref >> MAX >> IDENT(sib) >> SYLLCONT >> AGREE(place) >> **IDENT(-ant)**, IDENT(-sont) >> DEP

(25) mateix sol /məteʃ##səl/ [mə.tej.'səl]

/məte∫ı##s₂ɔl/	*[sib][sib]	IDENT	SYLLCONT	AGREE	IDENT	IDENT
		(sib)		(pl)	(-ant)	(-sont)
a. [mə.tet ₁ . ts_2 51]		*			*!	
☞ b. [mə.tej ₁ . s ₂ əl]		*				*

7. Fusion in Eivissan Catalan

Let us focus now on the behavior of EC when the first segment of the cluster is palatal and the second is alveolar. We have seen that in these cases a strict process of fusion which gives as a result a retracted alveolar sibilant applies.

As shown in the tableau of (26), with the ranking given in (22) we do not obtain the desired results, since candidate (26c) is wrongly selected as the optimal. And this is because the actual candidate, (26h), fatally violates IDENT(pl)-INRC. Moreover, the selected candidate, (26c), vacuously satisfies the IDENT(-ant) constraint, so that the apparent prominence of palatal segments is lost.

((26)	mateix sol /məte	(##sal/	[mb te 'sb]	1
١	40.	muicix soi mate	THEODY	IIII d. IC. SUI	

/məte∫ ₁ ##s ₂ ɔl/	Max- InRC	*[sib] [sib]	ID (-ant)	ID (sib)	ALIGN- W	Max	DEP	ID (pl)- INRC
a. [mə.te∫₁. ˈs₂ɔl]		*!				į		
b. [mə. te.∫ <u>1ə</u> . s ₂ ɔl]					*		*!	
6 *c. [mə.te. s₂ɔl]						*		
d. [mə.te.ˈʒ ₁ ɔl]	*!	1 1 1				*		
e. [mə.te. s _{1,2} 5l]		:	*!		*			
f. [mə.te.'∫ _{1,2} ɔl]		!			*	!		*!
g. [mə.te. s _{1,2} 5]					*	:		*!
h. [mə.tej ₁ . s ₂ əl]		, ! !		*!				

An additional constraint is necessary to discard a candidate like (26c). This constraint is MAX-FINALROOTC, according to which the final consonant of a root in the input must have a correspondent in the output. This constraint, placed in the same position as the alignment constraints and above IDENT(place)-INRC guarantees the discarding of a candidate with deletion of the final consonant of the base. The final ranking proposed for EC in (27) predicts that deletion only applies if the first segment is not part of a root, which is a desirable result if we take into account the greater prominence of stems and roots in relation to prefixes. The activity of this constraint explains the systematic application of fusion when the first consonant is palatal, because in Catalan there are no prefixes ending in a palatal segment. That is to say, all the cases where a palatal segment meets another consonant are instances of sequences of words and compounds.

- (27) Required positional faithfulness constraint
 MAX-FINALROOTC (MAX-FINRC): The final consonant of a root must have a correspondent in the output.
- (28) Definitive ranking for EC
 IDENTONS(sib), MAX-INRC, *[sib][sib] >> IDENT(sib), IDENT(-ant) >> ALIGN-Words, ALIGN-Pref, MAX-INRC >> MAX >> DEP >> IDENT(pl)-INRC

(29) mateix sol / mate (##sol / [ma.te. s_{1.2}ol])

/məte∫₁##s₂ɔl/	Max-	*[sib][sib]	ID	ΙD	ALIGN	Max-	MAX	DEP	ID(pl)-
	INRC	; !	(-ant)	(sib)	-W	FINRC			INRC
a. [mɔ te∫₁. s₂ɔl]		*!							
b. [mə.te.∫ ₁ <u>ə</u> . s ₂ əl]		! !			*	! !		*!	
c. [mə.te.s ₂ ɔl]		! !		:		*	*!		
d. [mə.te. 'ʒ ₁ ɔl]	*!	! ! !					*		
e. [mə.te.s _{1,2} əl]		! ! !	*!	<u> </u>	*	1 1 1			
		 - -		<u>!</u>	*	: - -			*
g. [mə.tej ₁ s ₂ əl]		!		*!		; ; ;			

To sum up, the analysis we propose for EC is that a process of deletion of the first consonant or fusion between the two consonants can apply when the first consonant is alveolar. This fact is captured, as seen, through the lack of dominance between two constraints that are in conflict with respect to deletion and fusion processes: ALIGN-W, ALIGN-Pref and MAX-FINRC. This unranked hierarchy motivates the selection of two candidates as optimal: the candidate with fusion and the candidate with deletion. Two positional faithfulness constraints determine the segmental quality that the fusion candidate acquires and the consonant that is preserved in case of deletion. Another positional faithfulness constraint, MAX-FINRC, prevents deletion when it would imply the loss of the final segment of a root. The activity of this constraint explains the lack of deletion and the systematic application of fusion when the first consonant of the cluster is palatal.

References

Beckman, Jill. 1998. *Positional Faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst. Available on Rutgers Optimality Archive-234. Benua, Laura. 1997. *Transderivational Identity: Phonological Relations between Words*. Doctoral dissertation, University of Massachusetts, Amherst. Available on Rutgers Optimality Archive-259.

- Boersma, Paul. 1998. Functional Phonology. Formalizing the Interactions between Articulatory and Perceptual Drives. The Hague: Holland Academic Graphics.
- Bonet, Eulàlia & Maria-Rosa Lloret. 1996. "Variació Dialectal i Optimitat: Epèntesi en el Grup Clític". Talk presented at the 6th Colloquium on Generative Grammar, València.
- Bonet, Eulàlia & Maria-Rosa Lloret. 2002. "OCP Effects in Catalan Cliticization". *Catalan Journal of Linguistics* 1. 19-39.
- Bonet, Eulàlia & Joan Mascaró. 1995. "On the Representation of Contrasting Rhotics". *Issues in the Phonology and Morphology of the Major Iberian Languages* ed. by Fernando Martínez Gil & Alfonso Morales Front, 103-126. Washington, D.C.: Georgetown University Press.
- Bonet, Eulàlia, Joan Mascaró & Maria-Rosa Lloret. 2003. "Phonology-Morphology Conflicts in Gender Allomorphy: a Unified Approach". Handout 26th Glow Colloquium, Lund University.
- Clements, G. Nick. 1990. "The Role of the Sonority Cycle in Core Syllabification". *Papers in Laboratory of Phonology I. Between the Grammar and Physics of Speech* ed. by John Kingston & Mary Beckman, 283-333. Cambridge: Cambridge University Press.
- Dols, Nicolau. 2000. *Teoria Fonològica i Sil·labificació. El Cas del Català a Mallorca*. Doctoral dissertation, Universitat de Mallorca.
- Jiménez, Jesús. 1997. *L'Estructura Sil·làbica del Dialecte Valencià*. Doctoral dissertation, Universitat de València.
- Jiménez, Jesús.1999. *L'Estructura Sil·làbica del Català*. València & Barcelona: Institut Interuniversitari de Filologia Valenciana & Publicacions de l'Abadia de Montserrat.
- Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. Doctoral dissertation, University of California Los Angeles.
- Lombardi, Linda. 2001. "Why Place and Voice are Different: Constraint-Specific Alternations in Optimality Theory". *Segmental Phonology in Optimality Theory* ed. by Linda Lombardi, 13-45. Cambridge & New York: Cambridge University Press.
- McCarthy, John J. 1986. "OCP Effects: Gemination and Antigemination". Linguistic Inquiry 17. 207-263.
- McCarthy, John J. & Alan Prince. 1993. "Generalized Alignment". *Yearbook of Morphology 1993* ed. by Geert E. Booij & Jaap van Marle, 79-153. Dordrecht: Kluwer.
- McCarthy, John J. 2003. "OT Constraints are Categorical", ms., University of Massachusetts, Amherst.

- McCarthy, John J. & Alan Prince. 1995. "Faithfulness and Reduplicative Identity". *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics* 18, ed. by Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk, 249-384. Amherst.
- Palmada, Blanca. 1994a. *La Fonologia del Català. Els Principis Generals i la Variació*. Bellaterra & Barcelona: Servei de Publicacions de la Universitat Autònoma de Barcelona.
- Palmada, Blanca.1994b. "Sobre la Semblança Excessiva en Fonologia". *Els Marges*, 49. Barcelona: Publicacions de l'Abadia de Montserrat.
- Pater, Joe. 2001. "Austronesian Nasal Substitution Revisited". *Segmental Phonology in Optimality* ed. by Linda Lombardi, 159-182. Cambridge & New York: Cambridge University Press.
- Pons, Clàudia. In preparation. Els Contactes Consonàntics en Balear. Descripció i Anàlisi. Doctoral dissertation, Universitat de Barcelona.
- Pons, Clàudia. 2003a. "Teoria de l'Optimitat, Processos Segmentals i Variació Dialectal en Català". *Zeitschrift für Katalanistik* 16. 55-76.
- Pons, Clàudia. 2003b. "Internal Palatal Segments: a Cross-Dialectal Analysis". Paper presented at the *First International Conference of Phonetics and Phonology in Iberia*, Lisbon University.
- Struijke, Caro & Paul De Lacy. 2000. "Explaining Overkill in Dissimilation". Handout NELS 31, Georgetown University.
- Vennemann, Theo. 1988. *Preference Laws for Syllable Structure*. Berlin, New York & Amsterdam: Mouton de Gruyter.
- Wheeler, Max W. 2003. "La Lenició en la Fonologia Catalana: Perspectives Teòriques i Empíriques arran dels Etnotextos de l'ALDC". *Interlingüística* 13. 83-100.
- Wilson, Colin. 2001. "Consonant Cluster Neutralisation and Targeted Constraints". *Phonology* 18, 147-197.
- Yip, Moira. 1988. "The Obligatory Contour Principle and Phonological Rules: A Loss of Identity". *Linguistic Inquiry* 19. 65-100.

-TU IN QUEBEC FRENCH AS A (SUPER)POSITIVE MARKER*

MARIE-THÉRÈSE VINET Université de Sherbrooke

1. Introduction

This paper proposes a unified analysis for the invariable item -tu in the grammar of Quebec French (henceforth QF), an enclitic which appears in root finite clauses with a variety of discursive or expressive effects. In these contexts, -tu is a bound morpheme which has a limited distribution in the sense that it is confined to finite verbs with overt subjects and it is precluded in non expressive root clauses. Moreover, when -tu cooccurs with the negative marker pas, it always deprives pas of its negative force. Traditionnally and even in recent theoretical studies, -tu in QF has been identified almost uniquely as an interrogative marker (cf. Noonan 1989 and many others cited in Vinet 1999: 383, footnote 2) but it is not the only role it plays. The main questions to be addressed are therefore the following:

- A) What is the function of -tu in this grammar?
- B) Why does -tu appear only in these identified environments and not in others?
- C) How can we license in a unified way all these expressions with -tu in QF?

^{*} I would like to thank the organizers of the Going Romance Conference for their warm welcome as well as the audience at the colloquium and at the Conference presented at Institut CharlesV in Paris (*Conférences du lundi*, December 2nd 2002) for their comments and their interest. I would also like to thank Anne Zribi-Hertz, Nomi Erteschik-Shir and François Poiré for discussions and comments on a previous version of this article. This research was partly supported by SSHRC grant #410-2001-0119 (Martineau-Vinet) and #410-2000-0954 (Tellier-Vinet).

In order to account for these expressions I will use certain elements from two different theoretical approaches. Similar structures, mostly interrogative forms in French and in Romance, have recently been studied through a Remnant IP approach in a layered CP, as in Kayne & Pollock (2001) and Poletto & Pollock (2000), a framework which accounts for V related and phrasal movements merged to the left, including Subject Clitic Inversion (SCII) and Complex Inversion (CI). From this point of view, -tu structures, just like ti, can be considered as a variety of CI structures. However, -tu structures in OF do not only involve yes/no question forms as in (1d) below. -Tu can also appear in evaluative-exclamative structures (1a) as well as change of state expressive constructions (1b) and alleviated forms of command structure using a modal verb (1c). In order to illustrate how the function of -tu is to signal an affirmative or a (super)positive context, since -tu always appears in positively polarized contexts, I will use certain aspects of an f-structure model, as developed by Erteschik-Shir (1997). The four identified contexts in which -tu can appear are exemplified in (1):

- (1) a. *Ça se peut-tu!* that REFL-CL can- TU "Can it be possible!"
 - b. *D'un coup, elle part-tu pas à crier* suddenly, she starts-TU Neg yelling "Suddenly, she starts yelling."
 - c. Tu veux-tu me laisser tranquille! you want-TU me leave peaceful "Why don't you leave me alone!"
 - d. Elles sont-tu responsables?3FPL are-TU responsible"Are they responsible?"

The paper is organized as follows. In the next section, I begin by summarizing previous results on the study of tu and ti as well as related forms of -tu in time and space. I then turn to the syntactic and semantic properties of this invariable enclitic and show that -tu must be merged to the left of IP, as an affirmative operator. Section 3 discusses the -tu pas eventive expressive structures in QF. In section 4, I present certain problems for the analysis of -tu within the framework proposed for ti by Poletto & Pollock (2000). Section 5 then explores the phonological aspect of Focus for -tu expressions in QF within the analysis of Erteschik-Shir (1997). The last section concludes briefly.

2. Previous studies and related forms in time and space

Previous studies on *-tu* (cf. references cited in Vinet 2000 and more particularly Noonan 1989) have analysed or identified this enclitic in QF as a yes/no interrogative marker only. The fact that it can also be an operator with a variety of expressive interpretations, as illustrated in (1), has usually been left aside, apart from my own work on the subject, namely Vinet (2000, 2001).

-Tu has often been studied, as in Pollock (2000), in a parallel fashion with the invariable ti form from a moribund dialect spoken in France or in Europe. Poletto & Pollock (2000) have identified the -t in ti as a "[question] morpheme" which is attracted in the derivation to delete an "uninterpretable [+interrogative] feature". However, contrary to the ti forms, the -tu constructions in QF are very productive in a variety of contexts.

Moreover, I want to demonstrate that all these interpretations in (1) have in common an emphatic affirmative feature which is incompatible with a negative force reading, a situation which illustrates clearly that the function of *tu* in this grammar is to signal a positively polarized context.

The empirical object of this study, QF -tu, is a quite recent grammatical phenomenon in Quebec. It developed in urban areas by the end of the first half of the 20th century (cf. Morin 1985, Picard 1992). Before that period, dictionaries, popular novels and plays which reported on the vernacular only mentioned the ti forms (Clapin 1894, Dionne 1909, Glossaire du Parler français au Canada 1930, Dunn 1880). However, it is rarely mentioned that the presence of such an enclitic form has also been noticed in certain French "patois" in France (Creuse, Nièvre, Puy de Dôme) (cf. Foulet 1921:341, cited in Vinet 2001: 42), as in (2a) and also (2b) from the non corrected text of the 17th century Agréables conférences, reported in Deloffre (1999). A first person enclitic form jou has also been used in yes/no questions (cf. 2c) in a moribund Picard dialect from the beginning of the 20th century, the Picard dialect of Demuin (see Hrkal 1910: 262, cited in Vinet 2001:32). Such invariable pronominal markers can also be found in Valdôtain (cf. Roberts 1993, Pollock 2000):

- (2) a. Tu kreye tu?/Tu kre tu? (Foulet 1921 :341) you believe TU/ you believe TU
 "Do vou believe?"
 - b. *Hé*, *d'* où *je venas-tu*, *de Nanterre?* (AC, III, note 30) (from Louis Richer, 17th c., cited in Deloffre 1999) hey, from where I came-TU, from Nanterre "Hey, where did I come from, from Nanterre?"

c. *Ils* y sont-jou? (Hrkal 1910 :262, cited in Vinet 2001) they there-CL are- JE "Are they there?"

This use of nominative subject pronouns of 1^{st} and 2^{nd} person as invariable enclitics in a variety of dialects, as illustrated in (2), must be clearly distinguished from 3^{rd} person enclitics which agree in gender and number with the lexical subject in Complex Inversion structures from Standard French. However, Complex Inversion (CI) and Subject Clitic Inversion (SCII) do share with *-tu* structures a common restriction to root contexts.

- 2.1 Syntactic and semantic properties of an invariable -tu
 The following characteristics identify its properties:
- a) -tu is an enclitic limited to root clauses;
- b) It is also confined to finite verbs with a lexical subject;
- It cannot be used with non expressive statements or 'residual V2' sentences;
- d) It is precluded with negative sentences.
- a) -tu is an enclitic form on a finite V and it is a Force operator limited to root clauses. It can indeed be used with root yes/no informative questions, as in (1d). It can also be used in other root sentences with several other interpretations where it reinforces a (super)positive reading connected to an expressive meaning, as in (1a-c).

Sentence (1a) is an evaluative sentence and -tu is translated as an intensifier. In (1b), -tu combined with pas serves to identify a segment of discourse which conveys new information and expresses unexpectedness and bewilderment. This last structure is always introduced by an aspectual change of state adverb, such as suddenly. It is observed that -tu and pas combine obligatorily in this structure as in the vlà-tu pas/vla ti pas "here is" expressive aspectual structure introducing an event in discourse (cf. Morin 1985, Vinet 2000).

Example (1c) is a root clause imperative form which is composed of a subject, the addressee, with a modal verb (*aller* "go" or *vouloir* "want") followed by a dynamic verb. I refer to such imperative clauses as 'alleviated' command structures because the command is perceived as less forceful than the injunctive form without an overt subject.

- b) -tu is confined to finite verbs with a lexical subject. It is therefore ruled out with infinitives and past participles. This also explains why -tu is rejected with real imperatives as in (3c) since they do not present an overt subject. -Tu only identifies a root Finite T which has a lexical subject.
- (3) a. *Ah! partir-tu en vacances!
 ah! (to) Leave-TU on holidays
 "If I could leave on holiday!"
 - b. *Elle a mangé-tu? she has eaten-TU
 - c. *Tais-toi-tu! keep quiet-TU
- c) -tu cannot be used with non expressive statements (4a). It is also unacceptable in 'residual V2' sentences with adverbs in initial position, as illustrated in (4b, c) or in hypothetical structures (4d). The well-formed QF sentences in (4) must therefore be used without -tu.
- (4) a. *Elle est-tu arrivée en effet. (statement) she has-TU arrived, indeed "She has arrived indeed."
 - b. *Sans doute elle est-tu malade.
 no doubt she is-TU ill
 - c. *Peut-être il va-tu revenir bientôt.
 maybe he go-TU come-back soon
 "Perhaps he will come back soon."
 - d. *Elle serait-tu arrivée de bonne heure que je l' aurais she would-TU arrive early that I 3FS-CL would jamais laissé rentrer.

 never let come in "Even if she had arrived early, I would never had let her in."
- d) -tu is precluded with negative sentences. In (5b), -tu deprives pas of its negative Force. This is illustrated by the fact that a negative polarity item like qui que ce soit "anyone" cannot be licensed when -tu is present. Without the negative polarity item, as in (5c), the sentence bears an acceptable expressive meaning in discourse:

- (5) a. Elle a pas renvoyé qui que ce soit. "She didn't fire anyone."
 - b. *Fak là, elle a-tu pas renvoyé qui que ce soit.
 all of a sudden, she did-TU not fire anyone
 "All of a sudden, she sent back anyone."
 - c. Fak là, elle l' a-tu pas renvoyé. all of a sudden, she 3MS-CL has-TU not sent back "Suddenly, she sent him back."

Considering the above characteristics and to answer our first question concerning -tu, I therefore propose the following descriptive generalization for -tu in OF:

(6) Descriptive generalization for -tu

-Tu in QF is an affirmative marker closely connected to a root Finite T with a lexical subject. Its function is to signal the existence of affirmative or super positively polarized contexts.

The affirmative reading associated with -tu therefore explains why it is always ruled out with a negative Force marker in informative yes/no questions as in (7) which can only be interpreted as gibberish in QF:

(7) *Ta mère est-tu pas là? your mother is-TU Neg there? "Is your mother not there?"

2.2 -tu as an operator

Let us recall that Noonan (1989) has proposed that -tu in QF is an operator. Within her framework, it was presented as an in situ operator on I. In this section, I present two other arguments for the operator status of -tu. An argument which illustrates clearly that -tu must be treated as an operator in QF is the following. Sportiche (1995) has shown that questions without a change in word order in French but with a rising intonation cannot license a negative polarity item such as qui que ce soit "anybody" in (8c):

- (8) a. *Il a vu qui que ce soit. (statement) he has seen anybody
 - b. A-t- il vu qui que ce soit?/Il n'a pas vu qui que ce soit. Has he seen anybody / He has not seen anybody

c. *Il a vu qui que ce soit? (question) he has seen anybody

However, when -tu is present in QF yes/no questions, both the negative polarity item qui que ce soit "anybody" or the positive polarity item quelqu'un "somebody" can be licensed. This situation follows from the well known fact that the polarity of yes/no questions can be either positive or negative:

If, as generally assumed, negative polarity items must be c-commanded by an operator [+ Neg] or [+Q] on their left, then the derivation in (8c) crashes because the negative polarity item is not c-commanded by any operator. Otherwise, in (9a), the operator -tu, identified as a Question operator, is merged to the left of IP and correctly c-commands the polarity item in object position. In other words, -tu in (9a) seems to have the same effect the inverted subject clitic operation has on the licensing of the sentence in (8b). This effect is clearly absent in (8c). The [+Q] operator in both (8b) and (9a) therefore 'neutralizes' the polarity of the sentence and the indefinite qui que ce soit in these sentences is therefore interpreted in both cases as the positive form quelqu'un. Let us recall that in standard logic representation, an operator must have scope from a left peripheral position and the domain of this operator is always to its right, as illustrated in (10):

b.
$$[Q]$$
 [$_{\text{IP}}$ A - t - il vu qui que ce $soit?$]

Another argument, based on evaluative structures with *-tu* as opposed to *donc*, runs as follows. In QF evaluative structures, *-tu* can be replaced by *donc*, an independent and stressed morpheme also translated by the degree word *so* in English. *Donc* always appears on the right of the inflected verb as in (11a). However, both forms cannot cooccur in the same clause with the same strong intonation (represented in capital letters) on TU and DONC, as the unacceptable example in (11b) illustrates:

(11) a. C'est donc choquant!

"It is so schocking!"
b. *C' est-TU DONC choquant!

it is-TU so schocking

The clash between -tu and donc in (11b) indicates that they play the same role in terms of focal stress. However, only -tu is an operator. In (12), only the sentence with the operator -tu is interpreted positively, the same sentence with donc is not. Since donc is not an operator which has scope from a left peripheral position, it is possible for the negation marker to bear a Negative Force reading and to negate the adjectival constituent in (12b). Donc is therefore not identified as an affirmative operator like -tu.

(12) a. C'est-tu pas gentil! it is-TU neg nice "How nice!"

b. C'est donc pas gentil! it is DONC Neg nice "How unpleasant!"

2.3 -tu as an affirmative operator

The operator -tu always combines with positively or superpositively polarized clauses. This explains why -tu in QF always appears as a positive reinforcer. Furthermore, when the negation marker pas combines with it, it always lacks negative Force. This situation is indeed observed in evaluatives (13a), it is also possible in alleviated command structures (13b) but, contrary to these last two structures, pas is always obligatory in aspectual expressive structures exemplified in (13c, d):

- (13) a. C'est-tu pas choquant! it is- TU neg schocking "It is so schocking!"
 - b. Tu vas-tu pas sortir de d' là! you go-tu Neg get out of there "Are you going to get out of there!"
 - c. Fak là, il part-tu *(pas) à crier. this has as a result that there, he starts-TU (Neg) to yell "Suddenly, he starts yelling."

d. Vlà-tu *(pas) Jean qui arrive. here-is-TU (Neg) Jean who arrive s "Here is Jean arriving."

It is well known that evaluative-exclamatives are superpositive clauses which can never be negated. Command structures, and sudden change of state expressions also, can never be negated. If *-tu* is an affirmative operator, the lack of negative Force in these structures is therefore expected.

-Tu in Question forms presents a slightly different situation. In order to illustrate my point I will use Erteschik-Shir's (1997) f-structure approach. The focus-structure theoretical model "assumes a theory of discourse which defines the state of the common ground both before and after the utturance of a sentence" (Erteschik-Shir 1997: 3). Yes/no questions can have many different f-structures. For instance, in sentences like the ones in (14), stressed low pitched tones L* in American English are associated with a request to choose between two entries:

(14) a.
$$Did\ he_{TOP}\ [see\ Susan]_{FOC}$$
 (Erteschik-Shir 1997)
 $L^*\ L^*$
 b. $Did\ he_{TOP}\ laugh_{FOC}$
 L^*

In answers to yes/no informative interrogative -tu clauses, it is the affirmative option that is stressed, as in (15), where the focal stress in on TU along with the predicate with which it forms a prosodic unit:

```
(15) Tu_{\text{TOP}}[viens-TU]_{\text{FOC}}?
you come-tu
"Are you coming?"
```

The question in (15) sets up a contrastive set consisting of the affirmed predicate and the negated predicate. The answer selects freely one of these two. It can then be claimed that the context with TU in (15) includes such a contrastive set and chooses the affirmative option of the question. As mentioned by Nomi Erteschik-Shir (p.c.), the function of -tu would be twofold: 1) to signal the existence of such a context and 2) to choose the affirmative option. Note that the affirmative option selected by the question form does not entail that the answer to the question will necessarily be positive. The intonation pattern of the question rather serves to indicate that one expects a

positive answer to the question. A reading which is not found within a verbsubject inverted yes-no question like *Vient-il?* "Is he coming?".

Moreover, in a disjunctive question with -tu in (16), it can be noted that the second clause can only appear with a negation marker pas, illustrating once again that -tu corresponds to the affirmative counterpart. Structures like (16a) in QF could perhaps be interpreted in a parallel fashion to disjunctive questions such as Tu viens ou quoi? "Are you coming or what?" found in a familiar European French (Anne Zribi-Hertz p.c.).

- (16) a. Tu viens- tu ou tu viens pas? you come-TU or you come Neg? "Are you coming or not?"
 - b. ?*Tu viens-tu ou tu viens- tu pas?
 you come-TU or you come-TU Neg
 (acceptable for some QF speakers)
 - c. Elle va-tu toujours à Montréal ou elle y va pus? she goes-TU always to Montreal or she there-cl goes no more "Is she still going to Montreal or is she not going any more?"
 - d. *Elle va-tu toujours à Montréal ou elle y va-tu pus? she goes-TU always to Montreal or she there goes-TU no more

Recall that the presence of the negation marker is completely ruled out in yes/no informative questions with -tu. The sentence in (17a) is interpreted as gibberish in QF, unless a conditional or an Irrealis tense on V turns it into a question with a different reading (17b) where pas rather lacks Negative Force. These results are expected within our analysis of -tu as an affirmative or superpositive marker.

- (17) a. *Ta mère est-tu pas là? your mother is-TU Neg there "Is your mother not there?"
 - b. Ta mère serait-tu pas là, par hasard? your mother would-be-TU Neg in, by chance "Would your mother be in, by any chance?"

3. -Tu pas in Eventive expressive structures and stage topic adverbs

Let us discuss in more detail the semantic type involved in eventive expressive structures with this emphatic affirmative operator which conveys new information in discourse. The following structures correspond to the more familiar *vla ti pas* "here is" constructions of regional French which also obligatorily appear with a marker *pas* that lacks Negative Force. Note that in these last sentences the sudden change of state situation is made explicit through the defective perception verb *voilà*.

Aspectual structures with -tu must correspond to an eventive reading. By definition, an event refers to a change of state, and in this case, more precisely, such structures refer to a sudden transition of one state into another in discourse. They do not possess a negation as a contrary. In other words, if a yell is an event, there is no non-yell:

- (18) a. *D'un coup, il part-tu pas à crier.* "Suddenly, he starts yelling."
 - b. *D'un coup, il part pas à crier.
 "Suddenly, he does not start to yell."

Events take place and they are located in time. The temporal location is expressed by tense and frame adverbials. Events which are felicitous with *-tu pas* must be measured out. They can be measured out or delimited by aspectual prefixes (19b) or by arguments in the verbal projection which refer to the property of an event bearing an inherent endpoint (Tenny 1987) (cf. 20b):

- (19) a. *Tout d'un coup, elle dort-tu pas. suddenly, she sleeps-TU neg
 - b. Tout d'un coup, elle s'endort-tu pas. suddenly, she falls-asleep-TU Neg "Suddenly, she falls asleep."
- (20) a. *Fak là, il court-tu pas. it so happens (that) there, he runs-TU Neg
 - b. Fak là, il court-tu pas au bout de la rue. it so happens (that) there, he runs-TU Neg to the end of the street

The adverbs which can introduce such change of state expressions are adverbs which would be identified by Cinque (1998) as higher sentence adverbs. Note that such change of state adverbs can also appear without -tu in Quebec French. In this case, however, the predicate requires more of a constrastive context, as illustrated in (21):

(21) *Tout d'un coup, elle PART à crier.* "Suddenly, she starts yelling."

Adverbs which are not change of setting adverbs, such as *generally* (22a), evaluative adverbs (22b), modal adverbs (22c) or edging expressions like *mais* in (22d) are not felicitous with *-tu*:

- (22) a. *Généralement, elle part-tu pas à crier. generally, she starts-TU Neg yelling
 - b. *Malheureusement, elle part-tu pas à crier. unfortunately, she starts-TU Neg yelling
 - c. *Probablement, elle part-tu pas à crier. probably, she starts-TU Neg yelling
 - d. *Mais tu l'as-tu pas ton passeport!

 (cf. Mais tu l'as ton passeport!)

 but you it have-TU Neg your passport

If -tu pas expressions are only licensed with preposed stage topic aspectual adverbs or equivalent forms, this explains why structures in (22) as well as structures in (4) above are all ill-formed. As discussed in Vinet (1999, 2000), static state verbs or predicates are always ruled out with -tu pas. Because tu pas is always related to a dynamic situation where change is involved, it is naturally compatible with verbs of achievements (discover), verbs expressing an inchoation of activity (begin to sing) or an inchoation of accomplishment (begin to build a house). These expressions all translate the 'here-and-now' of the discourse situation.

4. TI/TU and the Remnant movement approach

Poletto & Pollock (2000) and Pollock (2002) argue that ti is subdivided into two distinct morphemes. They identify -t as an H^o morpheme of main clause interrogatives or a finite tense morpheme merged in H^o which bears an uninterpretable [+interrogative] feature and attracts to its specifier a constituent marked in the same way. This is a first flaw in their analysis since they do not take into account the other discourse functions of ti which can also appear in exclamative-evaluative structures (C'est-ti $b\hat{e}te$! "How silly") or in change of state contexts such as ($Les\ vla\ ti\ pas\ ...$ "Here they are..."). In their framework, the input structure is then as in (23b) for an interrogative sentence like (23a):

- (23) a. Elle viendra-ti? (Poletto & Pollock :2000) she will+come-TI "Will she come?"
 - b. [HP [Ho -t [FP1 i [FP2 elle [FP3

Space limitation prevents us from discussing in detail their analysis which poses certain problems for the analysis of *tu* (cf. Poletto & Pollock 2000, Pollock 2002). Following their analysis, a form like *Elle va-tu lui prêter son char?* "Is she going to lend him her car?" in QF would need an input like the following:

(24) [tu [elle [va [lui prêter son char]]]

The problem is that this type of ordering with -tu in (24), based on data from Northern Italian dialects discussed in Poletto (2000), is counter-intuitive in French since the input presented in (24) never corresponds to a well-formed sentence in French, contrary to what is observed in the relevant Italian dialects.

Another main question concerning the *-tu* structures is the type of feature motivating movement since *-tu* bears several discourse functions. Following the remnant movement approach, *-tu* would need to be derived in a lower CP projection and would then need to combine with another projection in order to obtain the right interpretation as an evaluative-exclamative, an alleviated command structure or a sudden change of state situation in discourse.

This type of solution for -tu (pas) structures is therefore not obvious. Moreover, it does not capture the basic intuition that -tu corresponds to a single phenomenon when it is identified as an affirmative marker. I also refer the reader to a discussion by Newmeyer (2002), who has identified theoretical problems posed by a split CP approach. He then refers to Erteschik-Shir (1997) and more recent work by this author who proposes an alternative solution based on a system which derives intonation from focus structure rather than projecting focus structure from intonation.¹

5. Intonational properties of -tu (pas) expressions and f-structure

In this last section, I would like to explore the phonological aspect of Focus in connection with the intonational properties of *-tu* expressions of QF, an aspect which has been very poorly studied in the literature. It is interesting to observe that all structures with *-tu* present a similar f-structure in the sense of Erteschik-Shir (1997) where *-tu* or the predicate identifies the Focus and the subject is taken as the topic. The topic is what the statement is about and stress is assigned to the focus constituents. As discussed earlier, *-tu pas* structures are

¹ For a different perspective, see also Costa (2002), who discusses problems posed by a remnant movement analysis for VOS sentences with flat intonation in Portuguese.

different in that they must obligatorily present stage topics. It is also claimed that a sentence in discourse has only one main focus which is assigned to a syntactic constituent and this constituent can be an NP, a VP and sometimes the whole S.

Furthermore, it can be observed that all of the equivalent sentences without -tu in (1) require more of a contrastive context than the ones with -tu:

- (25) a. Qu' elle est BELLe! that she is beautiful "She is so beautiful!"
 - b. Tu VEUX me laisser tranquille!
 you want me leave quiet
 "Why don't you leave me alone!"
 - b. *D'un coup, il PART à crier.* of a sudden, he starts to yell "Suddenly, he starts yelling."
 - d. Tu VIENS?
 you are coming
 "Are you coming?"

The stress within the *-tu* sentences can be identified as follows in the following sentences:²

- (26) a. Elle est-TU intelligente! "She is so intelligent!"
 - b. Tu veux-TU me laisser tranquille! "Why don't you leave me alone!"
 - c. *D'un coup, il part-tu PAS à crier.* "Suddenly, he starts yelling."
 - d. Tu viens-TU?

 "Are you coming?"

² Thanks to François Poiré for helpful discussions on the intonational patterns in these QF examples.

Another interesting aspect is raised by the f-model. In a model of f-structure, Topic and Focus are annotated on the relevant syntactic constituents. The constituents are rearranged to meet the assignment of the focus and topic structures. It can therefore be observed that in a sudden change of state structure, as in (26c), the stress is on PAS, contrary to the other structures. As illustrated in (27), the prosodic unit therefore becomes [V+tu+pas]. It would be tempting to connect this unexpected prosodic unit here to the fact that *pas* is obligatory with *-tu* in this context, a situation which is not explained in a syntactic derivation. In other words, this prosodic constituent does not have a corresponding constituent in syntactic structure: f [part-tu] in (27) can be identified as a syntactic V constituent, [part-tu pas] cannot:

(27) *Il* [part-tu PAS] à crier. he starts-TU Neg to yell

In the f-model, this is accounted by Prosodic Incorporation or PI, a P-syntactic correlate of cliticization. PI is a phonological process. It is said to apply to unstressed constituents which then form a constituent with a preceding adjacent host.

To conclude this section, it must be pointed out that this research is part of ongoing research. More information is needed on word ordering in the f-structure model as well as a more precise study of the intonation pattern in QF.

6. Conclusion

In this research, I have identified *-tu* in QF as an affirmative operator merged to the left of IP, contrary to a previous claim by Poletto & Pollock (2000) who have proposed an analysis of *ti* and *tu* in a more or less parallel fashion, as interrogative markers only. It was rather claimed that the various contexts in which *-tu* appears all have either a superpositive polarity reading or identify with a question oriented towards an affirmation in the mind of the speaker, in the case of yes/no questions. Based on work by Erteschik-Shir (1997), it was here proposed that the function of *-tu* is therefore to signal such a context of affirmation. The *-tu pas* structures where *pas* is obligatorily present on surface also present a similar function and signal superpositively polarized contexts. These last structures only differ in that a) in order to be correctly licensed, they must rely on an overt stage topic, namely a preposed aspectual change of setting adverb and b) they present a different prosodic intonation.

References

- Cinque, Guglielmo. 1998. *Adverbs and Functional Heads*. Oxford: Oxford University Press.
- Clapin, Sylva. 1894. Dictionnaire canadien-français ou lexique-glossaire des mots, expressions et locutions ne se trouvant pas dans les dictionnaires courants et dont l'usage appartient surtout aux Canadiens-français. Montréal: Beauchemin.
- Costa, João. 2002. "VOS in Portuguese: Arguments against an Analysis in terms of Remnant Movement". mimeo. Universidade Nova de Lisboa.
- Deloffre, Frederick. 1999. "Agréables conférences de deux paysans de Saint-Ouen et de Montmorency sur les affaires du temps". Text written by Louis Richer (1649-1651). Critical Edition. Genève: Slatkine Reprints.
- Dionne, Narcisse-Eutrope. 1909. Le parler populaire des Canadiens-français. Ouébec.
- Dunn, Oscar 1880. Glossaire franco-canadien et vocabulaire de locutions vicieuses usitées au Canada. Québec: A. Côté.
- Erteschik-Shir, Nomi 1997. *The dynamics of Focus Structure*. New York & Cambridge: Cambridge University Press.
- Foulet, Lucien. 1921. "Comment ont évolué les formes de l'interrogation?". *Romania* XLVII .243-358.
- Glossaire du parler français au Canada. 1930. Québec: Action sociale.
- Hrkal, Émile. 1910. "Grammaire historique du patois picard de Démuin". Revue de philologie française et de littérature 24.118-119, 175-204, 241-277.
- Kayne, Richard & Jean-Yves Pollock. 2001. "New Thoughts on Stylistic Inversion". *Subject Inversion and the Theory of Universal Grammar*, ed. By Aafke Hulk & Jean-Yves Pollock, 107-162. New York: Oxford University Press.
- Morin, Yves-Charles 1985. "On the two French Subjectless Verbs *voici* and *voilà*". *Language* 61.777-820.
- Newmeyer, Frederick. 2002. "'Interpretable' Features, Feature-driven A-bar Movement and the Syntax-Discourse Interface". Glow Newsletter abstract, Tilburg University.
- Noonan, Máire B. 1989. "Operator Licensing and the Case of French Interrogatives". *Proceedings of WCCFL* 8.315-330. Stanford: Stanford Linguistics Association.
- Picard, Marc. 1992. "Aspects synchroniques et diachroniques du *tu* interrogatif en québécois". *Revue québécoise de linguistique* 21:2.65-75.
- Poletto, Cecilia. 2000. The Higher Functional Field, Evidence from the Northern Italian Dialects. New York: Oxford University Press.

- Poletto, Cecilia & Jean-Yves Pollock. 2000. "On the Left Periphery of some Romance Wh-Questions". mimeo. University of Padua & Université de Picardie à Amiens.
- Pollock, Jean-Yves. 2000. "Subject Clitics, Subject Clitic Inversion and Complex Inversion: Generalizing Remnant Movement to the Comp Area". mimeo. Université de Picardie à Amiens.
- Pollock, Jean-Yves. 2002. "Three Arguments for Remnant IP in Romance". unpublished manuscript. Université de Picardie à Amiens.
- Roberts, Ian. 1993. "The Nature of Subject Clitics in Franco-Provençal Valdôtain". *Syntactic Theory and the Dialects of Italian*, ed. by Adriana Belletti, 319-353. Turin: Rosenberg & Sellier.
- Sportiche, Dominique. 1995. "Sketch of a Reductionist Approach to Syntactic Variation and Dependencies". *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, ed. by Hector Campos and Paula Kempchinsky, 356-398. Washington D.C.: Georgetown University Press..
- Tenny, Carol. 1987. Grammaticalizing Aspect and Affectedness. Dissertation. MIT.
- Vinet, Marie-Thérèse. 1999. "-Tu pas : Aspect et Temps". Langues et grammaire 3: Syntaxe (Communications présentées au Colloque Langues et Grammaire III, Paris 1997, réunies par P. Sauzet). Université de Paris 8. 181-195.
- Vinet, Marie-Thérèse. 2000. "Feature Representation and -tu (pas) in Quebec French". Studia Linguistica 54.3:381-411.
- Vinet, Marie-Thérèse. 2001. D'un français à l'autre : la syntaxe de la microvariation. Montréal: Fides.

FEATURE CHECKING AND OBJECT CLITIC OMISSION IN CHILD CATALAN AND SPANISH*

KEN WEXLER, ANNA GAVARRÓ & VICENÇ TORRENS MIT. U. Autònoma de Barcelona & U. Nacional de Educación a Distancia

1. Introduction

In this paper we set out to substantiate by reference to two closely related languages, Catalan and Spanish, the claim that object clitic omission in child grammar has a non-accidental correlation with participle agreement. We argue that the correlation follows from the fact that in participle agreement languages objects need to double check with two functional projections; this double checking is at the source of clitic omission, following Wexler's (1998, to appear) Unique Checking Constraint (UCC). Catalan and Spanish constitute a relevant term of reference as their grammars are very similar except that Catalan is a participle agreement language while Spanish is not. Therefore, if our assumptions on participle agreement are correct, the UCC leads us to expect differences between the two in object clitic omission in child language development. As our expectations are born out, we are able to claim that variation in the development of the two languages under scrutiny can be accounted for on the grounds of a universal principle, the UCC, together with the parochial properties of the languages the child is exposed to.

-

^{*} We would like to thank Susagna Tubau for her help in running the experiment in Catalan, and the direction, teachers and children of the Escola Decroly de Barcelona for their collaboration; we would also like to thank Almudena Zurdo for her help in running the experiment in Spanish, Marisé Gil de Gómez for helping us find schools in the Madrid area, and the primary schools of Gallipatos and La Cañada. We are grateful for helpful comments to two anonymous reviewers and to the audiences of the workshop on language acquisition of Going Romance 2002 and the 5th Conference on the Acquisition of Spanish and Portuguese at the University of Iowa. The second and third author acknowledge the financial support of project BFF2000-0403-C02-02 and Generalitat de Catalunya respectively.

2. Basic assumptions: clitics and the UCC

We assume that clitics are verbal agreement morphemes (Uriagereka, 1995) that are base generated to the left of the verb. The evidence presented in this paper does not help us distinguish between syntactic theories of clitic doubling; the base-generation theory seems to cover a great deal of syntactic ground, and we assume that theory for concreteness. Following Sportiche (1996), clitics are heads of their own projection. We assume them to be embedded in a clause structure like that in (1) below (where CIP = Clitic Phrase, and $Agr_OP = Agreement Object Phrase$):

(1)
$$CIP$$

Spec CI'
 XP^2
 T
 Agr_0P
 Agr_0
 V
 V
 DP
 XP^1
 Pro

The specifier of the clitic phrase is a landing site of the coreferent DP, which is XP¹ at the base-generated position; XP¹ moves to the specifier of the clitic projection, XP². The XP matches in features (person, number, gender, Case) with the coindexed clitic.¹

With regard to the central hypotheses in the paper, we follow Wexler's (to appear) UCC principle:

(2) Unique Checking Constraint (on children in Optional Infinitive stage): the D-feature of DP can only check against one functional feature.

¹ Crucially, with accusative clitics – which are the object of this paper – the specifier of the Acc CIP is an A-bar position (under Sportiche's 1996 assumptions this is not so for e.g. dative clitics). So, *pro* raising to Spec, CIP is an instance of A-bar movement, leading to no minimality violations. As pointed out by an anonymous reviewer, Catalan and Spanish are (to varying degrees) clitic doubling languages, where sentences such as Catalan *Jo la veig a ella* (I cl-see her) are well-formed. The strong pronoun in such constructions would not raise to Spec, CIP; the possibility of clitic doubling is parametrised – see Sportiche 1996 for details.

The UCC acts in conjunction with 'Minimise Violations':

(3) *Minimise Violations*: Given an LF, choose a numeration the derivation of which violates as few grammatical properties as possible. If two numerations are both minimal violators, either one may be chosen.

Together, (2) and (3) grant the Optional Infinitive Stage (see Wexler 1998 for a full account), and also the possibility of object clitic omission for the same developmental period, as detailed below.

The UCC and Minimise Violations interact in the following way. An accusative DP in a language without participle agreement only has to check an uninterpretable D-feature (a feature of definiteness) in CIP (Clitic Phrase); thus the pro object, in a clitic construction, raises through Agr₀ to CIP where the definiteness feature is eliminated, and the derivation converges. In such a language no object clitic omission is expected. On the other hand, in a participle agreement language, a pro object has two D-features against which it must check: one in Agr_O (a case feature), one in Cl (the definiteness feature). In a given derivation, if double checking occurs, the UCC is violated (and the clitic then surfaces as in the adult grammar). If, on the contrary, the UCC is not violated and no double checking occurs, the derivation cannot converge with two unchecked uninterpretable features in CIP and Agr_OP: the only way for the derivation to converge is that one of the functional categories is not projected. Supposing CIP is eliminated, the derivation does not crash, but no clitic can be spelled out in CIP; pro moves to Agr_O and checks against the case features there – and clitic omission results, the only violation having taken place being the interface condition on the projection of CIP.

Comparing the two converging derivations, one involves a violation of the UCC, the other the interface condition that projection of CIP must occur; thus both derivations involve one violation. By Minimise Violations, the two derivations are equally bad – or equally good; both derivations are in competition and, as a consequence, optionality of clitic omission is granted.²

With respect to the languages under scrutiny, it should be pointed out that Catalan and Spanish clitic placement follow the same pattern: clitic pronouns precede finite verbs (4a, 5a) and follow non-finite verbs (4d, 5d):

² The UCC is conceived as a developmental principle, which dies out in the process towards an adult grammar; as expressed in Wexler (1998). This runs contrary to the common assumption that aspects of grammar that are innate are present from very early on (Elman et al., 1996; Quartz and Sejnowski, 1997). Rather, it is possible that innate mechanisms emerge late due to maturation.

la pentino (4) a. b. *pentino (Catalan) -la CL comb.1SG comb.1SG -CL *la pentinar c. d. pentinar -la CL comb.INF comb.INF -CL (Spanish) (5) a. la peino *péinola h. CL comb.1sG comb.1s.cl. *la peinar peinarla c. d. CL comb.INF comb.INF.CL

As indicated, in Catalan, unlike in Spanish, there is (optionally) participle agreement with a preceding direct object clitic in the perfect tenses. In most Catalan dialects there is no participle agreement for unaccusative verbs (for more details, see Cortés 1992).

- (6) La Marta les ha trobades/trobat the Marta CL-FEM-PL has found.FEM-PL/found "Marta has found them."
- (7) Marta las ha encontrado/*encontradas. (Spanish)
 Marta CL-FEM-PL has found/found.FEM-PL
 "Marta has found them."

Example (6) illustrates agreement with a feminine, plural object clitic; the morphologically unmarked form of the clitic is masculine, singular. In the Barcelona dialect, spoken by the subjects of our experiment, overt participle agreement is not as pervasive as in other varieties, but is certainly retained as a possibility, especially with feminine plurals. We follow Kayne in considering agreement between direct object and past participle as an instance of Spec-Head agreement (Kayne, 1989). To account for participle agreement in Catalan, we assume that the head of Agr_OP has an active uninterpretable feature. In Spanish, lack of participle agreement with a preceding direct object clitic results from the object passing through an Agr_O projection with no uninterpretable features.

3. An experiment on clitic elicitation

In order to test whether children produce or omit clitics in Spanish and Catalan, and whether children produce the correct clitic forms in both languages, we performed an elicitation task with 31 monolingual Catalan speaking children and 28 monolingual Spanish speaking children. We have grouped the children in 3 age groups: 2 year-olds, 3 year-olds and 4 year-olds,

on a cross-sectional design. This is the distribution of children by language and age:

(8) Catalan

age	age range	mean age
1-2 year-olds: 8	1;10 to 2;11,24	2;3,5
3-year-olds: 11	3;0,8 to 3;11,29	3;6,7
4-5 year-olds: 12	4;3,1 to 5;1,0	4;6,27
total: 31		

(9) Spanish

age		age range	mean age
2-year-olds:	8	2;6,7 to 2;11,6	2;8,18
3-year-olds:	10	3;5,2 to 3;11,13	3;7,14
4-year-olds:	10	4;4,9 to 4;11,23	4;8,13
total:	28		

The Catalan-speaking children lived in Barcelona and the Spanish-speaking children lived in suburbs of Madrid. All the children spoke the standard variants of their language. The age range is considered relevant because it is known that at these ages children speaking other languages, like Italian, omit clitic pronouns, in a period which coincides with the Optional Infinitive Stage.

The elicitation task, closely resembling that of Schaeffer (2000), comprised a tale performed with puppets, where an experimenter introduced the characters and told a story to the child. A second experimenter gave an incorrect continuation of the story, which the child was to correct. In (10) we present the Catalan version of the task – the Spanish version was the same, only some characters were changed, but we kept the number and gender of the characters. The context given strongly favoured a clitic object, rather than a full DP object: the object was known and had just been mentioned.

(10) Object clitic elicitation task for present tense

Experimenter 1: — Aquí hi tenim la Caputxeta Vermella. El rei la troba i pensa: "Mira si va despentinada!". I com que té una pinta, mira què fa.

Experimenter 2: – Ja sé què fa: renta la Caputxeta.

Experimenter 1: – No! Digues-l'hi tu: Què li fa el rei a la Caputxeta?

EXPECTED RESPONSE: - La pentina.

Experimenter 1: — Here we have Little Red Ridinghood. The king finds her and thinks: "Look what a mess her hair is!". And as

he has a comb, look what he does.

Experimenter 2: - I know what he does: he washes Little Red Ridinghood.

Experimenter 1: - No! You tell her: What is the king doing to Little Red Ridinghood?

EXPECTED RESPONSE: - He is combing her.

(11) Object clitic elicitation task for present perfect

Experimenter 1: – Aquest matí el cuiner ha començat a preparar el dinar. Ha agafat el trencanous i les nous i mira què ha fet.

Experimenter 2: – Ja sé què ha fet: s'ha menjat les nous.

Experimenter 1: - No! Digues-l'hi tu: Què ha fet el cuiner amb les nous?

EXPECTED RESPONSE: – *Les ha trencades*.

Experimenter 1: — This morning the cook started preparing lunch. He took the nutcracker and the walnuts and look at what he did.

Experimenter 2: - I know what he did: he ate the walnuts.

Experimenter 1: - No! You tell her: What did the cook do with the walnuts?

EXPECTED RESPONSE: — He broke them.

For each language two sentence types were tested: sentences in the present tense (4 items), and sentences in the present perfect (4 items). Verbs were transitive, and all the expected responses for the child consisted in a clitic pronoun and a verb. For the Catalan version, the clitics triggered optional participle agreement for gender and number. All children of all ages went through the same items.

4. Experimental results

In this section the results of the experiment are considered with respect to (i) the placement of clitics, (ii) the frequency of clitic presence vs. clitic omission, (iii) the actual production of participle agreement in the present perfect task in Catalan, and (iv) the morphosyntactic shape of the clitics

produced by the children when compared to the adult target. For all statistical analysis we used the Chi Square test.

4.1 *Clitic placement*

In line with what has been corroborated for quite a number of languages now (see e.g. Pierce 1992 for French, Guasti 1993/94 for Italian), our results show that children speaking Catalan and Spanish can distinguish correctly between finite and non-finite contexts: children place correctly the clitic pronoun before finite verbs and after non-finite verbs. There is a significant difference between the number of correct and incorrect contexts (P < 0.001). The table below shows the number of utterances produced by the children for each context.

(12) Catalan

	CLITIC + FINITE V	INF + CLITIC	*CLITIC + GERUND	
2-year-olds:	12	1		
3-year-olds:	57	4	1	
4-year-olds:	72	2		
5-year-olds:	14	1		
total	155	8	1	
% correct	163/164 (99.4%) corre	ct	1/164 (0.6%) incorre	ct

(13) Spanish

	CLITIC + FINITE V	INF + CLITIC	
2-year-olds:	53	9	
3-year-olds:	71	7	
4-year-olds:	72	8	
total	196	24	
% correct	196/196 (100%) correc	et 48/48 (100%) o	correct

Pre-verbal placement of clitics with finite verbs and post-verbal placement with non-finite verbs occurs with virtually no errors in Catalan and Spanish; the child is therefore sensitive to finiteness features from the earliest record and raising of finite verbs to T occurs systematically.

4.2 Clitic presence vs. clitic omission

Regarding the rate of clitic omission in present tense, we found that there is a significant difference in the number of sentences with an omitted clitic between Catalan and Spanish (P < 0.001). Children speaking Catalan omit

clitics more frequently than children speaking Spanish: in Spanish we found almost no omissions compared to Catalan. The differences between the two languages are found when all ages are collapsed (P < 0.001), and also for 2 year olds (P < 0.001), 3 year olds (P < 0.001) and 4 year olds (P < 0.001) separately. Below are the frequencies we found in Catalan and Spanish:

(14) Catalan, present tense

CLITIC	CLITIC OMISSION	FULL DP
1-2 year-olds: 7/31 (22.6%)	23/31 (74.2%)	1/31 (3.2%)
3-year-olds: 30/44 (68.2%)	11/44 (25%)	3/44 (6.8%)
4-5 year-olds: 45/47 (95.7%)	2/47 (4.2%)	0

(15) Spanish, present tense

	CLITIC	CLITIC OMISSION	FULL DP
2 year-olds:	32/32 (100%)	0	0
3-year-olds:	39/40 (97.5%)	1/40 (2.5%)	0
4 year-olds:	40/40 (100%)	0	0

These results for Spanish are consistent with those found in a study of spontaneous speech: Lyczskowski (1999) studied three Spanish speaking children (María, from 1;8 to 3;11; Juan from 2;6 to 4;11; Koki, from 1;7 to 2;11) and also found that these children very rarely omitted object clitics and rarely produced malformed or misplaced object clitics. The results of his study are summarised in (16).

(16) Spanish, spontaneous speech (Lyczskowski 1999)

	CLITIC	FULL DP	DOUBLE OBJ	MISSING OBJ	OTHER
D.O.	364	610	29	20	34
	34.44%	57.71%	2.74%	1.89%	3.22%
I.O.	355	10	35	2	16
	61%	2.39%	8.37%	0.48%	3.83%

With respect to the rate of clitic omission in present perfect tense, we found that Catalan speaking children omit clitics more frequently than Spanish speaking children. The number of sentences with omitted clitics is significantly different between Catalan and Spanish (P < 0.001). In Spanish we found only omissions for 2 year-olds. We have found differences between these languages when we collapsed all ages (P < 0.001), and also for 2 year olds (P < 0.001), 3 year olds (P < 0.001) and 4 year olds (P < 0.001) separately.

(17) Catalan, present perfect

CLITIC	CLITIC OMISSION	FULL DP
1-2 year-olds: 4/31 (12.9%)	26/31 (83.9%)	1/31 (3.2%)
3-year-olds: 30/42 (71.4%)	8/42 (19%)	4/42 (9.5%)
4-5 year-olds: 40/47 (85.1%)	3/47 (6.4%)	4/47 (8.5%)

(18) Spanish, present perfect

	CLITIC	CLITIC OMISSION	FULL DP
2 year-olds:	26/32 (81.25%)	5/32 (15.62%)	1/32 (3.12%)
3-year-olds:	39/40 (97.5%)	0	1/40 (2.5%)
4-year-olds:	40/40 (100%)	0	0

Clearly, there is no contrast between present and present perfect, neither in Catalan nor in Spanish. And Catalan patterns with Italian with respect to clitic omission (Schaeffer 2000): the rate of object clitic omission is very high in both languages at the early stages, in contrast to Spanish. Omission remits sharply at the age of 3 both in Catalan and Italian, to disappear by the age of 4, a strong age effect, absent in Spanish.

4.3 Participle agreement

With regard to the present perfect task, we consider whether children speaking Catalan prefer to produce agreement between the participle and the direct object clitic, or whether they prefer to produce the default masculine singular form for the participle. We find instances of participle agreement, but overall children prefer the construction without agreement: the number of sentences without agreement is higher than the number of sentences with agreement (P < 0.001). The actual percentage of participle agreement found in the children's productions is relatively low – as it is for many adult speakers of this variety. The table below shows the frequencies:

(19) Participle agreement

NO OVERT AGREEMENT	AGREEMENT W/ PARTICIPLE
1-2 year-olds: 16/21 (76.1%)	5/21 (23.8%)
3-year-olds: 25/28 (89.3%)	3/28 (10.7%)
4-5 year-olds: 27/38 (71.1%)	11/38 (28.9%)
total 68/87 (78.2%)	19/87 (21.8%)

There are altogether six children who produce some (or all) participles displaying agreement. In the following table the rate of clitic omission for present and present perfect are collapsed:

(20) Participle agreement and rate of clitic omission

```
CLITIC CLITIC OMISSION DP

2-year-olds w/ agr 6/16 (37.5%) 10/16 (62.5%)
w/o agr 5/40 (12.5%) 34/40 (85%) 1/40 (2.5%)

3-year-olds w/ agr 5/7 (71.4%) 2/7 (28.5%)
w/o agr 55/79 (69.6%) 17/79 (21.5%) 7/79 (8.9%)

4-year-olds w/ agr 23/24 (95.8%) 1/24 (4.2%)
w/o agr 62/66 (93.9%) 4/66 (6%)
```

The results in (20) indicate that, regardless of whether they produce agreeing participles or not, all children seem to behave in the same way with respect to clitic omission: there is no statistically significant difference between the clitic omission rate in children who produce some (or consistent) participle agreement and those who do not. This has implications for the characterisation of the optionality of participle agreement in Catalan: under one interpretation, even though some children's productions display no overt agreement, the structure generated may still involve an AgroP projection with an uninterpretable feature to be checked, as assumed above for participle agreement languages.

4.4 Clitic forms

Finally, let us consider the clitic form produced by the children. In the present tense task, at all ages children produce a percentage of target clitic forms in both Catalan and Spanish above-chance level (P < 0.001). Spanish-speaking children produce virtually no non-target forms. We have found a significant difference in the number of non-target clitics between Catalan-speaking children and Spanish-speaking children (P < 0.05), although this difference occurs only for 3 year olds (P < 0.05).

³ In Catalan, in the present tense task, the errors attested involved in two out of three cases the masculine instead of the feminine. We come back to this fact in the next section and the discussion.

(21) Catalan clitic form, present tense

	TARGET	[LI] [+ANIMATE]	NON-TARGET
2-year-olds:	4	3	0
3-year-olds:	17	10	3*
4-year-olds:	30	8	0
5-year-olds:	6	1	0
total	57/82 (69%)	22/82 (26.8%)	3/82 (3.6%)

(21) indicates that a new pronominal system may be emerging in Catalan, in which animacy is marked in the pronominal system, rather than the opposition between accusative/dative (*li* is otherwise a dative clitic in standard Catalan). The children in our experiment who produced *li* as an accusative clitic produced it systematically for animate objects, in no case for inanimates. What may be a new system, illustrated in (22), is not unique to children, and can be found in adult varieties, specially by Spanish native speakers; there is no study available of this phenomenon in adult language, but sentences such as (22) have not been included in the non-target class.

(22) Li pentina.

Cl combs

"He combs him/her (animate)."

(23) Spanish clitic form, present tense

	TARGET [LA]	TARGET [LE]	TARGET [LO]	NON-TARGET
2-year-olds:	14/32	13/32	5/32	0
3-year-olds:	17/39	21/39	2/39	0
4-year-olds:	18/40	21/40	1/40	0
total	111/111 (100%)	0	0	0

Depending on the variety the children had been exposed to (*leista* or not), they produced *le* or *lo* as clitics; they can both be regarded as target.

Also in the present perfect tense, collapsing all ages, children produce target clitic forms at above-chance level in Catalan (P < 0.05) and Spanish (P < 0.001). The percentage of errors found in Catalan appears in (24):

(24) Catalan clitic form, present perfect

	TARGET (LES)	NON-TARGET [L]
1-2 year-olds:	0	3/3 (100%)
3-year-olds:	6/25 (24%)	19/25 (76%)
4-5 year-olds:	31/35 (88.6%)	4/35 (11.4%)

Interestingly, all the errors found in Catalan are of the types illustrated in (25): [l] for *les*, that is, the masculine, singular (unmarked) form instead of the feminine, plural form in (26).⁴

(25) a. L' ha menjat.

AC.CL has eaten (target Les ha menjat/des)

b. *L'* ha menjades. (found marginally)

AC.CL has eaten.FEM.PL

(26) Les ha menjades .(target)

AC.CL-FEM-PL has eaten.FEM.PL

"S/he has eaten them (FEM)."

In Spanish an apparently similar phenomenon is taking place: target *las* co-occurs with *la* or [1], with no feminine marker:

(27) Spanish clitic form, present perfect

TARGET LAS	LA	[L]	OTHER
2yolds: 7/19 (36.85%)	0	11/19 (57.9%)	1/19 (5.2%)
3yolds: 8/32 (25%)	8/32(25%)	14/32 (43.75%)	2/32 (6.25%)
4yolds: 24/33 (72.7%)	7/33(21.2%)	2/33 (6%)	0

In the case of Spanish, it is possible to argue that these early reduced forms result from a phonological process of coda deletion (la for las) or simplification of the syllabic structure ([l] [a] for [la] [a] la ha); these phenomena are well attested in early Spanish, although there is considerable individual variation in their occurrence (Conxita Lleó, p.c.). Statistically, there is a significant difference between target and non-target clitic forms in Spanish-speaking children. We do not find a significant difference in the number of non-target clitics between Catalan-speaking children and Spanish

-

⁴ In the present tense task, no errors were produced in Spanish, but as mentioned three errors were attested in Catalan, of which two correspond to a masculine instead of a feminine form. As the target forms in the present tense task weren't always the same, no further analysis will be pursued, but the general pattern coincides with that in the present perfect.

speaking children. However, the source of the non-target clitics is not necessarily the same in Catalan and Spanish: although it can be argued that the non-target forms stem from a phonological process in Spanish, similar processes do not seem to be so readily available to Catalan-speaking children of the same age (rather, Catalan-speaking children appear to develop a more complex syllable structure at an earlier age). That leaves the possibility that the clitic forms exemplified in (25) are different in nature from those in Spanish.

In her study of clitic omission in child Italian, Schaeffer (2000) also found what she termed 'contracted plural object clitics', which appeared in the proportion indicated in (28) and are exemplified in (29b). We turn to the interpretation of the Catalan and Italian cases in the discussion.

(28) Proportions of target direct object clitics and contracted plural object clitics (Italian)

age	TARGET CLITIC	CONTRACTED PL CLITIC
2-year-olds:	62%	38%
3-year-olds:	77%	23%
4-year-olds:	90%	10%
5-year-olds:	92%	8%

(29) a. Li ha pettinati

CL.MASC.PL has combed.MASC.PL

"(He) has combed them."

- b. L' ha pettinati.
 - CL has combed.MASC.PL

5. Discussion

The results of our experiment allow us to corroborate for Catalan and Spanish one of the findings of previous studies regarding the development of object clitics in child grammar: object clitics appear in the right position with respect to the verb. Preverbal placement with finite verbs and post-verbal placement with non-finite verbs occurs with virtually no errors, and that indicates that the child is sensitive to finiteness features from the earliest record, and that raising of finite verbs to T occurs systematically. We do not expect children to have any problems identifying the morphosyntactic features of functional categories, and that is indeed what we find.

Second, Catalan object clitics are omitted in structures in which they are obligatory, as was found by Schaeffer (2000) for Italian (30), and for the same period: roughly the same stage in which Optional Infinitive effects are found in

non-null-subject languages. Up until the age of three, Catalan speaking children resort to omission, rather than clitic production, and omission does not disappear entirely until the age of four; this can be compared with the results for child Italian:

(30) Italian object clitic omission (Schaeffer 2000)

	CLITIC	CLITIC OMISSION	FULL DP
2-year-olds:	22%	64%	14%
3-year-olds	62%	15%	23%
4-year-olds:	89%	0%	11%
5-year-olds	91%	0%	9%

The results for Catalan sharply contrast with those for Spanish, since Spanish-speaking children produce obligatory object clitics from the first age group studied. The contrast attested between Catalan and Spanish is as predicted by the UCC (together with Minimise Violations), given the difference between Catalan and Spanish: while the first is a participle agreement language, the second is not. The optionality of object clitics in Catalan (and Italian) is as predicted by the UCC for a language displaying participle agreement, i.e. with checking of more than one uninterpretable feature by the object DP. On the other hand, the very low clitic omission rate found in Spanish is as predicted by the UCC if we assume that Spanish objects check against only one uninterpretable feature.

In the third place, the optional character of participle agreement in the tested variety of Catalan is of no consequence. We have to argue, then, that although participle agreement does not always occur, in the target grammar the *pro* object must check in all circumstances against two uninterpretable D-features, one in Agr_OP, one in CIP. The same holds for object clitics in the present tense: while participle agreement is only visible in the perfect tenses, the rate of clitic omission is the same in the present and the present perfect: so double checking must occur in a parallel fashion in both cases (this lack of contrast between different tenses is found not only in Catalan, but also in Italian; see Schaeffer 2000).

Finally, let us consider the form of the clitic in the children's productions. As pointed out in section 4.4., the clitics produced for Catalan were not always the target feminine, plural in the present perfect task; rather, they were systematically the unmarked clitic form (corresponding to the masculine, singular); the data available for Italian follow the same pattern. In Wexler (to appear) the suggestion is made that just as object clitic omission may be the

result of a convergent derivation under the UCC, clitics with a default case may also result from it. If the default case in Romance is the accusative⁵, then the unmarked clitics found in child Catalan (and Italian) may correspond precisely to such default forms. So under the UCC, we can explain these deviant forms found in Catalan and Italian by the inability of the child to double D-check: the default clitic [I] occurs when CIP is projected and Agr_O is not projected. This default clitic has as verbal counterpart the root infinitive. Thus we extend previous work on clitic development and argue that omission may not be the only outcome of the interaction of the properties of participle-agreement languages and the UCC.

References

Cortés, Corinne. 1992. *Issues in Catalan Syntax*. PhD Dissertation, University of California, Los Angeles.

Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, Mass.: MIT Press.

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

Jakubowicz, Celia, Natascha Müller, Ok-Kyung Kang, Beate Biemer and Catherine Rigaut. 1996. "On the Acquisition of the Pronominal System in French and German" *BUCLD* 20, ed. by Andy Stringfellow et al., 374-385.

Kayne, Richard. 1989. "Facets of Romance Past Participle Agreement". Dialect Variation and the Theory of Grammar ed. by Paola Benincà, 85-103. Dordrecht: Foris Publications.

Lizsckowski, David 1999. *On the Acquisition of Pronominal Object Clitics*. PhD Dissertation, Harvard University.

Pierce, Amy. 1992. Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars.

Dordrecht: Kluwer.

⁵ The assumption that accusative is the default case in Romance may be questionable, as pointed out by an anonymous reviewer. However, to our knowledge the only evidence that can be adduced to back up nominative as the default case is provided by hanging topic constructions, where a non-agreeing topic is nominative:

⁽i) Jo, m' agrada anar a la platja.1sNom 1sAcc appreciate go-inf to the beach.

[&]quot;Me, I like going to the beach."

- Quartz, Steven, & Terrence Sejnowski. 1997. "The Neural Basis of Cognitive Development: A Constructivist Manifesto". *Behavioral and Brain Sciences*, 20. 537-596.
- Schaeffer, Jeannette. 2000. *The Acquisition of Direct Object Scrambling and Clitic Placement*. Amsterdam/Philadelphia: John Benjamins.
- Sportiche, Dominique.1996. "Clitic Constructions". *Phrase Structure and the Lexicon* ed. by Johan Rooryck & Laurie Zaring. Dordrecht: Kluwer.
- Uriagereka, Juan. 1995. "Aspects of the Syntax of Clitic Placement in Western Romance". *Linguistic Inquiry*, 26. 79-123.
- Wexler, Ken. 1998. "Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation of the Optional Infinitive Stage". *Lingua*, 106. 23-79.
- Wexler, Ken. To appear. "The Unique Checking Constraint as the Explanation of Clitic Omission in SLI and Normal Development". *Essays on Syntax, Morphology and Phonology in SLI* ed. by Celia Jakubowicz, Lea Nash and Ken Wexler. Cambridge, Mass.: MIT Press.

INDEX OF LANGUAGES AND DIALECTS

This index lists languages and dialects that are mentioned in the texts and languages and dialects that are illustrated in the examples.

Passaua. 99 02 05	Carmon, 94, 99, 92
Basque: 88, 93-95	German: 84, 88-92
Catalan: 84-88, 184, 194-196, 253-	Italian: 49, 55, 62-63, 69, 85-95,
268	99-118, 119-123, 125-127,
Balearic: 217-234	128-131, 133-135, 137-138,
Central: 228	142, 156-159, 183, 184,
Eastern: 220	194-197, 257, 259, 261,
Eivissan: 218-220, 226-230	265-266
Majorcan: 217-227	Irish: 55
Minorcan: 217-227, 230	Japanese: 64
Dutch: 87-95, 131-132, 156-157,	Latin: 197
166-171, 173, 178, 204-205,	Portuguese: 183, 184, 190-193,
215	194-196
English: 18, 20, 37-38, 48-49, 51,	Brazilian: 1-16
54-65, 69, 76, 84-95, 107,	European: 1-16, 119-120,
124, 152, 157-159, 178,	125-131, 137
185, 187, 190, 199-216	Roumanian: 119-140, 142, 156-160
American: 243	Russian: 142
Belfast: 51, 53	Spanish: 38, 48, 54, 56-62, 64-65,
Black: 51	69-81, 84-91, 92-95, 119-
Finnish: 156-157	120, 122-123, 126, 128,
French: 18-29, 33-46, 55, 57, 84-	130, 133-134, 136-138,
95,142, 143, 156-158, 171-	183-198, 199-216, 253-278
179, 183, 184, 189, 194-	Dominican: 211-215
197, 201, 203, 237-238,	Swedish: 55, 156-157
240-241, 244, 245-247	Valdôtain: 237
Patois: 237	Welsh: 55
Quebec: 237-253	-
~	

SUBJECT INDEX

This index does not claim to be exhaustive. Specific language items are printed in italic.

iii italic.	
	Clitic-resumption: 99-118
A	Direct object clitic: 253-268
A'-dependencies: 103	Cliticisation: 35
A'-movement: 102, 113, 116	Competition: 183-184, 189, 194
Accusative Case: 157	Complete Functional Complex:
Adjacency: 183	104-106
Adverb: 1-16	Compositionality: 142-144, 146,
Frequency adverb: 1-16	151
Affirmative operator: 236, 242-244,	Correlation of events: 4-5, 10-11
249	
Affrication 225	D
Anaphor: 141, 143-144, 146, 149,	D-linking: 145, 169-175
151, 169-175	De: 192-193
Deep: 183, 185-186	Deep anaphor: 183, 185-186
Article: 188-189, 191-197	Deletion: 218-221, 223-224, 226-
Assertive Phrase: 6-9, 12, 13	232
Aux-insertion 91-93	Dependency: 141, 143
	Referential: 141
В	Determiner: 191-197
Bare plural: 152	Definite: 199, 200-202, 204,
Binding: 38-39, 103-110	208, 213
	Possessive: 200, 203, 211
C	Dislocation: 17-30, 35-39, 43-44,
Case: 156	99-118
Accusative: 157	Left-dislocation: 17-27, 99-
Genitive: 156, 191	118
Partitive: 156-157	Right-dislocation: 25-27,
Clitic: 33-45, 99-118, 190-191,	24-99-118
193-194, 253-268	Dissimilation: 217-221, 223-224,
Clitic-doubling: 103	227, 230
Clitic omissions 253-268	Distributed Morphology: 183-198

Distributive: Distributive-Universal Phrase: 6-8, 10, 11	Genitive Case: 156, 191 Gliding: 217-219, 230
Distributive interpretations: 205	H Head Movement: 47 - 67
E Ellipsis: 183-198 Nominal: 183-198	Internally-driven: 56-60, 62-65
Else: 141-162 Additive: 142, 148, 155 Exclusive: 142-162	I Inalienable possession: 199-216 Indefinite: 120, 122-124, 135-138,
Empty category: 183 EPP feature: 47, 49, 52-53, 60 Event argument: 3	145, 148, 156-157, 159 Individual Level Predicates: 40 Infinitives 83-95
Eventive expressive structures: 236, 244 Exclusion: 147-148, 150, 152-153 Existential reading: 36	Inflection 169-175 Inversion 48, 51-53, 65 IP-inversion: 100-101, 107- 108, 110, 116
Extension Condition: 50, 58, 66	K
F Feature:	Kind: 150-155, 201-203, 206, 214
Assertive: 7-8 Distributive-universal: 5-8, 10 Feature checking: 253-268 Interpretable: 199	L Late insertion: 183 Lexical Parsimony:58 Lexical restrictions: 206 Locality: 111-113
Negative: 120, 122, 124- 125, 130-136, 138 Quantificational: 120-121, 124-126, 130-132, 138 Focus: 33, 37-39-42, 44, 69-81,	M Merge: 99-118 Minimality: 113-116 Morphological identity: 183, 187
119-140 Contrastive: 72-73, 77 Identificational: 126	Morphological structure: 183, 187
Fusion: 218-219, 221, 225-232 G	Negative concord: 119-140 Negative Polarity Item: 120-124, 130-133, 135-136, 144
Gender: 186-188 Generic reading: 36	N-movement: 48, 60-64 Nominal ellipsis: 183-198

Nuclear Stress Rule: 101 Null subject 27-29 Number: 186-187 N-word: 119-140	Proper Binding Condition: 112 Proper name: 154 Prosody: 69-70, 79
O Object: 151-155 One-substitution: 185, 190 Optimality Theory: 217, 227-228 P Parameter: 199, 201-202, 206, 215 Parametric difference: 199, 201-204 Parametric variation: 48,	Q Quantification: 1-16
64-65 Parametrization: 199 Parametrized: 254fn Parasitic Gaps: 101-102 Participle agreement: 253, 255-256, 258, 261-262, 266	R Reconstruction: 104-108 Relation R: 152 Relative complementation: 147, 151
Partitive Case: 156-157 Partitive construction: 155-161 Partitivity: 141-162 Periphrastic 91-93 Pf-deletion: 183, 186-188 P-movement: 72 Phase: 191-193 Phonological Phrase: 70, 75-76 Phonology: 69-81 Pitch accent: 69-70, 76-79 Positional Faithfulness 225, 228-229, 232 Possessive: Pronoun: 203 Determiner: 200, 203, 211 Predicative nominal: 154-155 Pronoun: 194-197 Possessive pronoun: 203	S Scope: 108-110, 143-144, 146 Semantic competition: 199, 202 Sempre: 1-16 Confirmative/assertive reading: 1-2, 5, 7-9, 12-13 Position: 2-5, 8-12 Temporal reading: 1-5, 7, 10-11 Tense: 2-5, 8, 10, 12 Sibilants: 217-234 Spell-Out: 186-187 Stress: 69-73, 75-79 Subjacency: 113 Subject: 33-45 Superiority effects: 114 (Super)positive marker: 235-251 Syntax: 69-70

T	Type-shifting: 202
Tense: 2-5, 8, 10, 12	-
Past tense: 2-5, 8, 10	${f U}$
Present tense: 2-5, 8, 10, 12	Unique Checking Constraint
Tense Object Phrase: 6-8, 10	(UCC): 253-256, 266-267
Type-shifting: 202	Unranked hierarchy: 232
Token: 149-152	
Topic: 17-30, 33-45, 99-118, 127-	${f V}$
129, 136, 139, 171	Variable: 110-111, 144-146, 153,
Local topics: 113-116	156, 158-160
Multiple topics: 102, 114-	Pronominal: 110-111
116	V-raising: 54-61, 65
Non-local topics: 113-116	
Topic Phrase: 99-100	
Topicalization: 99-118	\mathbf{W}
Ti: 236-238, 245-246, 249	Weak Cross Over: 101-102
<i>Tu</i> : 235-251	Wh-extraction: 12
Type: 149-153, 155	Wh-word: 143-144, 148, 159-160

CURRENT ISSUES IN LINGUISTIC THEORY

E. F. K. Koerner

Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Berlin

The Current Issues in Linguistic Theory (CILT) series is a theory-oriented series which welcomes contributions from scholars who have significant proposals to make towards the advancement of our understanding of language, its structure, functioning and development. CILT has been established in order to provide a forum for the presentation and discussion of linguistic opinions of scholars who do not necessarily accept the prevailing mode of thought in linguistic science. It offers an alternative outlet for meaningful contributions to the current linguistic debate, and furnishes the diversity of opinion which a healthy discipline must have. A complete list of titles in this series can be found on the publishers website, www.beniamins.com

- 222 HERSCHENSOHN, Julia, Enrique MALLÉN and Karen ZAGONA (eds.): Features and Interfaces in Romance. Essays in honor of Heles Contreras. 2001. xiv, 302 pp.
- 223 FANEGO, Teresa, Javier PÉREZ-GUERRA and María José LÓPEZ-COUZO (eds.): English Historical Syntax and Morphology. Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000. Volume 1. 2002. x, 306 pp.
- 224 FANEGO, Teresa, Belén MÉNDEZ-NAYA and Elena SEOANE (eds.): Sounds, Words, Texts and Change. Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000. Volume 2. 2002. X, 310 pp.
- 225 SHAHIN, Kimary N.: Postvelar Harmony. 2003. viii, 344 pp.
- 226 **LEVIN, Saul:** Semitic and Indo-European. Volume II: Comparative morphology, syntax and phonetics. 2002. xviii, 592 pp.
- 227 FAVA, Elisabetta (ed.): Clinical Linguistics. Theory and applications in speech pathology and therapy. 2002. xxiv, 353 pp.
- 228 **NEVIN, Bruce E. (ed.):** The Legacy of Zellig Harris. Language and information into the 21st century. Volume 1: Philosophy of science, syntax and semantics. 2002. xxxvi, 323 pp.
- 229 NEVIN, Bruce E. and Stephen B. JOHNSON (eds.): The Legacy of Zellig Harris. Language and information into the 21st century. Volume 2: Mathematics and computability of language. 2002.
 XX. 312 DD.
- 230 PARKINSON, Dilworth B. and Elabbas BENMAMOUN (eds.): Perspectives on Arabic Linguistics. Papers from the Annual Symposium on Arabic Linguistics. Volume XIII-XIV: Stanford, 1999 and Berkeley, California 2000. 2002. xiv, 250 pp.
- 231 CRAVENS, Thomas D.: Comparative Historical Dialectology. Italo-Romance clues to Ibero-Romance sound change. 2002. xii, 163 pp.
- 232 BEYSSADE, Claire, Reineke BOK-BENNEMA, Frank DRIJKONINGEN and Paola MONACHESI (eds.): Romance Languages and Linguistic Theory 2000. Selected papers from 'Going Romance' 2000, Utrecht, 30 November–2 December. 2002. viii, 354 pp.
- 233 WEIJER, Jeroen van de, Vincent J. van HEUVEN and Harry van der HULST (eds.): The Phonological Spectrum. Volume I: Segmental structure. 2003. x, 308 pp.
- 234 WEIJER, Jeroen van de, Vincent J. van HEUVEN and Harry van der HULST (eds.): The Phonological Spectrum. Volume II: Suprasegmental structure. 2003. x, 264 pp.
- 235 LINN, Andrew R. and Nicola McLELLAND (eds.): Standardization. Studies from the Germanic languages. 2002. xii, 258 pp.
- 236 SIMON-VANDENBERGEN, Anne-Marie, Miriam TAVERNIERS and Louise J. RAVELLI (eds.): Grammatical Metaphor. Views from systemic functional linguistics. 2003. vi, 453 pp.
- 237 BLAKE, Barry J. and Kate BURRIDGE (eds.): Historical Linguistics 2001. Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001. Editorial Assistant: Jo Taylor. 2003. x, 444 pp.
- 238 NÚÑEZ-CEDEÑO, Rafael, Luis LÓPEZ and Richard CAMERON (eds.): A Romance Perspective on Language Knowledge and Use. Selected papers from the 31st Linguistic Symposium on Romance Languages (LSRL), Chicago, 19–22 April 2001. 2003. xvi, 386 pp.
- 239 ANDERSEN, Henning (ed.): Language Contacts in Prehistory. Studies in Stratigraphy. Papers from the Workshop on Linguistic Stratigraphy and Prehistory at the Fifteenth International Conference on Historical Linguistics, Melbourne, 17 August 2001. 2003. viii, 292 pp.

- 240 JANSE, Mark and Sijmen TOL (eds.): Language Death and Language Maintenance. Theoretical, practical and descriptive approaches. With the assistance of Vincent Hendriks. 2003. xviii, 244 pp.
- 241 **LECARME, Jacqueline (ed.):** Research in Afroasiatic Grammar II. Selected papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000. 2003. viii, 550 pp.
- 242 SEUREN, Pieter A.M. and Gerard KEMPEN (eds.): Verb Constructions in German and Dutch. 2003. vi, 316 pp.
- 243 CUYCKENS, Hubert, Thomas BERG, René DIRVEN and Klaus-Uwe PANTHER (eds.): Motivation in Language. Studies in honor of Günter Radden. 2003. xxvi, 403 pp.
- 244 PÉREZ-LEROUX, Ana Teresa and Yves ROBERGE (eds.): Romance Linguistics. Theory and Acquisition. Selected papers from the 32nd Linguistic Symposium on Romance Languages (LSRL), Toronto, April 2002. 2003. viii, 388 pp.
- 245 QUER, Josep, Jan SCHROTEN, Mauro SCORRETTI, Petra SLEEMAN and Els VERHEUGD (eds.): Romance Languages and Linguistic Theory 2001. Selected papers from 'Going Romance', Amsterdam, 6–8 December 2001. 2003. viii, 355 pp.
- 246 HOLISKY, Dee Ann and Kevin TUITE (eds.): Current Trends in Caucasian, East European and Inner Asian Linguistics. Papers in honor of Howard I. Aronson. 2003. xxviii, 426 pp.
- 247 PARKINSON, Dilworth B. and Samira FARWANEH (eds.): Perspectives on Arabic Linguistics XV. Papers from the Fifteenth Annual Symposium on Arabic Linguistics, Salt Lake City 2001. 2003. x, 214 pp.
- 248 WEIGAND, Edda (ed.): Emotion in Dialogic Interaction. Advances in the complex. 2004. xii, 284 pp.
- 249 BOWERN, Claire and Harold KOCH (eds.): Australian Languages. Classification and the comparative method. 2004. xii, 377 pp. (incl. CD-Rom).
- 250 **JENSEN, John T.:** Principles of Generative Phonology. An introduction. 2004. xii, 324 pp.
- 251 KAY, Christian J., Simon HOROBIN and Jeremy J. SMITH (eds.): New Perspectives on English Historical Linguistics. Selected papers from 12 ICEHL, Glasgow, 21–26 August 2002. Volume I: Syntax and Morphology. 2004. x, 264 pp.
- 252 KAY, Christian J., Carole HOUGH and Irené WOTHERSPOON (eds.): New Perspectives on English Historical Linguistics. Selected papers from 12 ICEHL, Glasgow, 21–26 August 2002. Volume II: Lexis and Transmission. 2004. xii, 273 pp.
- 253 CAFFAREL, Alice, J.R. MARTIN and Christian M.I.M. MATTHIESSEN (eds.): Language Typology. A functional perspective. xiii, 690 pp. + index. Expected Fall 2004
- 254 BALDI, Philip and Pietro U. DINI (eds.): Studies in Baltic and Indo-European Linguistics. In honor of William R. Schmalstieg. 2004. xlvi, 302 pp.
- 255 **MEULEN, Alice ter and Werner ABRAHAM (eds.):** The Composition of Meaning. From lexeme to discourse. 2004. vi, 230 pp.
- 256 BOK-BENNEMA, Reineke, Bart HOLLEBRANDSE, Brigitte KAMPERS-MANHE and Petra SLEEMAN (eds.): Romance Languages and Linguistic Theory 2002. Selected papers from 'Going Romance', Groningen, 28-30 November 2002. 2004. viii, 273 pp.
- 257 FORTESCUE, Michael, Eva Skafte JENSEN, Jens Erik MOGENSEN and Lene SCHØSLER (eds.): Historical Linguistics 2003. Selected papers from the 16th International Conference on Historical Linguistics, Copenhagen, 11–15 August 2003. x, 315 pp. + index. Expected Fall 2004
- 258 AUGER, Julie, J. Clancy CLEMENTS and Barbara VANCE (eds.): Contemporary Approaches to Romance Linguistics. Selected Papers from the 33rd Linguistic Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003. With the assistance of Rachel T. Anderson. viii, 404 pp. Expected Fall 2004
- 259 CARR, Philip, Jacques DURAND and Colin J. EWEN (eds.): Headhood, Elements, Specification and Contrastivity. Phonological papers in honor of John Anderson. 430 pp. Expected Winter 04-05
- 260 NICOLOV, Nicolas, Ruslan MITKOV, Galia ANGELOVA and Kalina BONTCHEVA (eds.): Recent Advances in Natural Language Processing III. Selected papers from RANLP 2003. ca. 400 pp. Expected Winter 04-05
- 261 KAY, Christian J. and Jeremy J. SMITH (eds.): Categorization in the History of English. ca. 260 pp. Expected Winter 04-05
- 262 VAJDA, Edward J. (ed.): Languages and Prehistory of Central Siberia. ca. 280 pp. Expected Winter 04-