

Anuario del Seminario de Filología Vasca «Julio de Urquijo», 2005-2

A NEW RELAY OF LINGUISTICS

Asier Alcázar, Irene Barberia,
Rebeca Campos-Astorkiza & Susana Huidobro
(eds.)



Gipuzkoako Foru Aldundia
Diputación Foral de Gipuzkoa
Donostia - San Sebastián

erman ta zabal zazu



Universidad del País Vasco Euskal Herriko
Unibertsitatea
Bilbo-Bilbao

Anuario del Seminario de Filología Vasca “Julio de Urquijo”
International Journal of Basque Linguistics and Philology
(ASJU)

Sortzaileak / Fundadores / Founded by
Manuel Agud (†) - Luis Michelena (†)

Zuzendariak / Directores / Directors
Joseba A. Lakarra (JUMI, UPV/EHU) - Ibon Sarasola (JUMI, UPV/EHU)

Argitaratzailea / Editor
Joseba A. Lakarra (JUMI, UPV/EHU)

Idazkaritza / Redacción / Board

Gidor Bilbao (JUMI, UPV/EHU)	Miren Lourdes Oñederra (JUMI, UPV/EHU)
Iñaki Camino (JUMI, UPV/EHU)	Javier Ormazabal (JUMI, UPV/EHU)
Ricardo Gómez (JUMI, UPV/EHU)	Íñigo Ruiz Arzalluz (JUMI, UPV/EHU)
Joaquín Gorrochategui (JUMI, UPV/EHU)	Blanca Urgell (JUMI, UPV/EHU)
Iván Igartua (JUMI, UPV/EHU)	Koldo Zuazo (JUMI, UPV/EHU)

Idazkaritza arduradunak / Coordinadores / Editor's Adjunts
Íñigo Ruiz Arzalluz (JUMI, UPV/EHU) / Blanca Urgell (JUMI, UPV/EHU)

Aholku Batzordea / Consejo Asesor / Advisory Board

Jacques Allières (Toulouse) (†)	Alazne Landa (UPV/EHU)
Patxi Altuna (Deustu, Donostia) (†)	Jesús María Lasagabaster (Deustu, Donostia)
Jesús Antonio Cid (Complutense)	Francisco Oroz Arizcuren (Tübingen)
Lyle Campbell (Utah)	Jon Ortiz de Urbina (Deustu, Bilbo)
Bernard Comrie (Santa Barbara / Max Planck, Leipzig)	Beñat Oyharçabal (Iker, CNRS)
Joan Coromines (Barcelona) (†)	José Antonio Pascual (Carlos III)
María Teresa Echenique (Valencia)	Georges Rebuschi (Paris III)
Ricardo Etxepare (Iker, CNRS)	Rudolf P. G. de Rijk (Leiden) (†)
Jon Franco (Deustu, Bilbo)	Patxi Salaberri (UPNA)
Jean Haritschelhar (Bordele III)	Iñaki Seguro (Orotariko Euskal Hiztegia)
José Ignacio Hualde (Urbana, Illinois)	Larry Trask (Sussex) (†)
Bernard Hurch (Graz)	Juan Uriagereka (Maryland)
Jon Juaristi (Alcalá)	Myriam Uribe-Etxebarria (UPV/EHU)
Itziar Laka (UPV/EHU)	Charles Videgain (Iker / UPPA)

ASJU 1954an sortutako Euskal Linguistika eta Filologiazko nazioarteko aldizkaria da, “Julio Urkixo” Euskal Filologi Mintegi-Institutuak (JUMI) argitaratua, eta iker-eremu horietatik edo horietarako ere interesgarri izan daitezkeenetik goi mailako artikuluko, ohar eta liburu-iruzkinak argitaratzen ditu. Urtero bi zenbaki ateratzen dira (orotara 750 bat orrialde). ASJU-k badu, orobat, *GEHIGARRI* sail bat, non artikuluko formatuaz gorako lanak argitaratzen diren, epe jakinik gabe (ikus zerrenda zenbakiaren amaieran).

Originalen buruzko harremanetarako ikus bitez bukatuak *EGILEENTZAKO OHARRAK*. Harpidetza eta eskarietarako: Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua, Leioa (luxedito@lg.ehu.es).

ASJU es una revista internacional de Lingüística y Filología Vasca fundada en 1954 y publicada por el Instituto-Seminario de Filología Vasca “Julio Urquijo” (JUMI). Se publican en ella artículos, notas y reseñas sobre los campos mencionados y afines, y otros de interés para los mismos. En la actualidad es de periodicidad semestral (unas 750 páginas). Sin regularidad preestablecida, ASJU publica en sus ANEJOS trabajos de formato superior al de un artículo (véase la lista al final del volumen).

Para correspondencia relacionada con los originales véase la *INFORMACIÓN PARA LOS AUTORES* al final del número. Para suscripciones y pedidos: Servicio de Publicaciones, Universidad del País Vasco (UPV/EHU), Leioa (luxedito@lg.ehu.es).

ASJU is an International Journal of Basque Linguistics and Philology founded in 1954, and published by the “Julio de Urquijo Seminar of Basque Philology” Institute (JUMI). It publishes high-quality papers, notes, squibs and reviews about the above mentioned and other related topics in two issues per year (up to a total of 750 pages). Longer works are published as SUPPLEMENTS to the regular issues of the ASJU (see the list at the end of the issue).

For correspondence about papers see the *INFORMATION FOR AUTHORS* at the end of this issue. For subscriptions and order: Servicio de Publicaciones, Universidad del País Vasco (UPV/EHU), Leioa (luxedito@lg.ehu.es).

Asier Alcázar, Irene Barberia, Rebeca Campos-Astorkiza & Susana Huidobro
(eds.)

A NEW RELAY OF LINGUISTS

Anuario del Seminario de Filología Vasca «Julio de Urquijo» XXXIX: 2 (2005)

A NEW RELAY OF LINGUISTS

Asier Alcázar, Irene Barberia, Rebeca Campos-Astorkiza & Susana Huidobro
(eds.)



Gipuzkoako Foru Aldundia
Diputación Foral de Gipuzkoa

eman ta zabal zazu



Universidad del País Vasco Euskal Herriko
Unibertsitatea

- © Asier Alcázar, Irene Barberia, Rebeca Campos-Astorkiza & Susana Huidobro
- © The authors
- © «Julio Urkixo» Basque Philology Seminar
- © «Julio Urkixo» Euskal Filologia Mintegia
- © Servicio Editorial de la Universidad del País Vasco
- © Euskal Herriko Unibertsitateko Argitalpen Zerbitzua

ISSN: 0582-6152

Depósito legal / Lege gordailua: BI -

Fotocomposición / Fotokonposizioa: Ipar, S. Coop.
Particular de Zurbaran, 2-4 (48007 Bilbao)

Impresión / Inprimatzea: Itxaropena, S.A.
Araba kalea, 45 (20800 Zarautz-Gipuzkoa)

INDEX

Preface	1
Introduction: «A new relay of linguists»	3
ANCHIMBE, Eric A. and ANCHIMBE, Stella A.: «Sociolinguistic variables in the <i>degeneracy</i> of English in postcolonial (<i>non native</i>) contexts»	13
ANCHIMBE, Eric A.: «Multilingual backgrounds and the identity issue in Cameroon»	33
ETXEBERRIA, Urtzi: «Quantification and compositional strategies»	49
FERNÁNDEZ-SALGUEIRO, Gerardo and MARLO, Michael R.: « <i>He himself</i> : rediscovering a non-local anaphor»	69
GALLEGO, Ángel J.: «Minimalist edge coordinations»	79
HADDICAN, Bill: «On verb focalization in central and western Basque»	105
MARUŠIČ, Franc: «What happens when phases get individualistic: On LF-only & PF-only phases»	121
MAYORAL HERNÁNDEZ, Roberto: «A typological approach to the ordering of adverbials: weight, argumenthood and EPP»	141
MONDOÑEDO C., Aysa: «Nominal participles, a case of categorial alternance: Eventive nominalizations in <i>-da</i> »	161
MURGUIA, Elixabete: «Antecedent-Gap relations and locality in verbal ellipsis»	175
RAVIV, Sofie: «Identifying and processing topicalization in Danish»	195
ROY, Isabelle: «Predicate nominals in eventive copular sentences»	213
SALZMANN, Martin: «Resumptive pronouns and Matching effects in Zurich German Relative clauses as distributed deletion»	237
WONG, May Lai-Yin: «The compilation of a sample PFR Chinese corpus of Skeleton-parsed sentences»	271

PREFACE

Many of our colleagues have contributed to make the first BIDE possible, starting from the early stages leading to the conference, all the way to the publication of the proceedings in this special volume of *ASJU*. We would like to take this opportunity to acknowledge them all.

Jon Franco and Jon Ortiz de Urbina from the University of Deusto deserve a special mention since they were involved in this project from its inception until the editing of the proceedings. Without their help the organization of the conference and the publication of this volume could have never been achieved. Jon Franco and Jon Ortiz de Urbina provided logistics and technical solutions as well as academic and stylistic recommendations. Their contribution goes beyond the encouragement and help developing BIDE, for they are indeed the driving force of the present generations of students in Linguistics at the University of Deusto. Their professional skills and amicable contact with the students have certainly encouraged the latter to pursue careers in linguistics. They have not only mentored students in their professional education, but they have also been a moral support in the difficult moments. Their classes have been suggestive and motivating, and they have always been close, accessible and helpful to their students.

We are also grateful to the scientific committee for their help in the review process. These are Alvaro Cerrón-Palomino (University of Southern California), Ricardo Etxepare (CNRS/LEHIA), Alexander Fraser (University of Southern California), Carolina González (University of California, Los Angeles), Leyre Goitia (University of Deusto/University of Nebrija), Nathan Klinedinst (University of California, Los Angeles), Elixabete Murguia (University of Deusto-Bilbao), Iván Ortega (University of Maryland), Leticia Pablos (University of Maryland), Lara Reglero (University of Connecticut), Ana Sánchez Muñoz (University of Southern California) and Carmen Silva-Corvalán (University of Southern California).

For financial and logistic support in hosting the conference, we would like to thank our sponsors. These include the School of Philosophy and Letters and the CIDE Program at the University of Deusto, the Basque Government, Bilbao Bizkaia Kutxa, *Bilbao Iniciativa Turística* and the Getxo Tourist Information Office.

Finally, we wish to express a special gratitude to Joseba Lakarra from *ASJU* Press for his guidance in the task of editing and for making possible the publication of this volume.

*Asier Alcázar
Irene Barberia
Rebeca Campos-Astorkiza
Susana Huidobro*

INTRODUCTION: A NEW RELAY OF LINGUISTS

Asier Alcázar

University of Southern California

Irene Barberia

Universidad de Deusto

Rebeca Campos-Astorkiza

University of Southern California

Susana Huidobro

University of the Basque Country/Stony Brook University

The papers included in this volume were presented at the first *Bilbao-Deusto Student Conference in Linguistics (BIDE'04)*, held at the University of Deusto on July 8-10, 2004. BIDE'04 is set within the broader context of other efforts in Europe like the Student Organization of Linguistics in Europe (ConSOLE) and in the Basque Country like the Workshop on Syntax and Semantics (WoSS), that offer students a forum to share their work in progress. BIDE'04, like the above-mentioned student conferences, aims to be an annual event and serve as a rendezvous, in a permanent location (Bilbao), for new generations of researchers from all areas of linguistics. BIDE provides students a welcoming atmosphere where they can present their own work and get feedback from peers as well as senior researchers. The goal of this collection is to make available to the linguistic community new, promising studies, which contribute to a deeper understanding of the nature of linguistic processes. The diverse linguistic phenomena covered in the articles are analyzed from different perspectives presenting evidence from a variety of languages. The common denominator of all these studies is their contribution to our understanding of the language faculty.

BIDE is the result of the conjoined effort of a group of former Deusto students now working in different international universities (Stony Brook University, University of the Basque Country, University of Deusto-Bilbao, University of Southern California), under the direction of two Deusto professors, Jon Franco and Jon Ortiz de Urbina. These professors are part of the long tradition of students from the Basque Country pursuing their graduate degree in overseas universities. The university of Deusto-Bilbao plays an important role in encouraging Basque students to develop their professional careers in the field of Linguistics. Deusto students were first introduced to the field of Generative Linguistics in the 70's by professor Peter Lavery and professor Manuel Brea and many of today's outstanding linguists took their first steps at the University of Deusto, including some of the most well known linguists such as Karlos Arregi, Gorka Elordieta, Jon Franco, Alazne Landa, Amaya Mendikoetxea, Elixabete Murguía, Jon Ortiz de Urbina, Lara Reglero, Itziar San

Martín and Juan Uriagereka, just to mention some. But not only students from Deusto have benefited from the instrumental role that this university plays; students from the University of the Basque Country were also encouraged and helped by Deusto faculty and students to pursue their graduate studies overseas. Nowadays, professors Franco and Ortiz de Urbina are taking care of the task that professors Lavery and Brea began. In other words, Deusto has been and continues to be an essential means to the development of the linguistic community not only in the Basque Country but also abroad, since some of its students have established and developed successful careers in overseas universities.

Former Deusto students share a spirit of engagement in the linguistic community. They take part in the organization of some of the most prestigious conferences, such as NELS, GLOW, ConSOLE, WECOL, among others. Despite being abroad, their ties with the home research centers are never broken. They continually return for talks, courses and even positions in the different institutions. The organizers of BIDE belong to this group of students who maintain a bond with the Basque linguistic community and their experience abroad is an enriching factor to the research carried out in such community. Being in direct contact with linguists from very different places, they bring a fresh and open-minded approach to the linguistic research. Thus, the idea of organizing a conference in Bilbao emerged on the one hand, as a tribute to the University of Deusto and on the other hand, as a way for other international linguists to become familiar with this research community. The organizers also oriented the conference towards students from a wide spectrum of international universities. This resulted in both Basque and non-Basque linguistics students taking part together and sharing their research.

One of the highlights of the conference was the presentation by the plenary speaker, Juan Uriagereka, a professor at the University of Maryland and the University of the Basque Country, and former student of the University of Deusto. Professor Uriagereka embodies the Deusto spirit. He specializes on syntax and semantics of Basque, Spanish, and other Romance Languages. He has directed a dozen of Ph.D. theses in Europe and the United States, and has been awarded with the Social Science and Humanities Research Prize of the Basque Country (2001). His contribution to the field is extensive. Uriagereka is the author of *Rhyme & Reason* (MIT Press, 1998), for which he received the 1998 best professional book in Language and Linguistics award from the American Association of Publishers, and *Derivations* (Routledge, 2003). He also coauthored *A Course in Minimalist Syntax* (Blackwell, 2005) and *A Course in GB Syntax* (MIT Press, 1988). Furthermore, he published a large number of articles in edited volumes and in prominent linguistic journals such as *Theoretical Linguistics*, *Linguistic Inquiry*, *Natural Language and Linguistic Theory*, and *Syntax*.

Bilbao and, more precisely, the Basque Country is a reference place for ongoing research in different aspects of human language. It is an important locus of linguistic research not only because of its privileged situation as a bilingual community, but also because of the growing number of researchers. Its two main institutions for linguistic research are the University of the Basque Country and the University of Deusto. The members of these centers enjoy the advantage of frequent and fruitful collaborations. It is common to encounter projects developed by researchers from both universities. A prime example is LEHIA (Lengoaiarako Euskal Herriko Ikerguna Aniztuna) or

Basque Center for Language Research. LEHIA was born at the end of the 90's as a result of the informal linguistic meetings that recent Ph.D. graduates were regularly holding at the University of Deusto. Nowadays, the situation of the field of linguistics in the Basque Country is characterized by a wide range of approaches and trends. There are several groups of consolidated researchers working in different disciplines, among them, syntax, semantics, phonology, psycholinguistics, neurolinguistics and second language acquisition, in all of which former Deusto students take part.

Papers in this volume

A significant number of abstracts were submitted for consideration, and 40% of these papers were accepted for presentation at the conference after a peer review process that was carried out by professors and advanced graduate students in the papers' appropriate subfields. This volume includes some selected papers, which cover a variety of topics in linguistics: syntax, semantics, computational linguistics and sociolinguistics.

Several papers in this volume present data from Basque in order to refine current theories with respect to different phenomena such as *do*-support and quantifiers. These papers show the possibility of looking at languages that have not been so referred to, as the case of Basque, to find new evidence to support or modify different accounts. This trend has proven useful in previous work by different linguists, who have used Basque as an ideal testing ground to verify different hypotheses about language universals.

In «Quantification and compositional strategies», Urtzi Etxeberria brings data from the Basque language that makes a contribution to Generalized Quantifier Theory. As in other well-surveyed languages, quantifiers in Basque also divide into two groups: weak and strong. The richness of the empirical data lies in the fact that strong quantifiers in Basque must appear with an article, but weak quantifiers cannot. The dual nature of Etxeberria's data support and reconcile two heretofore contradicting approaches: Matthewson's analysis (2001), which requires the presence of an article; and the standard analysis of generalized quantifiers (Barwise & Cooper 1981, Heim & Kratzer 1998), which does not consider it. In this new light, the Basque data vindicates the need for two different structures. One in which the quantifier merges with an individual of type $\langle e \rangle$ *a la* Matthewson, correlating with a strong, presuppositional interpretation, and another where the quantifier merges with a type $\langle e, t \rangle$ that produces indefinite readings.

In his study of the use of light verb *do* in Basque and Korean «On *egin*: *do*-support and verb focalization in central Basque dialects», William Haddican provides cross-linguistic breadth to the analysis of a familiar topic. *Do*-support is often considered a last-resort strategy to repair a sentence when the verb cannot rise further than structurally required. In this line, Beninca and Poletto (2004) propose that *do*-support has its underpinnings in theta-marking. However, Haddican argues that in Basque and Korean this argument could not apply, for in these languages the verb seems to be a nominal element. The author finds that this condition should be sufficient to motivate the repair strategy. One of the structures Haddican leans on is a type of light verb construction in Basque that typically instantiates unergative verbs

such as «*work*» or «*dance*» (Perlmutter 1978; also known as intransitive, Burzio 1986). Focalization of these verbs by means of *do*-support is not possible. This suggests that the light verb *do* in such constructions and *do*-support target the same position. If Basque unergatives are indeed a mirror image of *do*-support, then their nominal part serves as a basis to identify focalized verbs as nominal elements too.

The volume also includes a number of papers that adopt a comparative approach and present data from different languages with respect to processes such as eventive copular sentences and the status of the PF and LF interfaces. The aim of these papers is to provide a unified explanation to the different features or characteristics that the same phenomena might present across languages. The goal of the generative theory of language is to be able to accommodate for cross-linguistic differences without losing the universality of these patterns.

Isabelle Roy presents a comparative study of «Predicate nominals in eventive copular sentences» in French, German, Irish and Russian. These languages grammaticalize the contrast between eventive and non-eventive sentences through case-marking in Russian, the choice of copula in Irish and the presence or absence of an indefinite article in French and German. Based on cross-linguistic similarities, Roy argues for a unified account of eventive copular sentences. The properties shared by eventive predicate nominals in all of the above languages are the following: predicational and not identificational reading, no life-time effects (cf. Musan 1995), compatibility with spatio-temporal modifications, activity reading and possibility of functioning as small clause predicates. Following Bowers (1993), she proposes that the small clause is headed by a functional head Pred which must be realized, through P-licensing in Irish and Russian and through N-licensing in French and German. An overt preposition realizes the Pred head in Irish. On the other hand, in Russian Pred is licensed by a covert preposition, which assigns instrumental case to its complement. In French and German, a noun through head movement realizes Pred. This analysis is supported by the lack of agreement and the impossibility of adjectival modification in eventive copular constructions in these two languages. Finally English is shown to pattern with Russian and Irish due to the results of the agreement and adjectival modification tests, i.e., English P-licenses the head Pred with a covert preposition.

According to the latest developments in the Minimalist Program (Chomsky 2000, 2001), derivation proceeds in phases-stages, at the completion of which the constructed structure is shipped to the two interfaces, i.e., Logical Form and Phonological Form, for interpretation and pronunciation. It is generally assumed that every phase feeds both interfaces simultaneously. In «What happens when phases get individualistic: On LF-only & PF-only phases» Franc Marušič explores the possibility of a phase feeding only the LF interface, leaving in the derivation process what should have been shipped to the PF interface. In such a derivation, what has been interpreted at some early point can be pronounced in a higher position, resulting in a configuration parallel to what we understand as total reconstruction. According to Marušič, this derivation is actually superior to the previous analyses since it avoids undoing operations and explains the agreement facts that the PF movement analysis cannot explain. Independent evidence for the existence of non-simultaneous phases is presented to strengthen the claim.

Another line of research emerging from this compilation focuses on the importance of language processing. According to this view, ease of processing influences

syntactic structures and is able to shape them in a way that facilitates parsing of the constructions by the speaker. Adverbials positioning, topic fronting, and pseudo-gapping and gapping are some of the processes where processing seems to dictate the resulting structures and their grammaticality.

Roberto Mayoral Hernández sits at the forefront of a new generation of linguists that start integrating the use of electronic resources in their research. In «A typological approach to the ordering of adverbials: weight, argumenthood, and EPP», Mayoral Hernández focuses on the positioning of frequency adverbials in Spanish. To this end, he has taken advantage of the online corpus CREA (*Corpus de Referencia del Español Actual* or Modern Spanish Reference Corpus by the Royal Academy of the Spanish Language), from which he has drawn a specialized data set consisting of sentences containing frequency adverbials irrespective of the position they occupy. Interesting observations arise that are relevant to linguistic theory (Cinque 1999, Hawkins 2000). For example, in the absence of an overtly expressed subject, frequency adverbials show a predisposition to occur right before the verb. On these grounds the author concludes that adverbials may satisfy the Extended Projection Principle, a condition enforcing the requirement that sentences have subjects.

Sofie Raviv's «Identifying and processing topicalization in Danish» proposes a processing constraint on topic fronting for Danish, arguing that this fronting takes place at the phonological level. Her main piece of evidence comes from the asymmetry between pronouns and full DPs with respect to topic fronting. While pronouns can front freely, full DPs are restricted to certain cases, namely when the resulting structure is unambiguous. Raviv introduces a constraint (IDentification) that requires unambiguous interpretation of the subject and object either by their word order or by the presence of an identifier. The paper presents four types of identifiers: negation and adverbs, auxiliaries and modals, case marking and verbal selectional restrictions. The Danish data show that fronting of full DPs is possible *only* when any of these identifiers is present in the sentence. Raviv follows Ertschik-Shir (to appear) arguing that fronting is the result of a preference for placing the main topic sentence-initially, linking it to the previous discourse.

She formalizes this preference as an alignment constraint that requires the main topic to align with the sentence initial position. Thus, the first element is always interpreted as the main topic. Therefore, as Raviv shows, fronting only takes place with objects functioning as main topics in those cases where there is more than one topic in the sentence. Raviv compares her proposal with previous syntactic accounts of topic fronting based on features (e.g. Rizzi 1997). These latter approaches face a number of problems in view of the Danish data, such as the apparent violation of the Minimal Link Condition by object fronting and the possibility of fronting non-contrastive topics.

In order to account for the parsing of elliptical constructions, Elixabete Murguía's «Antecedent-Gap Relations and Locality in Verbal Ellipsis» analyzes three verbal ellipsis phenomena using the minimalist operations Merge, Move and Spell-Out (Uriagereka 1999), and assuming economy principles as in Weinberg (1999). She analyzes verb phrase ellipsis, pseudogapping and gapping in English in terms of the different locality restrictions that apply to them. She proposes locality as the interaction of three factors namely, the presence or absence of tense, low initial attachment

of coordinates and Spell-Out operations, which render syntactic structures unavailable. The author shows verb phrase ellipsis and pseudogapping not to be subject to locality restrictions, contrary to gapping. The challenge for parsing elliptical constructions is detecting the gap and resolving it. In verb phrase ellipsis and pseudogapping the presence of an auxiliary allows us to predict the deleted VP, as in *Anne loves Peter and Mary does too* and *I gave money to Susan and Peter did to Beth*, respectively. The VP is predicted in a top-down fashion, without resorting to the antecedent. On the other hand, in gapping constructions such as *Susan prepared lunch and John dinner*, the absence of an auxiliary forces us to consult the antecedent in order to assign any structure to the gap and to interpret it. Locality must be respected (**Susan prepared lunch and I think John dinner*). Otherwise the sentence is ungrammatical.

The phenomenon of cross-categories is dealt with in several papers, where Spanish participles, and non-local anaphors are analyzed. The latter belong to hybrid classes that show features of different types of entities or categories. This mixture of features is seen in the syntactic behavior of the cross-categories elements.

Aysa Mondoñedo's «Nominal Participles, a Case of Categorical Alternance: Eventive Nominalizations in *-da*» is a new contribution to the studies of syntactic and semantic nominalization in generative grammar. By studying participles in Spanish, the author analyzes the formation and behaviour of a group of deverbal nouns in this language, which are formed by a verbal base plus the suffix *-da* (fem.sing. «-ed»). The suffixation of this morpheme gives rise to eventive, resultative and object nouns. Mondoñedo proposes that the first type of nominalization is a subset of a more general deverbal noun group formed with the suffix *-DO* («-ed»). Both formations (*-da* and *-DO*) share morphological, syntactic and semantic properties. However, eventive action nouns in *-da*, such as *La leída del Quijote de Juana fue emocionante* («Juan's [V "read" *-da*] reading of Don Quixote was moving»), have more restricted characteristics as opposed to the general formations in *-DO*. When semantic ambiguity appears, deverbal nouns in *-da* have a more eventive reading, occur in constructions with light verbs, have a regular morphology and a register in the borders of the lexicon. Mondoñedo suggests that these nouns are formed from a verbal root with an eventive feature. These nouns merge internally with the perfective suffix *-da* along the derivation and check their eventive features with a nominal category, giving rise to partially deverbalized nouns with the features [+N, +V]. Then, they merge with the determiner and become part of the DP. Other nominalization formations composed of a verbal base and the suffix *-DO*, such as resultative action nouns, show nominal properties but not verbal ones. Mondoñedo suggests that the first step to lexicalization would be to turn an eventive noun to a resultative one, having [+N, -V] features.

Gerardo Fernández-Salgueiro and Michael R. Marlo present data involving what is called *nonlocal anaphors*. «The non-local anaphor itself» seeks to outline several of the relevant syntactic and semantic properties of this phenomenon and offer an analysis built on previous research, that explains without stipulation a subject-object asymmetry in the distribution of non-local anaphors. As it is well known the anaphor *himself* must be locally bound (Condition A). Therefore, anaphors such as *himself* are generally banned from subject positions. In contrast, pronominals such

as *he* are locally free (Condition B), and co-reference to non-local antecedents or discourse entities is possible. Fernández-Salgueiro and Marlo present new data in which the pronominal *he* and the anaphor *himself* function together as a non-local anaphor in subject position. This compound anaphor is co-referent with a non-local c-commanding antecedent. The authors assume that this phenomenon is qualitatively different from examples where the anaphor *himself* functions as a VP-adjunct, as in *John did the work himself*. The non-local anaphor *he himself* has hybrid properties of both pronominals and antecedents. On the one hand, *he himself* behaves as a pronoun since it can appear in subject position, receiving an external theta role and checking Nominative Case. On the other hand, *he himself* behaves as an anaphor since it is necessarily bound by a c-commanding antecedent. Fernández-Salgueiro and Marlo pursue the idea that *he himself* has the syntactic distribution of a pronoun but the interpretation properties associated with an anaphor. They propose an analysis for *he himself* that captures its grammatical behavior straightforwardly, *himself* is assumed to be acyclically adjoined to *he* after *he* checks its Nominative Feature in Spec-Tense.

The sociolinguistics papers in this volume offer different cases in which language is the source of social conflict. This clash is present in post-colonial countries, where the presence of non-native languages such as English gives rise to different problems. However, the normalization of the linguistic situation benefits from sociolinguistic analyses that show the similarities in the evolution of these post-colonial languages and languages in their native environments.

Eric Anchimbe's «Multilingual backgrounds and the identity issue in Cameroon» introduces a number of historical sociolinguistic problems that bear on the identity issue. The ethnolinguistic atmosphere in Cameroon shows a complexity that results from historical events: the multilingualism of the pre-colonial period, the colonial establishment of French and English as official languages and the spread of Pidgin English. Cameroon stands as a bilingual state (English and French) but as Anchimbe claims, this is not a reality since few people achieve the status of bilinguals. Thus, assuming that language is central to identity (cf. Giles and Coupland 1991), the author singles out and explains a number of different identities in the multilingual African country. The official language identity is used to obtain a national feeling and has a political dimension attached to it. The ethnic identity is related to the native languages and the tribal groups. The bilingual identity is enjoyed by a minority, like the individual identity, which is characteristic of people from higher social classes. These four identities may be incorporated by the same individual, who might switch from one to another depending on the context. The lack of a unified identity has given rise to sociolinguistic calls of discontent. Anchimbe focuses on the Anglophone movement. The members of this group share English as an identity marker and argue that their language, and by extension themselves, is discriminated in relation to French since there is an inequality use of the official languages in national life. Anchimbe finishes his paper mentioning some solutions to the problem, which include giving attention to regional representation and promoting bilingual education in future generations.

In their paper «Sociolinguistic variables in the degeneracy of English in postcolonial (nonnative) contexts», Eric Anchimbe and Stella Anchimbe study the situation

of postcolonial English in non-native contexts such as Africa and Asia. There are two main approaches: some view English in these contexts as degenerate (Quirk 1985) and some others claim that non-native varieties of English show the evolutionary power of this language (Kachru 1985). Their paper investigates the sociolinguistics variables that lead to the present situation of these non-native Englishes in order to falsify the degeneracy hypothesis. Since the European powers' main interest was economic, they did not develop linguistic projects in the colonies. This, together with the lack of native teachers, led to sparse exposure of the indigenous people to native English. Furthermore as the authors state, the colonized people did not have a favorable attitude towards English and they preferred the use of their native languages or Pidgin English. These factors, combined with the ecology of the new context, led to a characteristic evolution of English in these regions. According to Anchimbe and Anchimbe, the language gets restructured in different aspects such as syntax, discourse strategies and lexicon, in order to adapt to the sociocultural and physical realities of the new environment, namely the colonized areas. Thus, non-native English evolves similarly to other native varieties, developing expressions and structures unique to the postcolonial regions.

Another interesting and innovative approach to language is that proposed by computational approaches. Computational linguistics is a field on the rise and its advances help develop different aspects of linguistic theory. Some of the present contributions make use of computational resources such as corpora to contribute to the understanding of several syntactic constructions. A different approach adopted in one of the papers is to syntactically parse a corpus and modify the notation system as the parsing advances.

May L.Y. Wong is a computational linguist involved in a personal project that aims at enhancing the PRF Chinese corpus. In «The compilation of a sample PRF Chinese corpus of skeleton parsed sentences», her goal is to go beyond the current annotation based on part of speech tags and incorporate other syntactically relevant information. On the methodology side, Wong favors skeleton parsing rather than full parsing for practical reasons and aspires to develop her own set of annotation labels to improve the current standard.

Most papers present new data with respect to previously studied processes. These new pieces of evidence call for a revision of former analyses. The authors make use of tools already existing in their framework to account for their data. This shows the explanatory power of different approaches since the combination of some of their parts can be used to explain new cases. Sometimes, the data bridges different accounts by presenting instances where features from both analyses are manifested.

In «Minimalist Edge Coordinations», Ángel J. Gallego analyzes the syntactic nature of the structures displaying a pseudo-coordination of two correlative elements (the «correlates») introduced by the focus/polarity particles *not (only)* and *but (also)*. Bianchi & Zamparelli (2001) dub these structures «edge coordinations», noting that two non-trivially different patterns can actually show up: an adjacent one (e.g., *John went not to Boston, but to London*), and a nonadjacent one (e.g., *John didn't go to Boston, but to London*), depending on whether the correlates form a continuous string, i.e., some kind of linear cluster, or not. Irrelevant as this may seem, the choice between an adjacent and a non-adjacent structure will bring about conse-

quences for agreement, truncation and parallelism effects. Following Brucart (1999), Gallego argues that «edge coordinations» are a type of «corrective negation», a structure where some constituent is replaced by means of a coordinated adjunct (e.g. *John, (and) not Mary, did that*), crucially dispensing with ellipsis. Therefore, he assumes that «edge coordinations» contain a Boolean Phrase, whose arguments establish a contrastive polarity checking. Furthermore, Gallego endorses the approach to focus put forth by Herburger (2000) and Irurtzun (2003), according to which focus corresponds to the nuclear scope of an existential quantification over events, whereas the non-focused material maps to the restrictor. Gallego's proposal contrasts with Bianchi & Zamparelli (2001) who resort to a cartographic-like approach whereby both patterns, although ultimately departing from each other in technical details, crucially involve an elliptical process of reduction. In a nutshell, the author proposes a minimalist analysis of «edge coordinations» that avoids the complexities of Bianchi & Zamparelli (2001), by invoking neither ellipsis nor the number of functional projections proposed in order to get the semantic effects.

Recent studies (Aoun et al. 2001 and Boeckx 2003, among others) show that resumption is not incompatible with movement, at least in some languages. In «Resumptive Pronouns and Matching Effects in Zurich German Relative Clauses as Distributed Deletion» Martin Salzmann contributes to these studies by analyzing the grammar of resumptive pronouns in Zurich German restrictive relative clauses. He reveals a new pattern of resumption present in the UG and argues for a movement analysis of these pronouns. Resumptives can appear in Zurich German in positions where movement is expected to be able to occur. Consequently, they cannot act as an argument for or against locality violations. On the contrary, reconstruction does play an important role in determining movement in resumptive constructions. The account of reconstruction effects under matching proposed in this study on Zurich German relative clauses is compatible with a movement analysis. Resumption is crucially linked to the licensing of oblique case and PPs, that is, the pronouns are needed to license oblique relations unless this licensing is performed by the external case of the head noun via a modified chain.

References

- Aoun, J., L. Choueiri and N. Hornstein, 2001, "Resumption, movement and derivational economy", *LI* 32: 371-403.
- and Yen-Hui A. Li, 1993, "Wh-elements in situ: Syntax or LF?", *LI* 24: 199-238.
- Barwise, J. and R. Cooper, 1981, "Generalized quantifiers and natural language", *Linguistics and Philosophy* 4: 159-219.
- Beninca, P. and C. Poletto, 2004, "A case of do-support in Romance", *NLLT* 24, 51-94.
- Bianchi, V. and R. Zamparelli, 2001, "Edge coordinations". Ms. SNS di Pisa/U. di Bergamo.
- Boeckx, C., 2003, *Islands and Chains. Resumption and Stranding*. Amsterdam/Philadelphia: John Benjamins.
- Bowers, J., 1993, "The Syntax of Predication", *LI* 24, 4: 591-656.
- Brucart, J. M., 1999, "La elipsis". In *Gramática descriptiva de la lengua española*, (eds.) I. Bosque and V. Demonte. Madrid: Espasa, 2.787-863.

- Burzio, L., 1986, *Italian Syntax: A Government-Binding Approach*. Reidel: Dordrecht.
- Chomsky, N., 2000, "Minimalist Inquiries: The framework". In *Step by step: Essays on minimalism in honor of Howard Lasnik*, R. Martin, D. Michaels, and J. Uriagereka (eds.). Cambridge: MIT Press, 89-155.
- , 2001, "Derivation by phase". In *Ken Hale: A life in language*, (ed.) M. Kenstowicz. Cambridge: MIT Press, 1-52.
- Cinque, G., 1999, *Adverbs and functional heads: a crosslinguistic perspective*. New York and Oxford: Oxford U.P.
- Erteschik-Shir, N., (to appear), "The sound patterns of syntax: the case of object shift". In *Theoretical Linguistics*.
- Giles, H. and N. Coupland, 1991, *Language: Contexts and Consequences*. Minton Keynes: Open U.P.
- Hawkins, J., 2000, "The Relative order of prepositional phrases in English: Going beyond Manner-Place-Time", *Language Variation and Change* 11: 231-66.
- Heim, I. and A. Kratzer, 1998, *Semantics in Generative Grammar*. London: Blackwell Publishers.
- Herburger, E., 2000, *What Counts: Focus and Quantification*. Cambridge: MIT Press.
- Irurtzun, A., 2003, "The Locus of [+F]ocus". Ms. EHU/UPV.
- Kachru, B. B., 1982, "South Asian English". In *English as a World Language*, Bailey, G. and M. Görlach (eds.), Ann Arbor: University of Michigan Press, 353-82.
- Matthewson, L., 2001, "Quantification and the nature of crosslinguistic variation", *Natural Language Semantics* 9: 145-79.
- Musan, R., 1995, *On the Temporal Interpretation of Noun Phrases*. Ph.D. dissertation, MIT, Cambridge.
- Perlmutter, D. M., 1978, "Impersonal passives and the unaccusative hypothesis". In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, 157-89.
- Quirk, R., 1985, "The English Language in a Global Context". In *English in the World*, Quirk, R. and H. G. Widdowson (eds.), 1-6.
- Rizzi, L., 1997, "The Fine Structure of the Left Periphery". In *Elements of Grammar*, L. Haegeman (ed.). Dordrecht: Kluwer, 281-337.
- Uriagereka, J., 1999, "Multiple Spell-Out". In *Working Minimalism*, S. Epstein and N. Hornstein (eds.). MIT Press: Cambridge, MA, 251-282.
- Weinberg, G. A., 1999, "A Minimalist Theory of Human Sentence Processing". In *Working Minimalism*, S. Epstein and N. Horstein (eds.). MIT Press: Cambridge, MA, 283-315.

SOCIOLINGUISTIC VARIABLES IN THE *DEGENERACY* OF ENGLISH IN POSTCOLONIAL (*NON-NATIVE*) CONTEXTS

Eric A. Anchimbe and Stella A. Anchimbe

Ludwig-Maximilians (University of Munich) and The University of Yaounde I

Abstract

This paper tackles from a broad historical perspective the attitudes, media and strategies of transmission, and the interplay of English and identity in the world today. It traces the negative tendencies towards non-native Englishes resultant from British colonialism to the hangovers and strategic linguistic schemes adopted during colonialism. Here the appellations non-native, postcolonial, indigenised, New Englishes are used interchangeably without purporting to make a profound evaluation of the bias linked to them, especially the non-native. The paper concludes with the note that the claim of degeneracy of the New Englishes was ignited by colonial linguistic projects and later fuelled by social prejudices built basically on colonial skeletons. It has less linguistic evidence and if any exists its roots are strongly founded in colonialism.

1. Introduction

Over the last two decades and especially between the late 1980s and early 1990s various accusations (the Quirk concerns and the Kachru catch) were launched in various directions. Several catchy expressions like 'Liberation Linguistics' (Kachru 1991) and 'deficit linguistics' (Quirk 1988) were coined to describe the itinerary of English in its spread worldwide. Two basic groups cropped to the limelight: the adherents of native speaker norms and the advocates of non-native speaker norms in New Englishes contexts. Whereas the native speaker considered postcolonial Englishes as a *degeneracy* of *his* language (Quirk 1985, 1990, Abbot 1991) and called for a common standard in these areas being the native norm, the non-native proponents (Kachru 1985, 1992, D'souza 1986, Bamgbose 1998,) perceived these varieties as vital proofs of the vitality of a language that had ceased to be the sole property of its native owners.

Why were and/or are postcolonial Englishes considered *degenerate*? Are they actually thus? If yes, who/what is to blame? These questions are fundamental to resolving this query which though seems to be settled is receiving new perspectives

like that of Mufwene (2001) which considers these Englishes to have the same structural evolution processes as the so-called native varieties. This paper visits some of the determinant sociolinguistic variables that have been *forgotten* in the quest for blame in the *degeneracy* claim. These include: the colonialists' intention of (not) teaching English to colonised peoples, their materialistic priorities, the impact of a long existent pidgin English, the abhorrence of English as foreign (intrusive) force, the absence of adequate native teachers, ecology and the linguistic gap, native languages and the physical background, and the (in)dispensability of the colonial language. These factors cannot be overlooked in determining if English out of Great Britain suffered heresy (Prator 1968) or not.

2. Colonialism and the legacy of English

If today English enjoys the status of a 'language on which the sun never sets' it is exceedingly thanks to the colonial empire of Britain. Even though the later international activity of the US in post-World War II era contributed to this spread and consolidation, the initial foundations were laid by Britain's colonial expedition into Africa, Asia and the Caribbean. The work of the religious missionaries, who combed some of these areas even before the arrival of colonial authorities, cannot be neglected. But whereas they limited their intervention to spreading the gospel and winning lost souls for the kingdom of God, the colonial authority engaged in the expansion of boundaries beyond Europe—a move that required much more than just a passive presence.

The colonial governments were therefore concerned, at various levels, with constructing in the colonised people a feeling of *belonging* to the colonial empire. They were called, in the case of British colonies, "Her Majesty, the Queen's subjects". A major weapon that was used, as a double-headed serpent, was language. Contrary to pre-WW1 German annexation that made little efforts to systematically institute German as the official language, the later British and especially French colonisations insisted on complete education projects and language schemes predominantly (for the British) and exclusively (the French) in the coloniser's language. The double-headed weapon was aimed first at inculcating in the indigenes a sense of unconditional attachment to the colonial power and second, it maintained a gap of status quo between them and the colonialists since they, as elaborated below, had little access to the variety used by the colonial masters. In spite of whatever strategies that were attached to the linguistic schemes in these regions, they yielded one thing—the emergence of postcolonial varieties of English. And as Kachru (1986: 1) points out "the legacy of colonial Englishes has resulted in the existence of transplanted varieties of English having distinct linguistic ecologies—their own context of function and usage". Colonialism simply added another dimension to the already complex landscape of languages in colonial contexts. It meant that English, or whatever language that was spread through colonial expansion, was introduced into a contact situation with several other languages. This contact now constitutes the basic landmark for the description of postcolonial English varieties as poor, less educated, degenerate approximations of the native. Let us start with a cursory look at some of the attitudes expressed towards these varieties of English.

3. Attitudes towards the New Englishes

The very many appellations coined for varieties of English that took root after the colonial adventure of Britain is ample proof of the divergent nature of attitudes towards them. These Englishes qua *non-native, new, indigenised, localised, nativised Englishes* represent the often ironically hailed diversity of English but limitedly accepted local norms and standards that accrue therefrom. What therefore is responsible for this?

In and out of regions where the New Englishes are spoken, many negative assertions have been advanced about them. To Kachru (1982: 66) the first enemy of the New Englishes is the nation states in which these Englishes are used. This has rendered the varieties more or less “linguistic orphans in search of their parents”. These parents cannot be the *native* varieties because there exist many differences between them. They could and perhaps only have foster parents through their acceptance within these regions. The second major enemy has been the native speaker who seems biased and influenced by the glimpses of victory through colonisation to think *his* language must not be equally shared with the colonised. These two perceptions, which from a superficial view show the defence of one’s position, project two major classifications of attitudes towards the New Englishes. One, the native, is highly hinged on colonial hangovers and the other, the non-native, is fuelled by realities in language change and transmission. Fig. 1 below recaptures this rift in perception.

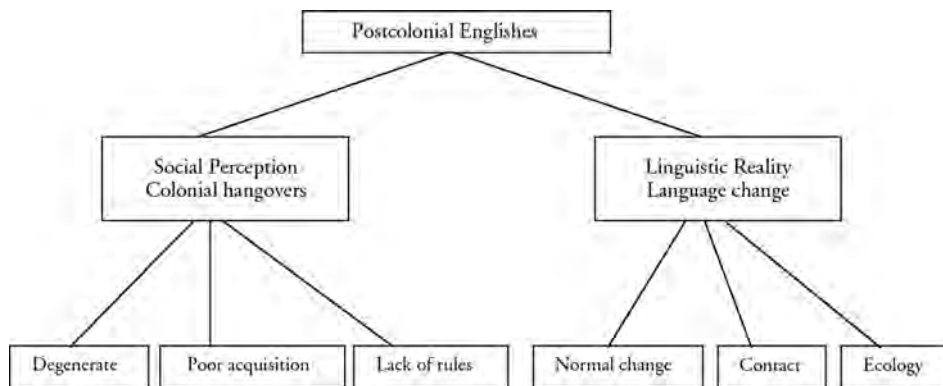


FIG. 1. Perceptions of Postcolonial Englishes

The classification in the diagram above shows how important a variable colonialism is in the perception of varieties of English and their media of diffusion in the world. It further indicates how non-linguistic parameters have been used in branding some varieties as *deficit* or *disintegrated* whereas language contact, the claim often used to justify this, is a constant residue of all languages whether colonially generated or not. Moreover, “linguistic change occurs even when no contact of languages is involved” (Mufwene 2001: 11). It becomes clear that whatever changes or variation occurred to English out of its native ecology forms part of the evolu-

tionary process of any natural language. It is not, as has been claimed, the exclusive outcome of non-native contexts nor the destruction of a *too perfect* language by *imperfect* users.

3.1. The Native vs Non-Native Divide

The battle of standards that was at the centre of debates on the New Englishes in the 80s and 90s was grossly rooted on the great divide of native and non-native speakers. While Quirk (1985, 1990), Abbot (1991) and so forth called for a monochrome standard both in writing and speech around the world (however difficult this could be), Kachru (1985: 92, 96) insisted on regions and nations developing individual standards according to the tastes and dictates of their societies. Quirk's (1985) preoccupation is that postcolonial nations do not have enough functions for English and so must not be granted the right to develop a standard for a language that they would not be able to master given their multilingual statuses. This position is tantamount to the fear of seeing English *degenerate*. It is further disputable that English has fewer functions in postcolonial countries. It is nevertheless the official language and is the major medium linking these countries to the rest of the world.

Even earlier than this period, Hocking (1974) made a rather unrepentant declaration about what the standard should be. To him, "the point is that what is correct in a language is just what native speakers of the language say. There is no other standard". His view, which perhaps influenced later proponents of this inclination, awards the native speaker an almighty control of his language. It confirms Chomsky's (1965) consideration of the native speaker as one who can make valid judgments about what is well-formed or ungrammatical in his language. Although this is unquestionable, its assumption in the context of postcolonial Englishes is too estranged to make any much meaning. Furthermore, Hocking (1974) does not realise as Trudgill (1998: 35) does that "most native speakers of English in the world are native speakers of some non-standard variety of the language". In all, with the amount of literature produced on this topic a great consensus seems to have been reached which favours local norms over foreign ones. The defence of a language that ceased to be the sole property of Britain yielded to the recognition of ecologically pertinent factors that rendered and continue to render homogeneity a fairy dream.

3.2. The ELT Industry and the claim of heresy

The ELT (includes ESL, EFL, TESOL, etc.) industry, which today is among the largest and fastest growing, was at its genesis in the 1960s shelled with horrifying missiles born predominantly from the "native speaker's fear of seeing *his* language disintegrate in the hands of (or shall we say, on the lips of) non-native users" (Kachru 1985: 34). This fear, a seemingly coordinated appendix to the colonial strategy of not teaching English too well or at all to colonised peoples, retarded many genuine attempts at vehiculating the language especially in former British colonies. As table 2 shows far less students were engaged in education in English in 1974 as opposed to later years when the stigma of colonialism started disappearing. To Prator (1968) for instance, it was pure heresy to teach English to non-natives

and even worse to grant them the right to 'own' a standard or norm. He (1968: 459) emphatically declares that

...the heretical tenet I feel I must take exception to is the idea that it is best, in a country where English is not spoken natively but is widely used as a medium of instruction, to set up the local variety of English as the ultimate model to be imitated by those learning the language.

The issue Prator (1968) seems not to be comfortable with is not English spreading but it spreading to non-natives who will not be able to use it properly. He fears the language would *disintegrate* or *degenerate* if allowed to evolve as an independent model. The cline of fears and exasperations expressed prompts Bamiro (1994: 58) to segment attitudes towards postcolonial Englishes into two schools. The first is the "sociolinguistic reality school" represented principally by Kachru and other advocates of the New Englishes. This school argues for the recognition and unprejudiced acceptance of the New Englishes as part of the diversity of the English language world. It posits above all that English has adopted and adequately sipped into the sociocultural environments in which it is used as a second language to a point that judging it in terms of native standards is absolutely unfair and illogical. The second school which Bamiro (1994: 58) terms the "pedagogic unreality school" and represented by its major exponents Quirk (1985, 1988, 1990), Abbot (1991) upholds that the "New Englishes are nothing but grammars or dialects of errors which are bound to have deleterious effects on the educational systems of many countries where English is used as a second or foreign language". The *degeneracy* claim studied in this paper was extensively sustained by proponents of this school. Their negative attitude towards non-native speakers was perhaps founded on the hierarchical order set in place by colonialism that equated colonised peoples to *tabula rasas* on whom the *civilisation* of Europe had to be written. Besides this, any meaningful assessment of a language out of its native habitat and spread and/or taught by non-native speakers must make enough allusion to change and variation. This like any other type of adaptation (human, environmental, etc.) foregrounds the replication of the new environment on the language which in effect is a favourable mid-range reconciliation between the foreign status of the language and its new habitat rather than a *disintegration* of the language as such. This reconciliation, often treated as *degeneracy*, was in part promoted by some sociolinguistic variables beyond the control of the non-native learners. Some of these have been outlined below.

4. Some sociolinguistic variables in postcolonial English claim

If English can be said to be *degenerate* in postcolonial contexts, the main blame cannot however be directed at the non-native users. This is because the colonial administration through whom English was substantially transmitted foiled the process with a series of strategic projects that unilaterally cared only for their prominence in power and authority in the regions. Change or subsequent evolution of the language in these areas was therefore conditioned by these factors. The

major ones are studied below. While the following variables are considered important to the path of English in these contexts, the ecological factor is not treated in any lesser importance.

4.1. Colonial (mis)conceptions of (not) teaching English well

Irrefutably, language is power or power takes a more decisive turn when it wields language. Language is an important attribute in identity creation and consolidation. It is moreover an effective tool for in-group exclusion and definition. These hints certainly guided the colonialists, among them the British, in the adoption of an ineffective language teaching approach in colonial areas. Whatever language they used with the indigenes was only as good as it made clearly evident the gap and distance between them and the indigenes. Any attempt at using the colonised peoples' language or letting them full access to the colonisers' language was interpreted as tantamount to levelling the great mounts of master-servant, ruler-ruled, privileged-unprivileged, modern-primitive, advanced-backward, etc distinction that existed between them. As Kachru (1986: 22) puts it, the colonialists "insisted on not teaching their language too well to 'non-group' Asians or Africans, the underlying idea being that the colonizer's code, if shared equally with the colonised, would reduce the distance between the rulers and the ruled". This was not exclusively the strategy of the British alone but of most of the colonialists. The Germans in Cameroon exposed this same attitude. Amvela (2001: 206) states that "some [German] officials also feared that the use of English may encourage Cameroonians to behave as the equals of their colonial masters".

The immediate outcome of this policy was the explosive growth of and reliance on Pidgin English: in Nigeria, Ghana, West Cameroon, etc, lingua francas: Swahili in Tanzania and Kenya, Krio in Sierra Leone and Liberia, Tok Pisin in Papua New Guinea and so forth. This growth was facilitated by the colonialists who intentionally promoted the use of these languages. In their daily communication with the indigenes they preferred a pidgin, Creole, indigenous language, as the case may be, in a bid to refurbish the master-servant relationship between them. In the case of East Africa, Abdulaziz (1991: 395) reports that "British settlers were most reluctant to use English with their native servants and with Africans in general, as they believed that this knowledge might 'spoil' them in the master-servant relationship that existed". Native languages also flourished but were checked by the generally excessive number in rather small communities such that pidgins, creoles and lingua francas were often invoked for wider communication. This notwithstanding, the linguistic foundation laid during colonialism that constituted in less effective usage of the language continued to thrive decades after. Many souring descriptions have been given which point to the state of English in these contexts. One of them is Gyasi's (1990: 24) who bemoans: "English in Ghana is sick. The cancerous tumours are numerous: wrong collocation; false concord; poor spelling ... mispronunciation; ... wrong omission or insertion of articles, misuse of preposition...". Gyasi's (1990) "cancerous tumours" might simply be some of the leftovers of the colonial experiment of *unteaching* or *misteaching* English as a power regulatory mechanism. While also accepting that societies

be granted the right to use language according to their demands and tastes, which accounts for much of the divergence in English speech today, the colonialists' impact cannot be altogether sidelined. Clearly, therefore, the claim that English suffered *heresy* or *degeneracy* in postcolonial or non-native contexts is prejudiced and socially motivated. As shown on fig. 1, it is simply an attempt to maintain colonially drawn skeletons.

4.2. Colonialists' notion of hierarchy through language

Societal stratification is often linked to and/or represented in its language. For instance, the most prestigious dialects or standard of the language is often identified with the socially privileged. Trudgill (1998: 39) in defining Standard English makes clear that "the further down the social scale one goes, the more non-standard forms one finds". What this means is that the various strata that can be identified in society can as well be graded on a linguistic ladder. European colonialists adopted this linguistic ladder framework to widen the distinction gap and to create a social elite situated between them and the common masses. As R.R. Roy and T.B. Macauley (1835) observe, teaching English to colonised people was directed at putting in place "a class who may be interpreters between us and the millions we govern, a class of persons, Indians in blood and colour, but English in taste, in opinion, in words and in intellect" (*qtd* Kachru 1982: 355).

A new stratum is being added to an already diversified society. In the first place, this suits what is discussed below as deprivation from normal exposure to the language. It further represents the refusal of belonging to a language that is bound to be one's official code of communication. The creation of such a class of people completely subservient to the colonial administration was motivated by fear of subversion and the desire to eternally assimilate and subjugate the colonised. So as Mazrui and Mazrui (1996: 272) clearly state, "many European settlers regarded the teaching of English to 'natives' as a potentially subversive force". Whereas military force played a great role in physically subjugating colonised people, the language policies (whether outlined or implied) adopted a psychological strategy that limited the range of perfection and proficiency of the out-group. This meant poor acquisition and subsequently poor performance in the language. Anybody, whether colonised or not, subjected to such circumstances would end up with the same or highly similar results. The poor rendition of the language must not be blamed on origin or status of speakers but rather on such mitigating situations as conditions of acquisition or transmission, variety of the language involved and the length of the period of transmission.

4.3. Material priorities over linguistic projects

The spread of colonial languages to (ex)colonial regions was not an exclusively linguistic experiment but a fall off from the materialistic incursion. The Green Revolution brought Europe to a sophisticated level with a flourishing industrial output that constantly needed more markets and more raw materials to keep it alive. Markets and raw materials were then found in the territories of Africa, Asia

and the Caribbean. However, this expansion coincided with the expansion of religious missions. So while the colonialists sought for material products and the expansion of political empires the religious missionaries embarked on preaching the gospel. They adopted whatever language could facilitate this objective. This ranged from Pidgin English, native languages to educated English. It must be mentioned that the British colonial administration in Cameroon benefited tremendously from the work of the missionaries who were in the region even before the German annexation of 1884.

The priority of material interests can be seen in the Germans' toleration of Pidgin English and English inasmuch as these languages facilitated the acquirement of raw materials and the construction of roads, railways and plantations. It is interesting that the pre-annexation undertaking and the 1884 annexation treaty between German authorities and Douala chiefs of Cameroon were both done in English and not German (see texts in Amvela 2001: 219-221). The English colonialists were not different from the Germans. They were more interested in consolidating the wealth, given the high costs of WWI, than in putting in place a solid linguistic project based on perfect English transmission. This is further explained by the fact that they tolerated (as opposed to the French who banned) education in native languages; themselves encouraged the use of Pidgin English in several sectors, for instance, trade and religion. Added to the biased vision exposed above, the lack of a devoted linguistic project that matched the creeds of allegiance to the queen and the British Empire recited by the colonised subjects meant a divergent acquisition of the language. This divergent acquisition, which however depicts a natural situation of language acquisition and evolution, has unfortunately been received as a destitution of the language and as, in the words of Whitworth (1907: 6) "linguistics flights ... which jar upon the ear of the native Englishman". Miraculously, the native colonialists' decision to *unteach* or *misteach* English as power regulatory mechanism has been forgotten. Non-native background and the status of secondness of the colonial subjects constructed during colonialism have been highlighted as sources of the divergence in English usage around the world. A divergence that has been generally treated as a *destruction*, *degeneracy*, *demeaning* or what Prator (1968) calls *heresy* of the language.

4.4. Impact and pull of Pidgin and/or Creole Englishes

Trade colonial expansion that debuted in the late 15th century heralded by Portuguese merchants installed Pidgin English along the coast of West Africa (Schneider 1974, Mbassi-Manga 1976, Mbangwana 1983). Although the Spanish, Swedes and Dutch were also involved their appearance on this coast was less regular compared to that of the Portuguese who engaged in the trade of diverse articles including spices (pepper), gold and slaves. When the British finally replaced the Portuguese on this coast following the build up to the abolition of slave trade, one of the pidgins that was used by English privateers on Portuguese boats gained more expansion. As rightly explained by Mbangwana (1983: 80),

inasmuch as the British were the first advocates of the abolition of the slave trade and at the same time practised the 'factory and trust' system of trade, which

brought them into very close contact with the native inhabitants, a language contact interaction emerged which served as a linguistic medium of communication.

The work of the missionaries further compounded this medium since it, beyond the scope of the notoriously many and diversified native languages, provided a broader spectrum of communication with the indigenes.

The arrival of the Germans in Cameroon after the 1884 Berlin conference did not create any much difference in the place and vitality of Pidgin English. Similarly the end of the slave trade marked the emergence of a more stable and mother tongue-like variety of Pidgin English. This was in communities such as Liberia, Sierra Leone and Fernando Po basically made up of resettled slaves. Menang (1979) advances that the variety of pidgin used by these communities eventually turned out to be not just a medium of contact and communication but a practical mother tongue for a group of divergent people who found themselves bound to live forever together. The German colonialists did much to impose their language on the Cameroonian natives, opening German-medium schools in Douala (1887), Victoria (1897), Garoua (1906) and Yaounde (1908); declaring German the only language in all education-related transactions after the Douala Educational conference (1907); issuing a special ordinance on April 24th 1910 “with the stipulation that grants-in-aids from the government to mission schools would be restricted only to those who adopted the government school programme based on German...” (Chumbow 1980: 284) and officially making the use of English illegal by March 1913 (Amvela 2001). In spite of all these rather radical measures, the colonial administration was unable to proceed in German given that it meant dismantling the whole edifice of Pidgin English constructed over several centuries. So it literally tolerated its use in the plantations, the road and railway construction sites, and sometimes used it as a contact code with the population.

When the First World War ended and Britain was rewarded with Cameroon, Pidgin English rather grew faster following the free language system adopted by the British. Even in ESL, the aim was not to recreate British English in colonised areas. Some reasons have been advanced above. Inasmuch as English was interpreted as a socially superior language its encroachment to the circles of communal and interpersonal transaction was limited. It is in this domain that Pidgin English triumphs. Its long existence as a contact and friendship code gives it more recurrence than English and partly explains why English expression is often considered dwindling or non-proficient. The table below exposes proficiency in these languages through ability to speak them. While only 1% of 433 people in Bamenda speaks only English, as much as 24% speak only Pidgin English and 43% both. This indicates that many more come into contact with Pidgin English than do with English; as many meet Pidgin English before any contact with English.

It is worth noting that all of the towns surveyed above are in the English-speaking part of Cameroon. Interestingly, more people speak French and Pidgin English in all of these towns than speak only English. This indicates that the lack of proficiency in the language, if this can be equated to the *degeneracy* claim, must be interpreted as a matter of language priority by the users rather than as a basic feature of

TABLE 1. Percentage of adults who speak official languages and Pidgin English in Cameroon (1983)

Town	No. of resp.	Pidgin only	English & Pidgin	English, Pidgin & French	French & Pidgin	English only	Total % of population
Victoria	371	29	50	16	2	—	97
Buea	185	11	49	29	11	—	100
Kumba	364	38	46	10	4	—	98
Mamfe	87	29	53	16	2	—	100
Bamenda	433	24	43	20	11	1	99
Kumbo	99	38	43	11	2	—	94

Source: curled from Koenig (1983: 51)

postcolonial Englishes. In Kumba, a predominantly business town along the Nigeria-Cameroon border, for instance, the difference between the pidgin only percentage (38%) and the English and pidgin percentage (46%) is less than 10%. But this is far larger in administrative headquarters like Bamenda where the difference is 19%, Buea 38% and Victoria (present day Limbe) 21%. In these towns reside state-employed workers, students and workers in education-related private jobs. *Degeneracy*, if at all must be admitted should therefore be founded on the preferences speakers give to languages in their quest for economic survival. It is not exclusively depended on native and non-native statuses as claimed by Hocking (1974), Prator (1968), Quirk (1990), Abbot (1991) among others.

4.5. Abhorrence of English as icon of invasion and exploitation

Like any intrusive force in communities of people, English suffered repudiation and abhorrence in colonial and postcolonial states. Such a staunch resistance to the colonial languages took strength from the fact that the African continent was considered an empty set into which colonial civilisation with all its components had to be stuffed. For instance, the *tabula rasa* approach of the French that “aimed at assimilating or absorbing France’s colonial subjects to the point where they would actually be Frenchmen linguistically, culturally, politically and legally” (Fanso 1989: 65), was certainly bound to meet with opposition. It was like wiping out any footage of the pre-colonial heritage and replacing it, like writing on a virgin sheet of paper, with the European cultural and linguistic heritage. The expression of abhorrence ranged from prohibiting African children from attending European schools, humiliating those who spoke English in non-official contexts to refusal to use the colonial language in certain (official) contexts. In Ghana, for example, Kwasi Duodu (1986: 3) in his support for the use of a Ghanaian language as official medium declares.

If we can’t decide on one Ghanaian language for the country after twenty-nine years of independence, then why shouldn’t a borrowed language be ‘butchered’ ... the youth, like many other silent Ghanaians, is protesting against an imposed language which prevents him from expressing himself in his own tongue.

Duodu (1986) implies that the youth, like many other Ghanaians are resisting English. Of course resistance would lead to lack of proficiency in the language since it is considered alien. Tchoungui (1983: 114) exposes the outcome of the adoption of a French-English bilingualism policy in Cameroon. To her, it “evinces a remarkable inability to live or to think out of well trodden colonial tracks, it actually opts for more educational wastage as children scrambling for more education are schooled in languages other than their own, or worse, in languages alien to their own cultures”. This negative perception, that reached fever pitch at the close of colonialism in the late 50s and early 60s, increased community and missionary work in the direction of promoting indigenous languages. This in part explains why up to 166 of the 270 living languages in Cameroon have been standardised (*Ethnologue*). However when national unity became threatened by the continuous empowerment of native languages and its corresponding political and social superiority, education in indigenous languages was banned in Cameroon in 1965. This ban that was enforced by forceful actions including confiscation of technical equipment and pedagogical environment, for instance in Dschang in 1966 (see Momo 1997), increased abhorrence for the official languages thereby limiting the extent of attachment to them. In this light therefore, lesser people became interested in English and the few who engaged in learning it did so with a rather reckless attention.

4.6. Absence of native teachers

Given that Britain practised distance administration with fewer British men on colonial ground, there was the stark absence and paucity in the number of native teachers to teach English to the indigenes. Moreover, the few that were available were too busy with colonial exploitation schemes to dedicate much time to teaching English to many Africans or Asians. This explains why only scores—a generation of interpreters—were educated and charged with vehiculating the language further. Similarly the missionaries easily adopted Pidgin English or Creoles and in some other cases the indigenous languages to spread the gospel.

If it can be truly claimed that English suffered *degeneracy* in these contexts, a substantial blame must be directed at the colonial authority for not providing enough native teachers to properly teach it. But if the language continues to be taught by non-natives as it has been since the end of colonialism, it would be bound to reflect the ecology of the areas in which it is being used. As reiterated below, this is not negative or detrimental; it simply adds a creative dimension to the language that exposes its vitality and adaptability.

4.7. Deprivation from exposure to English

It is undoubtedly true that colonialism was central to the spread of English. It is however also true that colonialism, as shown above, moulded the cline of performance and proficiency in the language. In the transplanted native varieties of English, transplanted so many centuries ago—America, Australia, New Zealand, Canada, the colonised people were allowed sufficient access to the language. There was, especially in the cases of America and Australia, the settlement of substantial

native English speakers. This colonisation pattern is different, in terms of its linguistic agenda from later ones witnessed in Nigeria, Kenya, India, etc, where “sparser colonial settlements maintained the precolonial population in subjection and allowed a proportion of them access to learning English as a second, or additional language” (Leith 1996: 181). The limitation of access to educated English meant the acquisition of the language in whatever manner possible and with whatever imperfection that could bring.

Even long after the colonialists were gone, many people were still far from exposed to the language. As late as 1974, education in English was still not completely accessible even in British ex-colonies that had English either as the only official language (Nigeria, Ghana) or together with other language(s) (Kenya, Tanzania, India, Cameroon and South Africa). This is evident in the following statistics, which reveal the number of students enrolled in English-medium schools in six British ex-colonies. Of a total of 195.452 million people only 14.9% (29.3 million) of the students were enrolled in classes with formal instruction in English. This number simply adds to that of other users of the language.

TABLE 2. Access to English in 1974

Country	Total pop est. 1974 (million)	Students enrolled in English (million)	Percentage
India	90.486	17.6	19.4
Nigeria	52.895	3.9	7.3
South Africa	22.458	3.5	15.5
Kenya	11.208	1.7	15.1
Ghana	8.631	1.6	18.5
Tanzania	9.774	1.0	10.2
Total	195.452	29.3	

The population estimates above are culled from the United Nations Environmental Programme (UNEP) African Population Database while the student enrolment statistics are supplied by Gage and Ohonnessian (1974, 1977). The total populations exceedingly drown the numbers of those studying in a language that is considered official to the nations. Only 3.9 million students out a total population of 52.895 million in Nigeria and a similarly small number, 3.5 million of 22.458 million people in South Africa, were added to the existing number of users of the language in 1974. Although such factors as the lack of schools, insufficiency of teachers, lack of educational motivation and so forth can be used to explain the insignificant number of those learning English, it can also be blamed on the colonial policy. This policy created a linguistic elite that served as a link between colonialists and the colonised. It drew the line between the languages or varieties reserved for them and those open to the indigenes. This implicit distinction distanced English from the people who thereafter considered it the colonialists' code. This is especially the case in Tanzania and Kenya where (Ki)Swahili was promoted even

more than English. It is not therefore surprising that only 10.2% and 15.1% of the total population was engaged in English instruction in Tanzania and Kenya respectively. For Kachru (1985) the above figures are impressive. However, the truth behind them can only be judged if the length of the colonial expedition is revoked. British colonial expansion in all of these areas lasted above half a century. Besides the work of the missionary churches and that of the colonial administration in instituting education and literacy, much was left undone given that only a microscopic elite benefited from it. And as said above it was directed at creating an educated minority elite for the expansion of the colonial administration and for the continuity of the colonial heritage even after independence.

The deprivation set up at colonialism and inherited at independence accounts in part for the varieties of English spoken in these areas. It thwarted every possible prospect of native-like varieties taking root and mounted the foundation on which these indigenised varieties are built. However, it also accentuated the cry of *degeneracy* of the language that was issued in the late 70s and 80s and re-echoed in the 90s. According to estimates by Graddol (1997: 11), while English received a tremendous increase in users around the world, second language users still numbered far less compared to the populations of their countries. How can it be explained that less than half (43 million) of the population of Nigeria (90 million) speak English although it (up till 1995) was the only official language? The following table further reveals how stagnant percentages of users of English have been.

TABLE 3. Speakers of English as second language (1997)

Country	Total pop. Est. 1990s (millions)	English users (millions)	Percentage
India	130.985	37.0	28.20
Nigeria	90.987	43.0	47.20
South Africa.	37.066	10.0	26.90
Kenya	22.214	2.576	11.50
Ghana	14.466	1.153	7.90
Tanzania	16.227	3.0	18.40
Total.	311.945	96.729	31.08

Of a total population of 311.945 million people from the six countries only 96.729 million (31.08%) understand English. Ghana records the least score with only 1.153 (7.9%) of its 14.466 million population being able to use English even though it is the official language. Only less than a third of the populations of India and South Africa, and even less than a quarter of Kenya, Ghana and Tanzania, as shown on Graddol's (1997: 11) statistics can be considered second-language speakers of English. This does not of course mean the rest of the population are first-language speakers because in essence there are none except for a handful in South Africa. Although this might be explained by the fact that the official languages are used basically for official functions that do not often concern the common man,

these statistics beat down the long-sung story of literacy. It further lays bare the limited initiatives made to promote the use and extensive acquisition of the language. One of the outcome of this has been the common believe that English in these areas is *deficient* or *degenerate*.

4.8. The gap between foreign English and the sociocultural environment

Languages like living organisms evolve in a given ecology. This ecology, which has an internal and external component, regulates in several consistent ways the life and evolution of the language. Ecology simply refers to the sociocultural and geophysical landscape within which a language evolves. Mufwene (2001: xii) arrives at a distinction between the external ecology that covers the socio-economic and ethnographic environment together with “the contact setting and power relations between groups of speakers” on the one hand and the internal ecology that extends to “the nature of the coexistence of the units and principles of a linguistic system before and /or during the change” on the other. Both are equally significant to judging how elaborately a language has changed in its new habitat. The New Englishes basically emerged from the transportation of English to new ecologies where it had to exist. Its successful existence meant it had to adapt to and adopt from these new ecologies in order to represent them properly. Kachru (1992: 2) clarifies that

once English was adopted in a region ... it went through various reincarnations that were partly linguistic and partly cultural. The reincarnations were essentially caused by the new bilingual (or multilingual) settings and by the new contexts in which it has to function.

Reincarnation or nativisation or indigenisation (it has been termed differently), of English in these contexts serves to fill the gap caused by the foreign status of the language in its new context of existence. It has to reflect and be reflected by the physical realities and the sociocultural emblems of the society of which it is now an integral part. Along the West African coastline several vocabulary items are shared which do not belong to the British English vocabulary. These include *bitter-leaf*, *corn-chaff*, *bush-meat*, *head-tie*, *watch-night*, *chewing-stick* and so forth (see Anchimbe 2004). Although all of these words are English if treated individually, they have been compounded in a way that reflects the region in which they are used. It is no longer strange to find native language words and other neologisms created to fill communicative gaps in second or foreign language contexts. The recreation of the ecology in language may extend beyond simply the creation and addition of new words to larger linguistic units as collocational preferences, analogical creations, sentence structure and discourse patterns. It might and often generally result in extensive restructuring of the language to suit the communicative habits of the speakers. So restructuring in this manner must not be *pro rata* to non-native or postcolonial heritage. Mufwene (2001) and Schneider (2000) uphold that the restructuring patterns are basically the same in all languages whether termed Creole, koinés, pidgins, non-native or native. In a nutshell therefore the evolution of the New Englishes cannot be singled out as cases of *degeneracy* or *deficiency* since English itself has had as much contact in Britain as any of the Englishes out of Britain.

Restructuring follows several ways. In the New Englishes, it adopts predominantly a straightening approach. It seeks to name the referent as accurately and descriptively as possible. One common example is the ascription of the *s-plural* to non-count nouns such as *advices*, *furnitures*, *equipments* just to name these. So rather than use many words just to create plural as in *pieces of advice*, most New Englishes simply apply the *s-plural*. This reduces the number of words (*advices*) and above all resolves a long (and perhaps illogical) explanation for exemption to this rule. In the table below, different words for the same referent from three Englishes are presented. These are Cameroon English, British English and American English (see Mbangwana 2002: 123-130 for more). Different restructuring processes account for the differences in the appellations. In CamE *concierge* may have been supplied by French, *roommate*, *face towel* and *cargo train* by exact descriptions of either the referent or what they are used for. A *roommate* is someone with whom you share the same room (see also *classmate*, *age mate*, *desk mate*, etc.); a *face towel* is used in cleaning the face and a *cargo train* carries cargo. Each appellation follows the priorities of the society.

TABLE 4. Lexical variation in three varieties of English

British English (BrE)	American English (AmE)	Cameroon English (CamE)
House of Commons	House of Representatives	House of Assembly
Lodger	Roomer	Roommate
Public prosecutor	District attorney	State counsel
Town hall	City hall	Municipal/City Council hall
Face flannel	Washcloth	Face towel
Goods train	Freight train	Cargo train
Pig-sty	Hog-pen	Pig fence
Caretaker	Janitor	Concierge
Sitting room	Living room	Parlour
Vest	Undershirt	Singlet

The differences in appellation do not make the varieties inappropriate media of communication within their communities. They do not show either whether any of the varieties is less effective than the other(s). What we must ride home with is that languages and/or varieties possess the capabilities of reacting to and adapting to the changes and tastes of the societies in which they are. Such changes or adaptations do not necessarily indicate negative turns in the language because if a language must spread, it must also be ready to change. Many factors have been cited above to illustrate how the change in the New Englishes, which has been interpreted by some bi-ased linguists as *degeneracy*, was triggered by certain sociolinguistic variables. But is this true?

5. Is postcolonial English *degenerate*?

After the above mitigations on the status of English, the overall question is, is postcolonial English *degenerate*? A legitimate answer to this question is found beyond the realms of social prejudice and bias that over the years veiled any genuine investigation. As Singh et al. (1995), Mufwene (2001) among others truly demonstrate there is adequate linguistic evidence that non-native English varieties evolve in the same evolutionary patterns as other normal (native) varieties. Language development, evolution, change and the contact variable are common to all language contexts not only the postcolonial. To consider them detrimental only in the postcolonial contexts is to shut out the many processes involved and to put into question theories of language evolution and change. It is simply the phobia of *our* language in *their* hands.

The phobia of the native speaker hatched social prejudices towards these varieties. Moreover colonial schemes, as explained above, moulded the cline of change. The prejudice gave the impression that non-native Englishes posed intelligibility obstacles. This orchestrated cautionary pieces of advice like the following from Trudgill and Hannah (1994: 122) who advance that “native speakers travelling to areas such as Africa or India should make the effort to improve their comprehension of the non-native variety of English ... rather than argue for a more English-type English of English in these areas”. In a similar manner, Adebija and Bello (2001: 105) advise that “as speakers of English move from one part of the English-speaking world to another, they need to make greater allowances for apparently unorthodox Englishes, senses and usages of words in their English lexical repertoire”. Indisputably the focus here is on native speakers and the fear that they may not understand non-native or what Adebija and Bello (2001) term “unorthodox Englishes”. No measure is taken for non-native speakers travelling to native or rather “orthodox Englishes” areas, whereas intelligibility is a mutual exchange and not the ultimate burden of the non-natives. This technically implies that the *degenerate* variety (as postcolonial English is believed to be) must live up to the normal —a perspective that has transformed research in the New Englishes to panoramic judgements of how *deviant* from native (British) English these varieties are.

6. Conclusion

Postcolonial Englishes are not as *rough* as thought. Instead there exists more logicity and easy-to-apply rules in these varieties, like the *s-plural* above, than in British English. While the sociolinguistic variables studied in this paper ignited a process of change and evolution, the ecology rolled the dice as in all contexts of language contact. These simply point to the vivacity of the language. So, “rather than act as if the language is being debased”, Yule (1996: 64) proposes that “we might prefer to view the constant evolution of new terms as a reassuring sign of vitality and creativeness in the way a language is shaped by the needs of its users”. Of course this foregrounds a broad range of differences across societies that use English. Difference of this nature has often been generally interpreted as a breakdown in inter-

national communication. This has not been the case, at least in writing, because the whole English language world is intricately linked through the language in education, diplomacy, publications, trade and business, aviation and so forth. Ogu (1986: 93), therefore rightly concludes that “difference or variation is not a deficiency, receptiveness is not necessarily a submission and that complementarity is what makes relationships between languages [varieties and users] possible and pleasurable”.

References

- Abbot, G., 1991, “English across Cultures: The Kachru Catch”, *English Today* 7 (4), 55-57.
- Abdulaziz, M. H., 1991, “East Africa (Tanzania and Kenya)”. In Chesire, J. (ed.), *English around the World: Sociolinguistic Perspectives*. Cambridge: CUP, 391-401.
- Adegbija, E. and J. Bello, 2001, “The Semantics of ‘okay’ (OK) in Nigerian English”, *World Englishes* 20 (1), 99-106.
- Amvela, Z., 2001, “The English Language in Cameroon: The German Period”, *Epasa Moto* 1, (4), 194-225.
- Anchimbe, E. A., 2004, “Creating New Names for Common Things in Cameroon English”, *The Internet TESL Journal* X (8) of August 2004. <http://iteslj.org/Articles/Anchimbe-CameroonEnglish.html>.
- Bamgbose, A., 1998, “Torn between the Norms: Innovations in World Englishes”, *World Englishes* 17 (1), 1-14.
- Bamiro, E. O., 1994, “Lexico-Semantic Variation in Nigerian English”, *World Englishes* 13 (1), 47-60.
- Bokamba, E., 1992, “The Africanization of English”. In Kachru, B. B. (ed.), 77-98.
- Chomsky, N., 1965, *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chumbow, B. S., 1980, “Language and Language Policy in Cameroon”. In Kale, N. K. (ed.), *An African Experiment in Nation Building: The Bilingual Republic of Cameroon since Reunification*. Colorado: Westview Press, 281-311.
- Duodu, K., 1986, “The Language Problem”, *People’s Daily Graphic* 5, Sept. 1986.
- D’Souza, J., 1986, “Codification of Non-Native English: Is it Necessary/Possible?”, *Studies in the Linguistic Sciences* 16, (1), 1-11.
- Ethnologue. Languages of the World*. <http://www.ethnologue.com>.
- Fanso, V., 1989, *Cameroon History for Secondary Schools and Colleges: The Colonial and Post-Colonial Periods*. Vol. 2. London: Macmillan.
- Gage, W. and S. Ohonessian, 1974, 1977, “ESOL Enrolments throughout the World”, *Linguistic Reporter* (1974) and later reproduced in *English Teaching Forum* (1977).
- Graddol, D., 1997, *The Future of English? The UK: The British Council*.
- Gyasi, I. K., 1990, “The State of English in Ghana”, *English Today* 6 (3), 24-26.
- Hocking, B. D. W., 1974, *All What I was Taught and other Mistakes: A Handbook of Common Mistakes in English*. Nairobi: Oxford U.P.
- Kachru, B. B., 1982, “South Asian English”. In Bailey, G. and M. Görlach (eds.), *English as a World Language*. Ann Arbor: University of Michigan Press.

- , 1985, “The Pragmatics of Non-Native Englishes”. In Smith, L. E. (ed.), *English for Cross-Cultural Communication*. Wiltshire: Macmillan, 15-39.
- , 1991, “Liberation Linguistics and the Quirk Concerns”, *English Today* 25, 7 (1), 3-13.
- (ed.), 1992, *The Other Tongue: English across Cultures*. Urbana: U. of Illinois Press.
- , 1996, “The Paradigms of Marginality”, *World Englishes* 15 (3), 241-55.
- Koenig, E. L., 1983, “Sociolinguistic Profile of the Urban Centres”. In Koenig et al. (eds.), 33-53.
- , et al. (eds.), 1983, *A Sociolinguistic Profile of Urban Centres in Cameroon*. U.S.A: Crossroads Press.
- Leith, D., 1996, “English - Colonial to Postcolonial”. In Graddol, D. et al. (eds.), *English: History, Diversity and Change*. London: Routledge.
- Mazrui, A. M. and A. A. Mazrui, 1996, “A Tale of Two Englishes: The Imperial Language in Post-colonial Kenya and Uganda”. In Fishman, J. A. et al. (eds.) *Post-Imperial English: Status Change in Former British Colonies, 1940-1990*. Berlin: Mouton de Gruyter, 271-302.
- Mbangwana, P. N., 1983, “The Scope and Role of Pidgin English in Cameroon”. In Koenig et al. (eds.), 79-91.
- , 2002, *English Patterns of Usage and Meaning*. Yaounde: Presses Universitaires de Yaoundé.
- Mbassi-Manga, F., 1976, “The State of Contemporary English in Cameroon”. In *Cameroon Studies in English and French (CASEF)*. Victoria: Presbook.
- Menang, T., 1979, *Trends in Cameroon Pidgin English Lexicology*. Unpublished Master’s Thesis, University of Yaounde.
- Momo, T. N., 1997, *Le Yemba: Histoire de la langue écrite dans la Menoua*. Dschang, Cameroon: Comité d’Etude de la Langue Yémba. www.ldc.upenn.edu/sb/papers/momo.pdf.
- Mufwene, S. S., 1997, “Native Speaker, Proficient Speaker, and Norm”. In Singh, R. (ed.), *Native Speaker: Multilingual Perspectives*. New Delhi: Sage, 111-23.
- , 2001, *The Ecology of Language Evolution*. Cambridge: Cambridge U.P.
- Ogu, J., 1986, “Variations in the Use of English in Fiction: A Look at Four Nigerian Novelists”. In S. Unoh (ed.), *Use of English in Communication: The Nigerian Experience*. Ibadan, Nigeria: Spectrum Books Ltd, 79-96.
- Prator, C., 1968, “The British Heresy in TESL”. In Fishman, J. A. et al (eds.), *Language Problems of Developing Nations*. New York: John Wiley & Sons, 459-76.
- Quirk, R., 1985, “The English Language in a Global Context”. In Quirk and Widdowson (eds.), 1-6.
- , 1988, “The Question of Standards in the International Use of English”. In Lowenberg, P. (ed.), *Language Spread and Language Policy*. Washington DC: Georgetown U.P., 229-41.
- , 1990, “Language Varieties and Standard Language”, *English Today* 21, 3-10.
- , and H. G. Widdowson (eds.), 1985, *English in the World*. Cambridge: Cambridge U.P.
- Schneider, E., 2000, “Feature Difusion vs Contact Effect in the Evolution of the New Englishes: A Typological Case Study of Negation Patterns”, *English World-Wide* 21 (2), 201-30.
- Schneider, G., 1974, *Masa Troki Tok Sey*. Athens, Ohio: Ohio University.

- Singh, R. et al., 1995, "On 'new/non-native' Englishes: A Quartet", *Journal of Pragmatics* 24 (3), 283-94.
- Tchoungui, G., 1983, "Focus on Official Bilingualism in Cameroon: Its Relationship to Education". In Koenig et al. (eds.), 93-115.
- Trudgill, P., 1998, "Standard English: What it isn't", *The European English Messenger* VII (1), 35-39.
- , and J. Hannah, 1994, *International English: A Guide to Varieties of Standard English*. 2nd ed. London: Edward Arnold.
- United Nations Population Division, 2002, *World Population Prospects: The 2002 Revision Population Database*. <http://esa.un.org/unpp/p2k0data.asp>.
- United Nations Environmental Programme (UNEP). *Africa Population Database*. <http://www.na.unep.net/globalpop/africalapp-2.php3>
- Whitworth, G. C., 1907, *Indian English: An Examination of the Errors of Idioms made by Indians in Writing English*. Letchworth.
- Yule, G., 1996, *The Study of Language*. 2nd ed. Cambridge: Cambridge U.P.

MULTILINGUAL BACKGROUNDS AND THE IDENTITY ISSUE IN CAMEROON

Eric A. Anchimbe

Ludwig-Maximilians (University of Munich)

Abstract

This paper seeks to establish at what point the language issue is relevant to the establishment of a Cameroonian identity which is devoid of complications (difficult though it may be) and which reflects the ideal citizen in this multilingual setting. It traces some of the historically linked sociolinguistic problems encountered in Cameroon to the quest for an adequately recognised identity vehiculated in a given language. This brings to light the divergent attachment to the official languages and the native Cameroonian languages and the various strata of identity they engender.

1. Overview

The multilingual nature of Cameroon (although officially termed and treated as bilingual in French and English) and its corresponding multiethnic complexity (with over two hundred identified ethnic groups) fuel the constant confusion that the ideal Cameroonian goes through as to where to lay the pledge of his identity. Of course in the building of a given identity, the linguistic component is of vital importance because as Giles and Coupland (1991: 107) postulate culture, for instance, “can be viewed as derived from, if not constituted in, communication and language practices”. On such a platform, therefore, where should the Cameroonian identify himself/herself? With the foreign English or French languages? With his/her native language (which number up to about 285) or with both? These questions are of great importance to the understanding of the linguistic behaviour and the subtle linguistic discords witnessed in the country since its independence.

2. Ethno-linguistic Atmosphere of Cameroon

From a purely historical perspective, and for more empirically based results to be achieved, studies of the linguistic complexity of Cameroon (or of any African state)

must be founded on the long and diverse itineraries of her history. Makoni and Meinhof (2003: 1) sum this up thus,

Pre-colonial migration, trade down the colonies, the radical displacements of slavery, the growth of print literacy and the decline of oral culture, arbitrary territorial changes under colonialism, industrial exploitation of natural resources, and the unprecedented rapidity of migration and urbanisation in the postcolonial period have brought language groups into contact and conflict, changing social and economic life and with it the shape, function and status of the languages within specific communities.

Cameroon is generally referred to as *Africa in miniature*, given that three of the four language families in Africa are represented within her boundaries (Greenberg 1966). These are the Congo-Kordofanian, Nilo-Saharan and Afro-Asiatic. The structures of the fourth, that is, the Khoisian have not been found in Cameroon. Besides just the resettlement of African peoples on Cameroonian soil, further outcomes of the historical admixture include the spread of non-African languages, the ramification of identity and the sowing of several discontentment seeds which today flourish in most parts of the continent. A complex foundation of multilingualism was therefore laid over a longer historical period than can be simply accounted for by colonialism and post-colonialism. It is not the intention to probe into this here (see Mufwene 2001, Makoni and Meinhof 2003, etc.) but to situate at what point the multilingual background contributes tremendously to the quest for a befitting identity.

2.1. Pre-colonial Period

Although the exploitation colonialism mechanism was built on the false premise that Africa, the black continent, was like a blank sheet of paper on which European civilisation had to be written, it is far from believing that pre-colonial Africa lacked a system of its own. This system, if not spread out on the margins of European-styled nation state, was in itself complete at the borders of tribal and cultural similarities bound together especially by shared languages. The French *mission civilisatrice* which “aimed at assimilating or absorbing France’s colonial subjects to the point where they would actually be Frenchmen linguistically, culturally, politically and legally” (Fanso 1989: 65) was based exclusively on this rather diminutive and prejudiced judgement of Africa.

Most studies of language evolution in Africa centre on the colonial and post-colonial periods. The multiplicity of languages, tribes and cultures is evaluated generally from the clash of the west and the south confounded by colonialism. However, other aspects like inter-tribal marriages, wars and struggle for supremacy, trade, quest for land and less hostile territories created inevitable contexts for multilingualism. As explained by Nurse and Spear (1985), Nurse (1997), Mazrui and Mazrui (1998), Mufwene (2001) and a handful of others, pre-colonial factors provide a more comfortable account for the origin of, for instance, coastal Swahili as being neither Arabic-based nor a Creole. Having been used over a long period of time as a vernacular or lingua franca by non-Semitic Africans, Mazrui and Mazrui (1998)

conclude that Arabic can as well be considered an indigenous language. Again Mufwene (2001: 169) makes it clear of North Africa that “Arabic ethnicity has depended more on assimilation to Islam and usage of Arabic as a vernacular than on race”. What stands out in these examples is that language is central to the delimitation of boundaries. It further indicates the fluctuations in identity that existed even before colonialism.

Furthermore, the “African confusion of tongues” (Fonlon 1969: 9) extensively represented in Cameroon was certainly realised if not initiated during this period. Today Cameroon counts over 200 identified ethnic groups who use a yet to be officially decreed number of native languages. However, *Ethnologue* purports a total of 285 languages of which 4 are extinct, 11 are threatened by extinction and 270 are living. This is a notoriously complicated infrastructure first to building a nation and secondly to achieving a binding identity for all, particularly because “it has become absolutely impossible to achieve, through an African language, that oneness of thought and feeling and will that is the heart’s core and the soul of a nation” (Fonlon 1969: 9).

2.2. Colonialism and the advent of Independence

The Berlin conference of 1884 confirmed the quest for territories and expansion of European boundaries beyond Europe. Cameroon was then authenticated as a German colony (1884-1916) until the defeat of Germany in the First World War. Evidence of the impact of the German language in Cameroon is less, given that the Germans were more interested in exploitation schemes than educational projects, which became a central focus in the later colonial expeditions of the British and the French; they were confronted by a pidgin that was in extensive use —bequeathed by the Portuguese trade colonial period along the West African coastline; and lastly they had just thirty years (1884-1916) to set up a German language policy which entailed a complete breakdown of the pidgin established over several centuries of trade (1495-1884). The subsequent colonial patterns of Britain and France, aided by the work of missionary churches and made more decisive by reconstruction efforts and consolidation of territories following the war, had a stronger focus on expanding not only the empire but expanding a feel of it through the colonial languages (see Wardhaugh 1987, Echu 2003). So the one-fifth of Cameroon administered by British colonial authority from Lagos, Nigeria now had English as its official language and the other four-fifth administered by resident French governors, used French in all activities, including education, for as Fonlon (1969: 20) advances “it was considered essential that instruction in the other subjects should be in French almost from the first day in school”.

English and French were able to gain considerable strength since Cameroon was first administered by Britain and France under the League of Nations Mandate (1919-1945) and later under the United Nations Trusteeship (1945-1960). They were the official languages in each of these regions. Even after January 1st, 1960 when French Cameroon became independent, followed a year later by the independence of British Cameroon these languages maintained the status of official languages. Although the British tried to give the native languages a chance in edu-

cation, this was met with strong resistance from the indigenes. In the South West area, the Bakweri people rejected the Duala language on grounds that “it is quite against the reason that our children should be educated in a barbarous tongue instead of a civilised one like German or English” (see Mbassi-Manga 1973: 39). When British Cameroon voted in the February 11th, 1961 referendum to unite with French Cameroon under a Federal Republic that was later transformed by Presidential Decree No. 27-270 into a unitary state on May 20th, 1972, French and English continued to serve as official languages over the complex number of native languages.

Of course, the language policy adopted at independence was the obligatory outcome of colonial presence, since Cameroon “inherited ... French and English; and has therefore been obliged to become, constitutionally, a bilingual State” (Fonlon 1969: 9). The adoption of state bilingualism in English and French was meant to create a unique national identity for Cameroon. As the then president Ahmadou Ahidjo (1964) explained,

[W]e must in fact refrain from any blind and narrow nationalism and avoid any complex when absorbing the learning of other countries. When we consider the English language and culture and the French language and culture, we must regard them not as the property of such and such a race, but as an acquirement of the universal civilisation to which we belong. This is in fact why we have followed the path of bilingualism since ... it offers us the means to develop this new culture ... and which could transform our country into the catalyst of African unity.

The adoption of bilingualism implied the creation of a completely new belonging and identity, one that would transform the country into the centrepiece of Africa and her entire existence. For Fonlon (1969: 35) “the target to aim at, for us, should be, not merely State bilingualism, but *individual bilingualism*: that every child that passes through our education system shall be able to speak and write both English and French”. The bilingualism project, though highly criticised for its lack of decisiveness (Tchoungui 1983, Kouega 1999, Anchimbe 2003), and its general failure since “*de jure*, Cameroon has become a bilingual state; but, *de facto*, it is a highly diversified multilingual, multi-cultural country” (Fonlon 1969: 28), succeeded in adding another dimension to the multilingual nature of Cameroonians and their quest for a befitting identity.

2.3. Language contact and contact varieties

The languages in Cameroon came into contact not only with English and French but also with one another. Although it cannot be claimed that the regional lingua francas are the outcome of contact, it can at least be partially upheld that the major lingua franca, that is, Cameroon Pidgin English (CPE) received more contextualisation through its contact with Portuguese (1495-1800), Educated English, Cameroonian languages and French. From a word count reported by Mbassi-Manga (1976) and Mbangwana (1983) the total lexicon of CPE is 80% English, 14% Indigenous languages, 5% French and 0.07% from other sources (among which are

Portuguese, creations and innovations, etc.). The regional lingua francas which Breton and Fohitung (1991: 20) identify to be Fulfulde, Ewondo, Basa'a, Duala, Hausa, Wandala, Kanuri, and Arab Choa, cut across several native language boundaries. They can be located within given regions and therefore limitable to a given group of speakers with shared characteristics or identities. CPE on the other hand is a non-man's code that is predominantly attached to the uneducated and the less privileged of the society. This negative appraisal stems from the fact that CPE is a non-literate code; it is not taught in school like English and French. It has therefore become the target for accusations of fallen and falling standards in English and education in general. Although no one wants to identify with it, everyone seems to use it. It is an important dimension in identity concealment given that its users cannot be traced to given regions, like English for Anglophone provinces, French for francophone regions, the native languages for their respective tribes and so forth.

3. Multilingual Backgrounds and the Identity Issue

With the current sociolinguistic and political state of the country set up at the close of colonialism which makes a primary distinction to anglophone and francophone parts, the attachment to these languages as icons of linguistic identity was made prominent. It became so close to another ethnicity as observed by Wolf (2001: 223); "the feeling of unity is so strong that 'being Anglophone' denotes a new ethnicity, transcending older ethnic ties". This rather linguistically motivated and sometimes derogatory distinction as in such slang as *anglofoul anglofool* for anglophone and *frogs/francofoul/francofool* for francophone, which is "tantamount to group definition and membership is too strong that it excludes non-group members, and transcends ethnic contours" (Anchimbe 2003: 3). But because the role of native or tribal communion and cultural integrity is still strongly linked to language, the native languages are far from completely subsuming themselves to the hegemony of the official languages which are understood to be icons of political identification, minority-majority classification and which remind the Anglophone of marginalisation.

The Anglophone-Francophone divide fashioned on the commonality of English and French is far more profound than just the use of these languages. Although grossly multi-cultural and having internal differences and discontents, this divide has hatched new identity icons which Anchimbe (2003) refers to as Anglophonism and Francophonism. At a higher level, that is beyond the communality of the tribe and the village, these icons "constitute superior sociolinguistic groupings above the individual ethnic groupings, whose languages are less represented in education and less useful in cross-ethnic communication" (Anchimbe (2004: 3). On the other hand, the native languages honourably referred to as national languages, carry the special attribute of one's roots and origin. No one can claim in this case not to have a native language. Such a person would be treated as a vagabond, one who has nowhere to retreat to or to identify with when the higher icons, anglophonism and francophonism become less representative. This is basically the case with those referred to as the 11th province (we return to it later). Here bilinguals in French and

English belong yet to another class. One that is more flexible, given that its members can switch, on basis of performance, from English to French, as contexts require. However, the multilingual backgrounds have created identity confusion and forged a reshape of belonging and attachment at various levels, some of which are examined below.

3.1. Where do I belong?

The claim to an identity is incomplete if it has no language through which it is vehiculated. Jaffe (1996: 818) commenting on what he calls the “European political ideology of language” upholds that “linguistic identity is a prerequisite for cultural identity and political stability”. Moreover, culture, if it has to be vocal and immediately comprehensive, must “be viewed as derived from, if not constituted in, communication and language practices” (Giles and Coupland 1991: 107). The discontent calls issued so far in Cameroon are attached to a shared cultural background that has a strong linguistic icon. It must be remembered that the Soweto Massacre of 1976, a historical landmark in South African history, was a response to the refusal of black school children to use Afrikaans as medium of instruction. The imposition of Afrikaans by the apartheid regime was interpreted as a systematic attempt to erase their rights and belonging to a more prestigious language, English and limiting them to a secluded range of activities.

If language is thus central to identity, it is then clear that many identities can be traced in Cameroon. These may not be easily traced to groups of individuals that can be pinned down by such sociological factors as regions of origin, gender, profession, age, level of education and so forth. Rather the same individual may incorporate various identities depending on the context in which he/she is. Four of such identities can be easily encountered. They include the official language identity, the ethnic identity, the bilingual identity and the individual identity.

3.1.1. *Official Language Identity*

Being more expansive levels of identification, and given that the Anglophone provinces are sometimes treated as one region, the official languages serve as basis for the procurement of a national feeling. This is significantly because Cameroon is known as a bilingual country. A national identity can only therefore be reached through one of its official languages. It is construed on the English-French or Anglophone-Francophone categorisation. It is less attached to education and other education-related activities because geophysical regionalism applies very strongly. An Anglophone in this case is one born in either the North West Province or the South West province. With the political tension that erupted over a decade ago as to the equality of anglophones and francophones in the country, the official language identity has had a strong political dimension. Political ideas are spread in these languages because they transcend the borders of the native languages.

Unfortunately, the gross differences that exist between tribes render the official language identity rather weak and overgeneralised. It is too general and less decisive,

it provokes suspicion since ethnic groups are often caught up in subtle competition. Competition and the desire to maintain a worthwhile esteem make ethnic ties far stronger. The ethnic group, contrary to the official language classification, provides a closer set-up made more comfortable by the native language, which very often is not understood by non-tribe members. English and French on the other hand are too widespread and lack the reserve of secrecy and exclusion enjoyed by the native languages.

3.1.2. *Ethnic Identity*

As mentioned earlier, it is often identified as the roots. It is closed to others who do not belong to it. Ethnic groups are generally in a subtle competition either for social esteem or for political achievement. This breeds insecurity and victimisation of less politically backed groups by those that wield power. The 1992 post-presidential election violence (fuelled by accusations of rigging) shifted from political to ethnic violence and victimisation. Settlers from the anglophone and Western provinces in especially the Southern province (home of the incumbent president) were earmarked to be chased and their property, especially businesses, was looted and destroyed. This therefore indicates that ethnic ties often drown political ideologies since the ethnic language carries far more than a simple message.

3.1.3. *Bilingual Identity*

Fonlon's (1969) notion of state bilingualism required that all educated Cameroonians must be fluent in both English and French. President Ahidjo (1964) certainly had this in mind when he declared that bilingualism "could transform our country into the catalyst of African unity". Unfortunately as it is often said, it is Cameroon that is bilingual and not Cameroonians. Nevertheless, a few people are. These include graduates of the Advanced School of Translation and Interpretation (ASTI) Buea, Anglophones who grew up and studied in francophone areas and vice versa, as well as others who majored in bilingual studies at the university. Although this group may be too complicated to accurately circumscribe following the path of bilingualism, it is an important arm of success in certain domains. However, in terms of identity creation, it is often opportunistic and determined by contexts and situations. They have the unique chance of benefiting from francophone and Anglophone opportunities, if the regional criterion is kept out.

3.1.4. *Individual identity*

This is less significant and less uniform given that individuals try to give themselves befitting esteems. These are usually socially well-ranked people in the society; those exposed to foreign-influence; and those trying to live above the limit of their class, that is, "apes of their betters" (Passé 1947: 33, about Sri Lanka). As a result of this, they are often treated as showing off. As shown on the following figure, it is the smallest identity group. It has less political strength but does often enjoy political favours given that these are generally wealthy people.

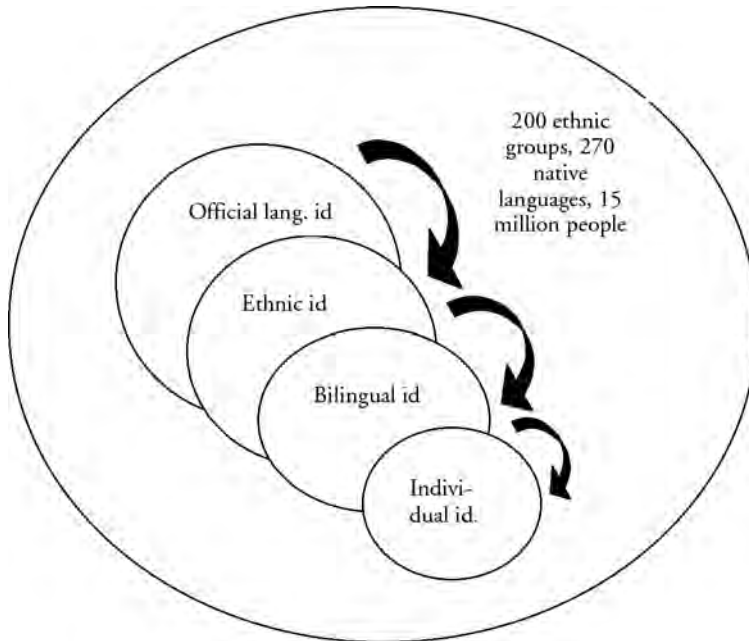


FIG. 1. Identity belonging and overlapping in Cameroon

These identities, following the diagram, open into each other. They are not limited to identifiable and exclusive groups of people. Individuals portray a cross-section of these identities. As mentioned earlier, this simply means that these identities can be accumulated. Everyone presumably belongs to an ethnic group (ethnic identity), is either an Anglophone or francophone with the exemption of those referred to as “11th province” (official language identity), may or may not be bilingual (bilingual identity) and may want to be regarded as socially superior (individual identity). However, this is not as smooth as it seems to be. These identities have turned out to be makeshift responses to several social discontent calls. Shifting from one to the other is a means of coping with the challenges presented by the reunion of people of diverse origins and races within the same country.

4. Sociolinguistic Calls for Identity

Cameroonians seem to be caught in an identity web from which various kinds of calls based on social discontentment have been issued, for example: the reference to the 11th province (since they are neither Anglophone nor francophone), the Anglophone problem (since they are marginalised by the populous francophones) and the Southern Cameroons secession issue (since the political union at independence has failed to guarantee equity for Anglophones). The only significant similarity between these sociolinguistic disputes is the search for a better recognised identity.

4.1. Bases of Discontentment

In the face of the forceful national boundaries created by colonialism and protected today by United Nations Charters; the seemingly heavy distant-presence of the (ex)colonial powers in certain nations and with the continuous expansion of (ex)colonial languages, many cries of discontentment have been issued in post-independence Cameroon. The discontentment has forged a general torrent of insecurity that has led to the construction of identity borders based predominantly on the commonality of language. Although mobility in these identities is common, they serve given purposes each time they are adopted.

The lack of equal opportunities in political appointments and marginalisation in other aspects of national life are projected by the Anglophones as a major source of their dissatisfaction. This state-of-affairs has launched them into the defence of a culture presumed to be their own, that is, the Anglo-Saxon tradition, and into the call for the creation of an independent state along the boundaries set up by British trusteeship up to 1961.

The quest for lasting alliances and the search for security in the face of the divide-and-rule system adopted by the regime in power, as well as the robust pride and political strength of the francophones and the tribes that wield political power and enjoy political favours, have been central to what people say they are. This heightened attachment to the official languages and with it feelings of anglophonism and francophonism; strengthened tribal ties and tremendously fuelled efforts for the secession of the former British colony, Southern Cameroons. Discontentment therefore varies from region to region and from one group or tribe to the other. We will look at three of them.

4.2. The Anglophone Movement

It is interesting to note how tremendous the impact of British colonialism (1919-1961) is still psychologically present in Cameroon. The part of the country colonised by Britain is generally said to have an Anglo-Saxon tradition. This was crowned in 1991 by the creation of an Anglo-Saxon styled state university in Buea. Constituting basically two out of the ten provinces and using English instead of French as in the rest of the country, the Anglophones have issued many reprimands in what is generally referred to as "the Anglophone Problem". Over the years much blame has been directed at the Anglophone elite who sat in the Fumban Conference (1961) and signed for union with East (francophone) Cameroon. Although several reasons have been advanced for the return to federation, for instance the treatment of the Anglophone in unequal standing with the francophone, the major source of these calls is that the Anglophones are united under the umbrella of the English language, (inherited from British colonialism) which to them is their identity-marker (see The Buea Declaration 1993). Several movements and associations cropped up during the 80s and 90s to fight for the rights of the Anglophones among them: the Cameroon Anglophone Movement (CAM) now called Southern Cameroon's Restoration Movement (SCARM), the Ambazonia Republic, the Free West Cameroon Movement (FWCM) and the Southern Cameroon's National Congress (SCNC).

In 1993 the First All Anglophones Conference (AACI) was convened in Buea under the banner of CAM to give a real look into their problems and decide on their future. It came up with the Buea Declaration (1993), which summed up the grievances and proposed some urgent remedial. Prominent among these grievances was the inequality in the use of the official languages in national life. The national television was singled out as a case in point. As the Declaration upholds,

Television films and programmes originally made in English are shown in Cameroon only after they have been translated into French, and only in their French version. Broadcast time on Radio and television is very unevenly divided between English and French programmes, even though it does not take longer to inform, educate or entertain in French than it does in English.

The Buea Declaration was a subtle fact-giving manifesto. It was followed a year later by the Bamenda Declaration (1994) which resolved that if after a “reasonable time” there was no reaction from the government, then Southern Cameroons would proceed to declare her independence from La République du Cameroun (see 4.4).

Again, the reference to a linguistic identity is central. No one seems to call for the inclusion of native language programmes on the national television. No one seems to bother no native language programme exists on the national television although these languages are officially called national languages. This further confirms Wolf’s (2001) position that being an Anglophone or francophone (official language identity) constitutes higher ethnic entities. The Buea Declaration further says, “in the end, Anglophones who share equally in the burden of financing Cameroon Radio and Television get far less than 1/4 of the service provided by this public utility”. To be able to judge this stance, let us have a look at some excerpts of the national television (Cameroon Radio and Television Corporation, CRTV) programme. Table 1 below indicates the sluts and the languages used.

The above table reports the vital parts of the national television programme for three days (Mondays, Tuesdays and Wednesdays: see Appendix I for the rest of the week) for the period August-September 2002. It is not altogether complete because it is often drawn in advance and impromptu features are inserted as need arises. However this is the basic skeleton found on the CRTV website (www.crtv.cm).

Within these three days, a total of 1.046 minutes of airtime is spent on the diverse programmes. French programmes occupy 541 minutes (51.7%), English 300 minutes (28.7%), bilingual (English-French) 70 minutes (6.7%) and neutral (images or music) occupy the remaining 135 minutes (12.9%) (see Appendix II below). More than 50% of the airtime goes to French language programmes and only 28.7% to English programmes. Although a bilingual country only 6.7% of the 1.046 minutes is dedicated to bilingual programmes. The Anglophones consider the paucity of English language programmes as a neglect not only of that part of the diversity but also predominantly of them. Their tastes are not taken into account. The attachment to English makes it possible for them to consider its prominence as representative of their own existence.

TABLE 1. Programme schedule of CRTV television

Day	Time	Duration	Programme	Language
Monday	13:55-14:00	5mn	Overture d'antenne	Image/sound
	14:00-14:10	10mn	Flash Bilingue	Bilingual
	15:10-15:40	30mn	Clip Box	Music
	17:02-18:00	58mn	Fou Fou Foot	French
	19:30-20:00	30mn	The 7:30 News	English
	20:00-20:30	30mn	Terra Nostra	French
	20:30-21:00	30mn	Le Journal	French
	21:18-21:44	26mn	Déviations	French
	22:00-00:00	120mn	The Monday Show	English
Tuesday	13:55-14:00	5mn	Overture d'antenne	Image/sound
	14:00-14:30	30mn	Au Coeur de la société: Nkongsamba	French
	15:10-15:40	30mn	Clip Box	Music
	17:00-17:57	57mn	Sports Parade	French
	18:30-19:00	30mn	You and the Law	English
	19:30-20:00	30mn	The 7:30 News	English
	20:00-20:30	30mn	Terra Nostra	French
	20:30-21:00	30mn	Le Journal	French
	23:00-23:30	30mn	CRTV Late night News	Bilingual
	23:35-00:05	30mn	Méli-Mélo	French
Wednesday	13:55-14:00	5mn	Overture d'antenne	Image/sound
	14:00-14:30	30mn	Groove	French
	14:30-15:00	30mn	Clip Box	Music
	16:05-16:32	27mn	Déviations	French
	16:32-17:32	60mn	Women and Development	English
	18:00-18:30	30mn	Le quotidien des provinces	Bilingual
	18:30-19:00	30mn	Santé magazine	French
	19:30-20:00	30mn	The 7:30 News	English
	20:00-20:30	30mn	Terra Nostra	French
	20:30-21:00	30mn	Le Journal	French
	21:17-21:30	13mn	Q.D.O	French
	21:30-22:30	60mn	Un Siècle d'histoire	French
	23:00-23:30	30mn	CRTV Late Night News	Bilingual
Total		1046mn		

4.3. The Eleventh Province

The search for a source of living prompted many displacements in Cameroon during and after colonialism. Among these is the migration of North Westerners to the South West as labour force for the Cameroon Development Corporation (CDC), the settlement of people from the Western province around Kumba for business reasons given that Kumba is a major business spot on the border with Nigeria, among others. These migrations were initially meant to be temporary but have ended up being permanent. These patterns of settlement have resulted in a

number of conflicts. Prominent among them is the politically induced “come-no-go” issue of 1998 and the 11th province reference. The “come-no-go” refers to North Westerners who moved to the South West either voluntarily or forcefully (under the German colonial rule) to work in the CDC but have ever since replaced one generation with another. Having lived so long in the area the immigrants who wanted to join politics were rejected on grounds that they were foreigners.

On the other hand, the 11th province refers to those from the Western province who by virtue of their origin are considered francophones but who have lived, worked all their lives and have been delivered children in the South West province, an Anglophone zone. So they are francophones by origin but Anglophone by upbringing and linguistic expression. That is why they are referred to as an 11th province having its own qualities different from the other ten. Cameroon has ten provinces that are generally categorised on English (anglophone) and French (francophone)-speaking basis. This group is extensively bilingual and does not squarely fit into any of these linguistic categorisations. They are therefore considered to belong to a different province—the 11th. Some people have even gone as far as asking for the creation of a province for them. Again, the linguistic background is used here as the basis for determining belonging. Children born into this situation have had difficulties returning to the francophone area, where they are treated as outsiders and integrating in the Anglophone region in which they were born, where they are also considered outsiders. Although they may own property in the South West province, they are not treated as part of the tribal heritage and so have no claim to it. Most of them do not even speak French, but since their parents or grand parents came from a French-speaking zone, they are considered francophones.

4.4. The Southern Cameroons Secession Attempt

After several years of verbal requests for the respect of the rights of the Anglophone, issued at the First All Anglophones Conference in Buea, re-echoed at the Second All Anglophones Conference (AACII, 1994) in Bamenda, the SCNC and other group favouring the secession of former British Southern Cameroons, on October 1st, 1999 declared the autonomy of the Republic of Southern Cameroon (also referred to as Ambazonia) and followed this declaration with the hoisting of her flag. The flag is modelled on the defunct Federal Republic of Cameroon flag with two stars. For the Ambazonia Liberation Party (ALIP) whose aim is to “fight and end this bitter episode as a conquered people and restore our sovereignty and independence” (Party Manifesto 2004), secession is the only solution given that the case made at the United Nations is not forthcoming and the Cameroonian government is far from accepting any such secession. The major claim is once more related to the linguistic colonial inheritance. According to ALIP,

For the powerful of Cameroon and the government of Cameroon, to come from Ambazonia or to look Ambazonian, is reason for discrimination, scorn, shame, suspicion and hatred.

Preamble of the Manifesto of Ambazonia Liberation Party
(ALIP)

Again, Ambazonians (as they are called) cannot be identified by any other parameter than the use of English. The official language identity therefore adopts a political and a regional dimension that makes it to spread over several other identities. The Anglophones, it should be mentioned, are very culturally diversified and would not have any claims to similarity if English had to be taken away from them. This confirms the initial claim in this paper that linguistic background or expression is very essential to the creation of an identity.

5. Ways Out

This paper has attempted to situate at what point linguistic background moulds identity(ies). It also traced historically the place and vitality of the linguistic component in some of the social and political discontentment common in Cameroon. From the subtle complaints of the Anglophones to the declaration of independence of the Federal Republic of Southern Cameroon, one thing is made quite clear. That is, the quest for a modest and worthy identity is intricately linked to one's language and the strong attachment to it. This is because all of the other characteristics that make the group a common entity must be transmitted through a language. And since many languages exist side-by-side one another, it becomes possible for individuals to expose many different identities.

The situation as presented above is not a dead-end as such, nor has it deteriorated into the quagmire of violence either. This indicates that if the sovereign nation has to continue intact several concessions and reforms have to be made. The first of them is levelling political mounds that have been raised by ethnicity and regionalism and thereafter giving regional representation a real attention. This can certainly make sense if a national identity built preferably on Fonlon's (1969) vision of official bilingualism is sought and implemented. This would mean promoting bilingualism beyond the threshold of instrumental necessity and creating an integrative dimension for it. In the past bilingualism was the goal of Anglophones wanting to integrate francophone zones, francophones seeking to go abroad, francophones seeking the Commonwealth scholarship, Anglophones seeking jobs in francophone zones and so forth. The encouragement of an integrative attachment to bilingualism in English and French would reduce if not erase the geophysical and psychological boundaries of francophonism and anglophonism. In this vein, the minority-majority, oppressor-oppressed gap hitherto created by this divide would definitely disappear.

6. Conclusion

The above recommendations cannot be achieved overnight. It requires much more than a single generation to implement a bilingualism scheme that would ensure that everyone who successfully completes secondary school should normally be fluent in French and English. This is however possible. If the generation of bilinguals in nursery and primary schools in urban centres (see for instance Anchimbe's

(2003) study of Yaounde) were to be succeeded by yet two others, then the issue would be near a definitive solution. This is because these children, born into mixed anglo-francophone families or entirely francophone families are considered neither as francophones nor as Anglophones. They are enrolled in an English-medium school; study exclusively in English (with French as a subject) and only speak French at home with their parents and/or neighbours. As indicated in Anchimbe's (2003) survey, 54.6% of the 194 parents would consider their children Anglophones given they will have studied all along in English. But in a follow up question, 75.2% objected to considering these children in the same manner as Anglophones from the English-speaking zones. The children consider themselves as bilinguals and nothing more. They lay no claims, either geophysical (origin) or linguistic, to the origins and linguistic (identity) belonging of their parents, either of whom may be from one of these classifications. If this annihilation of regional and strict official language division is attained, then a stable identity would be achieved, more convenient alliances would be born, and lastly, ethnicity would be an added spice rather than a deadly dessert.

References

- Adams, B., 2004, "Ambazonia Political Party formed". In *The Post News Online*, <http://www.postnewsonline.com/2004/07/strongambazonia.html>.
- Ahidjo, A., 1964, "Message to the Nation, Buea, 1st October (extract)", *Abbia* 7 (4).
- All Anglophones Conference, 1993, *The Buea Declaration*, 2nd and 3rd April 1993. Limbe: Nooremac Press.
- Amvela, Z., 2001, "The English Language in Cameroon: The German Period", *Epasa Moto* 1, (4), 194-225.
- Anchimbe, E. A., 2003, "Anglophonism and Francophonism: The Stakes of (Official) Language Identity in Cameroon". Paper presented at SignPost Forum (A postgraduate literary forum of the Dept. of English, the University of Yaounde I) on 11th June 2003.
- , 2004, "Lexical Markers of Social (Youth) Group Communication in Cameroon". In *Proceedings of 4th Postgraduate Forum on Linguistics*, Hong Kong University, <http://www.hku.hk/linguist/conf/prf/4prf-proceedings.pdf>
- Breton, R. and B. Fohtung, 1991, *Atlas Administratif des Langues Nationales du Cameroun*. Yaounde/Paris: Cerdotola/CREA - ACCT.
- Dieu, M. and P. Renaud, 1983, *L'Atlas Linguistique du Cameroun*. Paris, Cerdotola and ACCT.
- Echu, G., 2003, "Coping with Multilingualism: Trends in the Evolution of Language Policy in Cameroon". In *PhiN* 2003 (31), <http://www.fu-berlin.de/Phin/x/dir.cgi?ptest>. *Ethnologue: Languages of the World*. www.ethnologue.com
- Fanso, V., 1989, *Cameroon History for Secondary Schools and Colleges: The Colonial and Post-Colonial Periods*. Vol. 2. London: Macmillan.
- Fonlon, B., 1969, "The language Problem in Cameroon", *Abbia* 22, 5-40.
- Giles, H. and N. Coupland, 1991, *Language: Contexts and Consequences*. Minton Keynes: Open U.P.
- Greenberg, J., 1966, *Languages of Africa*. The Hague: Mouton.

Jaffe, A., 1996, "The Second Annual Corsican Spelling Contest: Orthography and Ideology", *American Ethnologist* 23, 816-835.

Koenig, E. et al, (eds.), 1983, *A Sociolinguistic Profile of Urban Centres in Cameroon*. U.S.A: Crossroads Press.

Kouega, J., 1999, "Forty Years of Official Bilingualism in Cameroon: An Appraisal", *English Today* 15 (4), 38-43.

Makoni, S. and U. H. Meinhof, 2003, "Introducing Applied Linguistics in Africa", *AILA Review* 16, 1-12.

Mazrui, A. M. and A. A. Mazrui, 1998, *The Power of Babel: Language in the African Experience*. Oxford: James Currey.

Mbangwana, P., 1983, "The scope and role of Pidgin English in Cameroon". In Koenig et al. (eds.), 79-91.

Mbassi-Manga, F., 1976, "The State of Contemporary English in Cameroon". In *Cameroon Studies in English and French (CASEF)*. Victoria: Presbook.

Mufwene, S., 2001, *The Ecology of Language Evolution*. Cambridge: Cambridge U.P.

Neba, A., 1982, *Modern Geography of the Republic of Cameroon*. 2nd. Ed. New Jersey: Neba Publishers.

Nurse, D., 1997, "Prior Pidginization and Creolization in Swahili". In Thomason, S. (ed.) *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins, 271-363.

—, and T. Spear, 1985, *The Swahili: Reconstructing the History of an African Society 800-1500*. Philadelphia: University of Pennsylvania Press.

Passé, H., 1947, *The English Language in Ceylan*. Unpublished Doctoral Dissertation, University of London.

Tchoungui, G., 1983, "Focus on Official Bilingualism in Cameroon: Its Relationship to Education". In Koenig et al. (eds.), 93-115.

The Cameroon Radio and Television website, www.crtv.cm.

The Southern Cameroons Website, <http://www.southerncameroons.org>

Wardhaugh, R., 1987, *Languages in Competition*. Oxford: Basil Blackwell.

Wolf, H. G., 2001, *English in Cameroon*. Berlin: Mouton de Gruyter.

Appendix I. CRTV programme of the rest of the week

Day	Time	Duration	Programme	Language
Thursday	13:55-14:00	5mn	Overture d'antenne	Image/sound
	15:00-15:10	10mn	Bulletin infos	Bilingual
	15:10-15:45	30mn	Clip Box	Music
	16:35-17:00	30mn	Groove	French
	18:00-18:30	30mn	Le quotidien des provinces	Bilingual
	19:30-20:00	30mn	The 7:30 News	English
	20:00-20:30	30mn	Terra Nostra	French
	20:30-21:00	30mn	Le Journal	French
	21:00-21:30	30mn	Annonces	Adverts
	21:30-22:30	60mn	Espace Francophone	French
	22:30-23:30	30mn	CRTV Late Night News	Bilingual

Day	Time	Duration	Programme	Language
Friday	13:55-14:00	5mn	Overture d'antenne	Image/sound
	15:00-15:10	10mn	Bulletin infos	Bilingual
	15:10-15:45	30mn	Clip Box	Music
	17:00-18:00	60mn	Sports Vision	English
	18:30-19:00	30mn	Youth fights AIDS in rural area	English
	19:00-19:30	30mn	Scalpel	French
	19:30-20:30	60mn	Connaissance de l'Islam	French (Arabic)
	20:30-21:30	60mn	Journal Bilingue	Bilingual
	21:30-21:45	15mn	Annonces	Adverts
	23:00-23:30	30mn	CRTV Late Night News	Bilingual
Saturday	11:55-12:00	5mn	Overture d'antenne	Image/sound
	12:00-12:15	15mn	Bulletin infos	Bilingual
	12:30-13:30	60mn	Evasion	French
	13:30-14:00	30mn	Stargate SG I	French
	14:00-15:00	60mn	Délire	French
	15:00-15:30	30mn	Copain copine	French
	18:00-18:56	56mn	Tube Vision	French
	19:00-20:00	60mn	Edito	French
	20:30-21:30	60mn	Journal Bilingue	Bilingual
		22:00-22:30	30mn	Universalis
Sunday	08:55-09:00	5mn	Overture d'antenne	Image/sound
	09:00-09:30	30mn	Keep fit	French
	09:30-10:00	30mn	Chorales	Choral singing
	10:00-11:00	60mn	Le culte protestant	Varies: French/English
	12:00-12:30	30mn	Super Book	French
	12:30-13:00	30mn	The World this week	English
	13:00-13:00	120mn	Tam-Tam Weekend	Bilingual
	15:45-16:15	30mn	Youth Rhapsody	French
	17:30-18:30	30mn	Journal d'Afrique	French
	19:00-20:00	60mn	Actualité Hebdo	French
	20:30-21:30	60mn	Journal Bilingue	Bilingual

Source: Compiled from the CRTV website, www.crtv.cm. The programme dates back to 2002 when it was last updated, but not much has changed in the news and major programme slots. However some series and serials like Terra Nostra, Stargate, Super Book have ended but programmes like Actualité Hebdo, Tam-Tam Weekend, Connaissance d'Islam, Evasion, Tube Vision are still broadcast.

Appendix II. Time Allocation per week in minutes

Days	Total	French		English		Bilingual		Neutral	
		No.	%	No.	%	No.	%	No.	%
Monday-Wednesday	1046	541	51.7	300	28.7	70	6.7	135	12.9
Thursday-Sunday	1536	776	50.5	180	11.7	425	27.6	155	10.2
Total	2582	1317	(51.2)	480	(18.5)	495	(19.1)	290	(11.2)

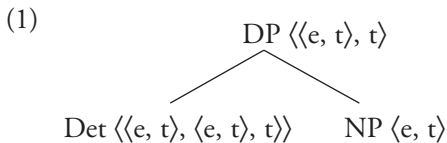
QUANTIFICATION AND COMPOSITIONAL STRATEGIES

Urtzi Etxeberria

EHU-U.Basque Country/LEHIA

1. Introduction*

The standard analysis of quantification claims that the compositionality of a Generalized Quantifier, which denotes a set of sets and is of type $\langle\langle e, t \rangle, t\rangle$, comes from combining a determiner quantifier of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t\rangle\rangle$ (following Montague (1973)'s notation) with a Noun Phrase (NP) predicate of type $\langle e, t \rangle$ as shown in (1) (cf. Barwise & Cooper (1981), Chierchia & McConnell-Ginet (1990), Dowty et al. (1981), Keenan (1987, 1996, 2002), Keenan & Stavi (1986), Heim & Kratzer (1998), Partee (1995), and references therein).



However, when applied to crosslinguistic scrutiny this analysis has been shown to have little comparative bite, since there are many languages that lack this standard construction (see Bach et al. 1995, Marlett 2000 among others).

This paper introduces an additional, very interesting problem that Basque nominal quantification poses to the standard analysis of Generalized Quantifiers and provides a new compositional¹ analysis for Basque nominal quantificational elements.

* Many people have contributed in the creation of this paper. I am especially grateful to Fernando Garcia Murga and Ricardo Etxepare for their time, guidance, encouragement and endless brainstorming sessions. Different parts of this paper were also discussed with (in alphabetical order) Xabier Artiagoitia, Gennaro Chierchia, Danny Fox, Aritz Irurtzun, Brenda Laca, Itziar Laka, Joseba Lakarra, Louise McNally, Luisa Martí, Javier Ormazabal, Juan Uriagereka, and Myriam Uribe-Etxebarria. I am also indebted to the audience of BIDE'04 for helpful comments. Of course, none of them necessarily believe anything contained herein, and I am alone responsible for errors. This research is supported by the research project grants BFF2002-04238-C02-01 of MCYT-FEDER and 9/UPV00033.130-13888/2001 of EHU-U.Basque Country, and the research grant UPV 033.129-HA036/98 by EHU-U.Basque Country.

¹ The compositionality principle can be expressed as follows: *The meaning of an expression is a function of the meanings of its parts and of the way they are syntactically combined* (see Partee 1982, 1995).

First of all, in section 2, we proceed to make a division between the so-called *strong* and *weak* Basque quantifiers considering some logical and linguistic properties that generalized quantifiers show. Taking this division as a starting point, section 3.1 introduces the problem that Basque nominal quantification poses to the standard analysis of Generalized Quantifiers. Then, section 3.2 presents and analyses the Basque article and proposes its semantic denotation, which as well as being able to account for the denotation both of count (singulars and plurals) and mass terms, also allows us (following Matthewson 2001) to make a correct semantic composition of Basque generalized quantifiers in section 3.3. A summarizing and concluding section follows.

2. Basque nominal quantification

The Basque quantifiers that are going to be analyzed in this paper are GUZTI (all), DEN (all), GEHIEN (most), BAKOITZ² (each), BATZU(E)K (some), ZENBAIT (some), ASKO (many), GUTXI (few), UGARI (many), NUMERALS, NUMERAL BAINO GEHIAGO (MORE THAN Numeral), NUMERAL BAINO GUTXIAGO (less than numeral).³

2.1. Basque strong and weak quantifiers

Barwise & Cooper (1981) offer a definition to divide quantifiers into strong and weak. The terms *strong* and *weak* are taken from Milsark (1974, 1977), where he states that only weak quantifiers can appear in *there*-insertion contexts. Milsark argues that weak determiners are *cardinal*, whereas strong quantifiers are *quantificational*. Barwise & Cooper (1981: 182) give the following definition in order to formalize the distinction:

- (2) DEFINITION: A determiner is *positive strong* (or *negative strong* resp.) if for every model $M = \langle E, \parallel \parallel \rangle$ and every $A \subseteq E$, if the quantifier $\parallel D \parallel (A)$ is defined then $A \in \parallel D \parallel (A)$. (Or $A \notin \parallel D \parallel (A)$, resp.). If D is not (positive or negative) strong then D is weak.

They also provide a test-sentence which can decide whether a quantifier has to be classified as positive strong, negative strong, or weak.

- (3) D N is a N/ are Ns

To classify the quantifiers it is necessary to check whether the above sentence is automatically valid, contradictory, or contingent.

According to Barwise & Cooper a quantifier is positive strong if the statement created is a tautology in every model in which the quantifier is defined. A quantifier will be negative strong if it is a contradiction, and a quantifier will be weak if the truth of the statement depends on the model.

² Cf. Etxeberría (2001, 2002a) for an analysis of the behaviour of this inherently distributive quantifier of Basque.

³ Basque does have more quantifiers than the ones analysed in this paper. Note however that the analysis offered in this paper can also be applied to those quantifiers. See Euskaltzaindia (1985, 1993) and Euskal Hiztegia (1996) for a mention of some of them.

The sentences that we form with Basque quantifiers (following this test sentence) are the ones we have in (4).

- (4a) Baserritar guzti-ak baserritarrak dira.
farmer all-ART.abs.pl farmer be.pl
'All the farmers are farmers.'
- (4b) Baserritar den-ak baserritarrak dira.
'All the farmers are farmers.'
- (4c) Baserritar gehien-ak baserritarrak dira.
'Most farmers are farmers.'
- (4d) Baserritar bakoitz-a baserritar bat da.
'Each farmer is a farmer.'
- (4e) Baserritar batzuek baserritarrak dira.
'Some farmers are farmers.'
- (4f) Zenbait baserritar baserritarrak dira.
'Some farmers are farmers.'
- (4g) Hainbat baserritar baserritarrak dira
'Some farmers are farmers.'
- (4h) Baserritar asko baserritarrak dira.
'Many farmers are farmers.'
- (4i) Baserritar gutxi dira baserritarrak.
'Few farmers are farmers.'
- (4j) Baserritar ugari baserritarrak dira.
'Many farmers are farmers.'
- (4k) Bi baserritar baserritarrak dira.
'Two farmers are farmers.'
- (4l) Bost baserritar baino gehiago baserritarrak dira.
'More than five farmers are farmers.'
- (4m) Bost baserritar baino gutxiago dira baserritarrak.⁴
'Less than five farmers are farmers.'

According to Barwise & Cooper a quantifier is positive strong if the statement created is a tautology in every model in which the quantifier is defined. A quantifier will be negative strong if it is a contradiction, and a quantifier will be weak if the truth of the statement depends on the model.

Universal quantifiers will be easy to classify since they always come out true even in an empty domain. Therefore, *guzti*, *den*, and *bakoitz* will be defined as positive strong. The quantifier *gehien* will be defined only if there are elements in the domain, and, when that is the case, it will be described as positive strong (following Barwise & Cooper 1981: 182). Therefore, the Basque quantifiers *guzti* (all), *den* (all), *bakoitz* (each) and *gehien* (most) should be considered positive strong quantifiers since they give tautologies as a result.

⁴ The sentences created by *gutxi* and *numeral baino gutxiago* are ungrammatical unless they appear preverbally as the examples in (4i-m) show. See Etxeberria (to appear) for a possible analysis of these facts.

It is interesting to mention the fact that there are no negative strong quantifiers in Basque, that is to say, Basque does not have quantifiers such as English *neither*. An English sentence like (5) is translated into Basque making use of a more complex construction (notice that the sentence is not completely grammatical).⁵

- (5) Neither superhero is a superhero.
 (6) ??Superheroi [bakar bat ere] ez da superheroia.
 superhero [single one even] not be.pres.sg superhero

The characterization of weak quantifiers depends on the model, in other words, if a sentence is sometimes true and sometimes false, it is neither a tautology nor a contradiction, and the quantifier is therefore classified as weak. Following this reasoning, a quantifier like *batzu(e)k* (some) in (4e) is described as weak because when the model contains two or more than two farmers the sentence is judged true; but when the model contains less than two farmers the sentence will be false. The characterization of *asko* (many) in (4h) as a weak quantifier is based on its behaviour in models that contain less than ‘many’ farmers (whatever the context-dependent number happens to be). In such a model the sentence comes out false; in all other models the sentence will be true. The Basque quantifiers that behave this way are, in addition to the already mentioned *batzu(e)k* and *asko*; *zenbait* (some), *hainbat* (some), *asko* (many), *gutxi* (few), *ugari* (many), *numerals*, *numeral baino gehiago* (more than numeral), *numeral baino gutxiago* (less than numeral).

2.1.1. Logical Properties of Basque Quantifiers

2.1.1.1. Symmetry and Intersectivity

One of the logical properties that can be used to differentiate strong quantifiers from weak quantifiers is symmetry. The equivalence that weak quantifiers fulfil is the one introduced in (7).

- (7) $D(A)(B) \leftrightarrow D(B)(A)$

The prediction is then that *den*, *guzti*, *gehien*, and *bakoitz* (described as strong quantifiers in the previous section) will not show the equivalence presented in (7), this prediction is fulfilled as the examples (8a-b) show.⁶

- (8a) Euli guzti-ak itsusiak dira.
 fly all-ART abs.pl ugly be
 ‘All the flies are ugly.’
 ((\leftrightarrow)) (gauza) itsusi guzti-ak euliak dira.
 thing ugly all fly be
 ‘All the ugly (thing)s are flies.’

⁵ Cf. Etxepare (2003).

⁶ From now on—in order to make the reading easier—only two strong and two weak quantifiers will be used in the examples. The properties that these strong and weak quantifiers show should also be taken to apply to the rest of strong and weak quantifiers.

- (8b) Euli gehien-ak itsusiak dira.
 ‘Most flies are ugly.’
 ((↔)) (gauza) itsusi gehien-ak euliak dira.
 ‘Most ugly (thing)s are ugly.’
- (8c) Zenbait euli itsusia ??da/ itsusiak dira.
 ‘Some flies are ugly.’
 ↔ (gauza) itsusi zenbait eulia da/ euliak dira.
 ‘Some ugly (thing)s are flies.’
- (8d) Euli asko itsusia ??da/ itsusiak dira.
 ‘Many flies are ugly.’
 ↔ (gauza) itsusi asko eulia ??da/ euliak dira.
 ‘Many ugly (thing)s are ugly.’⁷

The quantifiers in the examples (8c-d) fulfil the equivalence and must therefore be considered weak. These quantifiers give the cardinality of a set that is defined as the intersection of the sets A (flies in (8)) and B (be ugly in (8)). Since intersection is symmetric, the quantifiers described as symmetric are also intersective, and the roles of the A set and the B set in the relation can not be different.

- (9) A = the set denoted by *euli* (fly).
 B = the set denoted by *itsusia izan* (be ugly).

(10) Strong Quantifiers

[[A guzti-ak B dira]]	$A \subseteq B$
[[A den-ak B dira]]	$A \subseteq B$
[[A bakoitz-a B da]]	$A \subseteq B$
[[A gehien-ak B dira]]	$ A \cap B > A - B ^8$

Weak Quantifiers⁹

[[A batzu(e)k B dira]]	$ A \cap B \geq 2$
[[Zenbait A B da/ dira]]	$ A \cap B \geq 2$

⁷ The interpretation of the quantifiers *asko* (many), *gutxi* (few), and *ugari* (many) varies from context to context; it seems as though determining how many individuals count as many or few is strongly context-dependent. See Barwise & Cooper (1981), Keenan & Stavi (1986) among many others.

Observe also that the weak quantifiers *zenbait* (some), *hainbat* (some), *asko* (many), *gutxi* (few), and *ugari* (many) can agree with the verb in plural or otherwise not show any kind of agreement. When these quantifiers do not agree with the verb the sentences in (8c-d) are not grammatical unless the QP appears in focused position (cf. Etxepare 2000, Etxeberria 2001, to appear).

⁸ It is due to the fact that *guzti*, *den*, and *bakoitz* are universal quantifiers that the three elements are described the same way in (68), that is, the set denoted by the common noun is a subset of the set denoted by the VP. *Gehien* is not a universal quantifier and what $|A \cap B| > |A - B|$ means is that the cardinality of the intersection of the set A and the set B is bigger than the cardinality of the set of elements in A that are not B. The important thing to observe is that with these four quantifiers it is necessary to know the elements in A (that are B) in order to derive the truth values of the sentence, and the rest of the set B is not relevant to the truth conditions.

⁹ Weak quantifiers are assumed to be ambiguous between a strong and a weak reading. See among others Barwise & Cooper (1981), Keenan (1987), Keenan & Stavi (1986), Diesing (1992), Herburger (1997, 2000), de Hoop (1992), Milsark (1974, 1977), Partee (1988). See also Etxeberria (2004, to appear).

[[Hainbat A B da/ dira]]	$ A \cap B \geq 2$
[[A asko B da/ dira]]	$ A \cap B = \text{asko}^{10}$
[[A gutxi B da/ dira]]	$ A \cap B = \text{gutxi}$
[[A ugari B da/ dira]]	$ A \cap B = \text{ugari}$
[[Numeral A B dira]]	$ A \cap B = \text{numeral}$
[[Numeral A baino gehiago B dira]]	$ A \cap B > \text{numeral}$
[[Numeral A baino gutxiago B dira]]	$ A \cap B < \text{numeral}$

While the strong quantifiers of Basque express a proportion of the A set, weak quantifiers denote the cardinality of the set derived from the intersection of the sets A and B. For example, in a sentence like *euli guztiak itsusiak dira* the quantifier *guzti* denotes a relation between the sets A (flies) and B (the individual that are ugly). For the sentence to be true the set A must be a subset of B. If there is a member of the set A that is not also a member of the set B, the sentence will be considered false.

2.1.2. Linguistic Properties of Basque Quantifiers

2.1.2.1. Existential Sentences

It was Milsark (1974, 1977) who first pointed out that the strong/weak distinction plays an important role in the interpretation of existential *there* constructions. Existential sentences exhibit the so-called *definiteness effect*, which means that although some noun phrases are acceptable in *there*-insertion sentences, others are not. As a consequence, we distinguish between weak quantifiers which are acceptable in the position after *there be*, and strong quantifiers, which are unacceptable and anomalous in this context.

According to Milsark (1974, 1977), the quantifiers that occur in the postverbal position of *there*-sentences (or weak quantifiers) are cardinality markers, elements whose function “is to express the size of the set of entities denoted by the nominal with which they are construed” (Milsark 1977: 23). Milsark’s explanation of why only cardinality markers are allowed in *there*-sentences is as follows. Strong quantifiers are quantificational. Since *there be* is an existential quantifier, a *there*-sentence containing a postverbal strong quantifier “would have two quantifications on the NP... [which] should certainly be expected to be anomalous” (Milsark 1977: 24). If, on the other hand, the postverbal NP is cardinal (non-quantificational), no double quantification on the NP would occur and no anomaly arises.¹¹

¹⁰ Partee (1988) concludes that these quantifiers (*many* and *few*) are ambiguous between a cardinal and a proportional reading. According to her, the cardinal-weak interpretation (of English *many* and *few*) is intersective and symmetrical. The same applies to correlative Basque quantifiers *asko*, *ugari* and *gutxi*.

¹¹ Note, contra prediction, that the sentences in (ia/b) are completely grammatical,

- (ia) There was every kind of professor at that school.
- (ib) There were both varieties of rice for sale.

See McNally (1992, 1998), where she defends that the elements appearing in *there be* constructions must be predicative.

As predicted by Milsark (1974, 1977), and contrary to the behaviour of weak quantifiers (11c-d), Basque strong quantifiers are not acceptable in existential sentences (11a-b).

- (11a) *Badira zientzilari guzti-ak laborategi honetan.
yes-be.pl scientist all-ART.abs.pl laboratory this-in
'*There are all the scientist at this laboratory.'
- (11b) *Badira zientzilari gehien-ak laborategi honetan.
yes-be.pl scientist all-ART.abs.pl laboratory this-in
'*There are most scientist at this laboratory.'
- (11c) Bada/Badira zenbait zientzilari laborategi honetan.
yes-be.sg/pl some scientist laboratory this-in
'There are some scientists at this laboratory.'
- (11d) Bada/Badira zientzilari asko laborategi honetan.
yes-be.sg/pl scientist many laboratory this-in
'There are many scientists at this laboratory.'¹²

2.1.2.2. Presuppositionality

The property of presuppositionality with regard to quantificational elements has been classically illustrated by means of a paradigm introduced by Lumsden (1988) (examples (12a-d) are taken from Zucchi 1995).

- (12a) If you find every mistake, I'll give you a fine reward.
(12b) If you find most mistakes, I'll give you a fine reward.
(12c) If you find some mistake, I'll give you fine reward.
(12d) If you find four mistakes, I'll give you a fine reward.

According to Lumsden or Zucchi, the examples (12a-b) presuppose the set denoted by the common mistakes to be non-empty, whereas (12c-d) sound neutral with respect to this regard.¹³ As a consequence they claim that only strong quantifiers are presuppositional.

As expected, Basque strong quantifiers behave presuppositionally,¹⁴ in (13a-b) *den*, *guzti*, *gehien*, and *bakoitz* presuppose the existence of the set denoted by *akats* (mistake). Weak quantifiers on the other hand do not necessarily do so (13c-d).

¹² Existential sentences can also be constructed with the verb *egon* (be in a location) in most Basque dialects. Freeze (1992) claims that both the structures in (i) and (ii) are derived from the same basic structure.

(i) Badago gizon bat atean. (ii) Atean gizon bat dago.
yes-is-egon man one door-at door-at man one is-egon
'There is a man at the door.' 'There is a man at the door / A man is at the door.'

¹³ See Zucchi (1995: 77) for arguments in support of the idea "that the presuppositional characterization of strong NPs yields the most accurate predictions concerning which NPs are allowed in *there*-constructions".

¹⁴ Some authors claim that all the quantifiers are presuppositional (see von Stechow 1994). This might be so, but there is still a clear difference between the interpretations we obtain in (41a-d) where the set of mistakes is presupposed, and the interpretations we obtain in (41e-l) where this is not necessary.

- (13a) Akats guzti-ak aurkitzen badituzu, goxoki bat emango dizut.
mistake all-ART find if-aux. candy one give aux.
'If you find all the mistakes, I'll give you a candy.'
- (13b) Akats gehien-ak aurkitzen badituzu, goxoki bat emango dizut.
mistake all-ART find if-aux. candy one give aux.
'If you find most mistakes, I'll give you a candy.'
- (13c) Zenbait akats aurkitzen baduzu/badituzu, goxoki bat emango dizut.
some mistake find if-aux.sg/pl candy one give aux.
'If you find some mistakes, I'll give you a candy.'
- (13d) Akats asko aurkitzen baduzu/badituzu, goxoki bat emango dizut.
mistake many find if-aux.sg/pl candy one give aux.
'If you find many mistakes, I'll give you a candy.'

Once we have applied all these properties to the Basque nominal quantificational elements, the final division we get is presented in the following chart:

(14)

	LOGICAL PROPERTIES		LINGUISTIC PROPERTIES	
	symmetry	intersectivity	existential sentences	presuppositionality
<i>GUZTI</i>	NO	NO	NO	YES
<i>DEN</i>	NO	NO	NO	YES
<i>GEHIEN</i>	NO	NO	NO	YES
<i>BAKOITZ</i>	NO	NO	NO	YES
<i>BATZU(E)K</i>	YES	YES	YES	NO
<i>ZENBAIT</i>	YES	YES	YES	NO
<i>HAINBAT</i>	YES	YES	YES	NO
<i>ASKO</i>	YES	YES	YES	NO
<i>GUTXI</i>	YES	YES	YES	NO
<i>UGARI</i>	YES	YES	YES	NO
<i>NUMERAL</i>	YES	YES	YES	NO
<i>NUM. BAINO GEHIAGO</i>	YES	YES	YES	NO
<i>NUM. BAINO GUTXIAGO</i>	YES	YES	YES	NO

The chart in (15) offers the division of Basque quantifiers according to the strength/weakness criterion:

(15)

<p>Strong Quantifiers: GUZTI (all), DEN (all), GEHIEN (most), BAKOITZ (each).</p> <p>Weak Quantifiers: BATZU(E)K (some), ZENBAIT (some), HAINBAT (some), ASKO (many), GUTXI (few), UGARI (many), NUMERALS, NUMERAL BAINO GEHIAGO (more than numeral), NUMERAL BAINO GUTXIAGO (less than numeral).</p>

3. Towards a Compositional Analysis of Basque Generalized Quantifiers

3.1. Problem for the Standard Analysis of Generalized Quantifiers Theory

One crucial difference between the strong and weak quantifiers is that Basque strong quantifiers (*GUZTI* (all), *DEN* (all), *GEHIEN* (most), *BAKOITZ* (each)) must necessarily appear with the article *-A/-AK* as examples (16-17) shows.

- (16a) [Ikasle **guzti-ak**] berandu etorri ziren.
 [student all-ART.pl(abs)] late come aux.past.pl
 ‘all the students came late.’
- (16b) *[Ikasle **guzti**] berandu etorri dira.
- (17a) [Ikasle **gehien-ak**] berandu etorri ziren.
 [student most-ART.pl(abs)] late come aux.past.pl
 ‘most student came late.’
- (17b) *[Ikasle **gehien**] berandu etorri ziren.

On the other hand, and in opposition to what happens with strong quantifiers, Basque weak quantifiers (*BATZU(E)K* (some), *ZENBAIT* (some), *ASKO* (many), *GUTXI* (few), *UGARI* (many), *NUMERALS*, *NUMERAL BAINO GEHIAGO* (more than numeral N), *NUMERAL BAINO GUTXIAGO* (less than numeral N)) do not take *-A/-AK* as observed in the following example.

- (18a) [**Zenbait** ikasle] berandu iritsi zen/ziren.
 [some student] late arrive aux.sg/aux.pl.past
 ‘some students arrived late.’
- (18b) *[**Zenbait(-ak)** ikasle(-ak)] berandu iritsi ziren.
- (19a) [Ikasle **asko**] berandu iritsi zen/ziren.
 [student many] late arrive aux.sg/aux.pl.past
 ‘many students arrived late.’
- (19b) *[Politikari **asko-ak**] berandu iritsi ziren.
- (20a) [Ikasle **batzu(e)k**] berandu etorri ziren.
 [student some] late come aux.past.pl
 ‘some students came late.’
- (20b) *[Ikasle(-ak) **batzu(e)k(-ak)**] berandu etorri ziren.

The problem that Basque poses to the standard theory of Generalized Quantifiers derives from the necessity that Basque strong quantifiers have to appear with the article. These quantifiers always make use of an element more than the ones presupposed by the standard theory; and besides the combination of the Noun Phrase with the Quantifier, the presence of the article is essential if the sentence is going to be grammatical [NOUN+STRONG QUANTIFIER+ARTICLE].

Now that the problem has been introduced, and maintaining the compositionality of Basque Generalized Quantifiers as the final aim, the following subsection is dedicated to decide the function that the Basque article fulfils both in simple DPs and inside QPs.

3.2. The Basque article: *-A/-AK*

In principle, it seems as though the Basque article behaves like the definite article of Spanish or English. The English sentences in (21a-b), for example, are translated into Basque making use of the article *-A* (singular in this case) as can be seen in the sentences (22a-b).

(21a) Peru read the book.

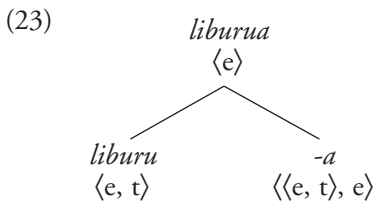
(21b) The dog ate the candy.

(22a) Peru-k liburu-*a* irakurri zuen.
Peru-erg.sg book-ART.sg(abs) read aux.past.sg

(22b) Txakurr-*ak* goxoki-*a* jan zuen.
dog-ART.sg(erg) candy-ART.sg(abs) eat aux.past.sg

Observing these sentences we could come to the conclusion that the interpretation forced by the Basque article is a specific interpretation. In the sentences in (22) it is possible to say that the speaker has a particular referent in mind, that is, presupposes¹⁵ the existence of a book in particular, a dog in particular and a candy in particular respectively.

Consequently, the lexical entry of the Basque article can be said to be equivalent to the definite article of Spanish or English. According to this definition, the article (of type $\langle\langle e, t \rangle, e\rangle$) combines with a set (of type $\langle e, t \rangle$), which is formed by a single element, and gives as a result an entity (of type $\langle e \rangle$) that is that single element of the contextually relevant set.



The described specific interpretation is the most natural (and in many occasions) the only possible interpretation (as just seen in examples (22)). However, there are situations where the article can also force a non-specific interpretation (indefinite-like) as can be seen in (24) (cf. Artiagoitia 1997, 2001, Laka 1993 among others).

(24) Peru-k kotxe-*a* erosi zuen.
Peru-erg.sg car-ART.sg(abs) buy aux.past.sg

Specific: 'Peru bought the car'

Non-specific: 'Peru bought a car'

¹⁵ The classical definition of presuppositional relations can be described as follows:

The sentence A presupposes B if and only if A entails B and the negation of A entails B.

It follows from this definition that (in bivalent systems) B has to be true for A to have a truth-value. In other words, if B is false A can not receive a classical truth-value. This is the logical test to see whether a sentence A presupposes a sentence B.

The non-specific interpretation of *car* in (24) can not be expressed making use of the terms described for the definite article.¹⁶ In this particular case, the speaker does not have a particular referent in mind; therefore, the interpretation obtained can not be presuppositional. This interpretation of the article is can be said to be similar to the interpretation of the indefinite article; thus, it is possible to conclude that the non-specific interpretation of the Basque article introduces variables over Choice Functions (the car chosen from the set of cars by the contextually relevant choice function) (see Reinhart 1997, Winter 1997, Kratzer 1998, Matthewson 1999). There is no change in the semantic type of the article ($\langle\langle e, t \rangle, e \rangle$ type), that is, the combination of the Noun Phrase and the article yields again individuals of type $\langle e \rangle$. The only difference between the specific and the non-specific interpretation is the presuppositionality, which is only present in the specific interpretation.

The ambiguity is much more general when it comes to plurals and mass terms. Sentences containing [NP+plural article] sequences (except when in external argument positions where they are interpreted in a specific way obligatorily) can be interpreted both specifically and non-specifically even in constructions where the singular article Basque article -A does not admit non-specific interpretations (see example (22a)).

- (25) Peru-k liburu-ak irakurri zituen.
 Peru-erg.sg libro-ART.pl(abs) read aux.past.pl
 Specific: ‘Peru read the books’
 Non-specific: ‘Peru read books’

In contrast to the singular article (which creates singular individuals of type $\langle e \rangle$), the plural Basque article creates plural individuals (of type $\langle e \rangle$ too), but its semantic type will be equal to the singular article -A, $\langle\langle e, t \rangle, e \rangle$ (see Link 1983). The different interpretations available depend on the functional application chosen, just like with the singular article

Mass terms, which must necessarily appear with the singular article,^{17,18} show exactly the same properties shown by the plurals and can be interpreted both specifically or non-specifically (depending again on the functional application used as in the example (26)).

¹⁶ Bosque (1996) considers that sentences like the following are stereotypical and analyses them in terms of incorporation, see (i). Rodriguez (2003) assumes this analysis and applies it to Basque.

(i) Peru se ha comprado coche
 Peru refl.pronoun have buy car
 ‘Peru has bought a car’

¹⁷ Crosslinguistically, mass terms lose their mass denotation when combined with the plural article. Basque is not an exception.

¹⁸ One interesting property of Basque Determiner Phrases is that the presence of an overt Determiner seems obligatory, even in Determiner Phrases (as mass terms or plurals) that are described as bare in many languages.

- (26) Peru-k ardo-*a* edan zuen.
 Peru-erg.sg wine-ART.sg(abs) drink aux.past.sg
 Specific: 'Peru drunk the wine'
 Non-specific: 'Peru drunk wine'

The non-specific interpretation, where we refer to *mass* in general, is obtained when the speaker does not have a specific referent in mind. In this case, the article -A introduces a variable over Choice Functions. The Choice Function selects an element (a quantity) of the contextually relevant set (of wine).¹⁹

This parallelism between plural forms and mass terms can be explained if we assume the analysis offered by Chierchia (1998), where the so-called 'Inherent Plurality Hypothesis' is proposed. This hypothesis defends that "mass nouns come out of the lexicon with plurality already built in and that that is the (only) way in which they differ from count nouns" (Chierchia (1998: 53).

The next section proposes a new compositional analysis of Basque nominal quantificational elements.

3.3. Compositional Analysis of Basque Generalized Quantifiers

As already mentioned in section 3.1, the problem that Basque poses to the standard theory of Generalized Quantifiers derives from the necessity that Basque strong quantifiers (*guzti* (all), *den* (all), *gehien* (most), *bakoitz* (each)) have to appear with an element more than the ones presupposed by the standard theory; the article (singular or plural) -A/-AK. In the proposal that is going to be developed in this paper²⁰ the compositionality of Basque strong quantifiers does not maintain the word order shown by the surface structure [NOUN+QUANTIFIER+ARTICLE] (see examples (16-17)). Instead, and following Matthewson (2001), first, the Nominal Phrase combines with the article yielding an individual of type $\langle e \rangle$ (singular or plural), and in a second step, the quantifier quantifies over parts of the created individual.

Let us construct an example so that we can see and check how the composition of a Basque strong quantifier proceeds. The nouns with which the article combines will be marked as count terms or mass terms in the Lexicon.

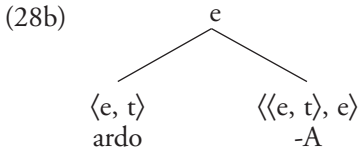
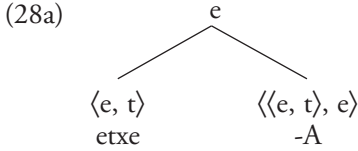
- (27a) *etxe* (house) [+ COUNT / - MASS]
 (27b) *ardo* (wine) [- COUNT / + MASS]

When -A is attached to a count noun like the one in (81a) it creates a singular individual and asks to find in the context a singular individual with the property of being a house. When the article -A is combined with a mass term (*ardo* in (81b)) it will be necessary to find in the context an individual with the property of being wine. Thus, the article combines with a one-place predicate of type $\langle e, t \rangle$ and creates

¹⁹ The function fulfilled by the Basque article -A, that must necessarily appear with mass terms, can be related to the function fulfilled by the French partitive element 'du' in 'du vin' (French mass terms can not be bare in French). Both -*a* (in Basque) and 'du' (in French) select a select a quantity of the contextually relevant set, wine in this particular case.

²⁰ See also Etxeberria (2004).

an individual. *-A/-AK* can be said to be of type $\langle\langle e, t \rangle, e\rangle$, a function that takes set denoting expressions and gives individuals as a result.



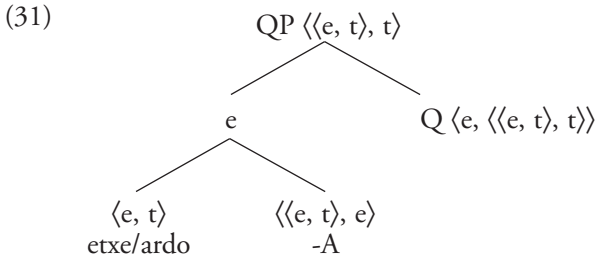
Once the structures in (28) have been built the quantifier comes into play; what the quantifier does is quantify over parts of the individual denoted by the combination of the noun and the article. When the quantifier joins something like *etxe+A*, i.e., a count noun, it will quantify over parts of a contextually relevant single house, an atom that can be divided into smaller pieces (bedrooms, kitchen, toilet, etc.) but is not already considered a house.

(29) *etxe + A + GUZTI*: quantification over parts of the singular individual (house).

The procedure will be the same when the quantifier combines with a mass term like *ardo+A*, again the quantifier will quantify over parts of the contextually relevant quantity of wine.

(30) *ardo + A + GUZTI*: quantification over parts of the mass singular individual (water).

Following Matthewson’s proposal, I assume that the quantifier creates a function which takes an individual of type $\langle e \rangle$ and gives a generalized quantifier of type $\langle\langle e, t \rangle, t\rangle$ as a result. The quantifier, then, must be of type $\langle e, \langle\langle e, t \rangle, t\rangle\rangle$.



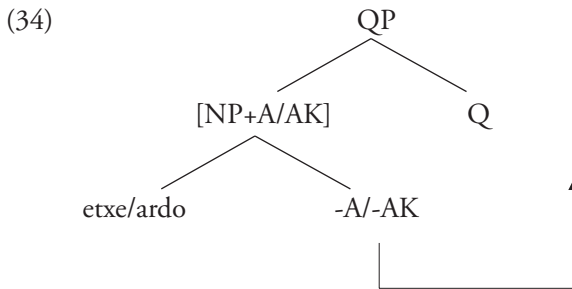
The compositional structure of a QP does not change when instead of *-A* the plural form of the article is combined with the common noun. Thus, first *-AK* combines with the common noun (count or mass) and creates an individual (plural), and then the quantifier quantifies over parts of the plural individual denoted by the [NP+Article] sequence.

One difference between the singular and the plural article is that the plural -AK asks to find in the context a plural individual (instead of a singular one) with the property of being houses (count) or with the property of being wines (count) respectively (in the cases at hand). As it has already been mentioned the mass term comes with the plurality built in and when it is combined with the plural article it loses its mass denotation and becomes a count noun (Chierchia 1998). Therefore, when the strong quantifier is added to the plural individual it will quantify over parts of the plural individual, in this case, different houses and different wines, in other words, it quantifies over different atoms.

- (32) *etxe* + AK + GUZTI: quantification over parts of the plural individual (houses).
 (33) *ardo* + AK + GUZTI: quantification over parts of the (-mass) plural individual (different types of wine).

It is important to note that the Basque article loses its indefinite (non-specific) interpretation when combined with a quantifier; as a consequence, the only possible interpretation will be that where the speaker has a particular referent in mind, that is, the presuppositional (specific). Presuppositionality is a property that strong quantifiers of natural languages show (see section 2.1.2.2).

Once the whole structure of the (strong) generalized quantifiers is built the article must move to the final position of the quantifier phrase by some Phonological Form (PF) requirement.²¹



The proposed compositional strategy will be the one used by the Basque strong quantifiers *guzti*, *den*, *gehien*, and *bakoitz*. The first three can combine both with a singular and a plural individual, depending on whether the common noun is combined with the singular form or the plural form of the article respectively.

Bakoitz, the fourth strong quantifier that has not been mentioned, can only combine with a singular individual [NP +A]. Considering what it has been said about the contribution of the quantifiers it seems as though this quantifier could only quantify over parts of a contextually relevant singular individual of type ⟨e⟩ (parts of house or parts of wine). However, and contrary to the prediction, *bakoitz*

²¹ Elordieta (1997: 189) claims that “the determiner always appears attached as a suffix to the last element in the NP. [...] The determiner is a bound morpheme, a suffix, and it attaches to the last element of the phrase that precedes it. Thus, it would be a phrasal clitic [...] or a lexical clitic.”

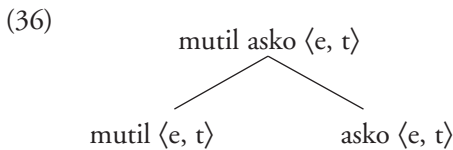
quantifies over different houses or (kinds) of wine. This might be due to *bakoitz*'s inherent distributive properties and to the fact that in order to obtain distributive interpretation the agreement with the verb must be made in singular in Basque.²²

- (35) Futbolari bakoitz-**ak** / ***ek**²³ gol bat sartu zuen.
 football player each-ART.sg(erg) / pl(erg) goal one score aux.past.sg
 'each student scored a goal.'

Up until now I have explained how the compositionality of a Basque strong quantifier would be according to this analysis, but I have not said anything about the compositionality of weak quantifiers, let us move on to that then.

As has been shown during the paper (section 3.1) weak quantifiers do not combine with an individual of type $\langle e \rangle$, that is, they do not combine with a sequence [NP+article].²⁴ The creation of a weak quantifier will not be a two-step process and they appear to be combining directly with an NP predicate of type $\langle e, t \rangle$.

Weak quantifiers are assumed to be ambiguous between a strong (proportional/presuppositional) and a weak (cardinal/non-presuppositional) interpretation (cf. Milsark 1974: 19). Following Milsark's observations, I will assume that the cardinal reading seems to be basically adjectival/modificational and not intrinsically quantificational; this reading will be the one that is weak, intersective, and symmetrical. Partee (1998: 14) claims that "these properties would follow from an analysis that treated these NPs as Kamp-Heim indefinites". Therefore, the structure we would obtain is the one we have in (36) where the weak quantifier of $\langle e, t \rangle$ ²⁵ combines with a common noun of type $\langle e, t \rangle$ to yield a predicative element of type $\langle e, t \rangle$ again.



²² See Etxeberria (2001, 2002a) for the different interpretations that different quantifiers force; in these papers it is claimed that distributive interpretations are obtained when quantificational phrases agree with the verb in singular.

²³ Historically: *-ak + (e)k* [caso ergativo] > **a-gek* > **-aek* > *-ek*.

²⁴ Note however that numerals can appear with the plural article: *Hamalau ikasle-ak* in (ib). At the moment I do not have explanation of why the order of *bost ikasleak* is Q+NP+AK, but some authors have argued that numeral quantifiers in Basque can fill two different syntactic positions, initial and final (specifier and head position). Note that other languages do also make use of these kinds of constructions that give raise to presuppositional interpretations.

- Basque: (ia) Hamalau ikasle berandu iritsi ziren.
 (ib) Hamalau ikasle-ak berandu iritsi ziren.
 English: (iia) Fourteen students arrived late.
 (iib) The fourteen students arrived late.
 French: (iiia) Quatorze étudiants sont arrivés tard.
 (iiib) Les quatorze étudiants sont arrivés tard.
 Spanish: (iva) Catorce estudiantes llegaron tarde.
 (ivb) Los catorce estudiantes llegaron tarde.

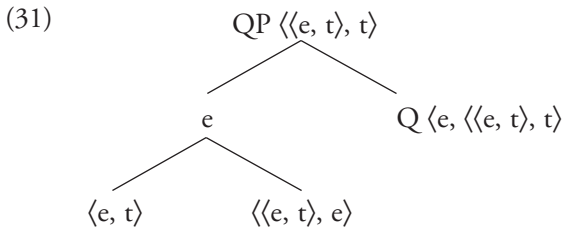
²⁵ The common noun and the weak quantifier (both of type e, t) will combine through Predicate Modification (cf. Heim & Kratzer 1998).

In order to combine the outcome of the weak quantifier and the common noun (of type $\langle e, t \rangle$) with the monadic verb phrase (*etorri* ‘come’) of type $\langle e, t \rangle$ in a sentence like (37a) the so called Existential Closure have been postulated (see Kamp 1981, Kamp & Reyle 1993), and a sentence like (37a) would have the logical form in (37b). What the Existential Closure allows us to do is combine two $\langle e, t \rangle$ type elements by introducing a quantification (since weak quantifiers do not have quantificational force themselves) and saying that their intersection is (that of *neska asko* (many girls) and *etorri ziren* (came) in (37)) is not empty, but filled with many girls that came.

(37a) Neska asko etorri ziren.
 girl many come aux.
 ‘Many girls came.’

(37b) $\exists x$ [neska asko(x) & etorri ziren(x)]²⁶

While the cardinal weak interpretation will derive from the structure offered in (36), where the cardinal quantifier directly combines with the noun; the proportional strong interpretation will be claimed to derive from a structure similar to that in (31) (repeated here for convenience) where first the noun and the article combine and then the quantifier comes into play to create a generalized quantifier of type $\langle \langle e, t \rangle, t \rangle$.



This analysis makes us think that there is an extra element in the strong interpretation of weak quantifiers which is responsible of creating an individual²⁷ and making the structure similar to that of strong quantifiers (two step process); a structure that necessarily forces a presuppositional reading. The overt version of this covert element is the partitive NP-ETATIK (‘of the NP’) shown in (37).

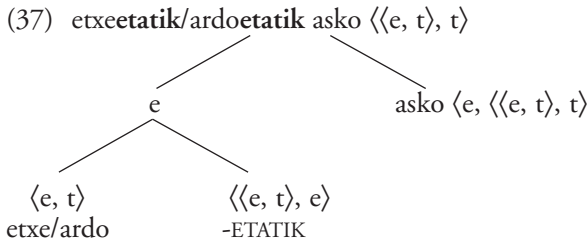
²⁶ In Discourse Representation Theory (DRT) the logical form of a sentence like (i) is expressed as in (ii) (Kamp 1981, Kamp & Reyle 1993):

- (i) Many professors arrived yesterday
 (ii) DRS for cardinal weak reading: existential quantifier implicit.

x professor (x) many (x) came (x)
--

Truth conditions: there is (was) a group of professors whose cardinality is many and which arrived yesterday

²⁷ *-etatic* in (37) derives from joining the article -AK and the partitive -TIK. Following Matthewson (2001), we are forced to treat partitive *of* as semantically vacuous, contra Ladusaw (1982). See Matthewson (2001) for extensive discussion of this point.



4. Final Remarks

This paper has explored how the internal structure of the Basque Quantifier Phrases could go. Section 2 has been dedicated to make a division between Basque strong and weak quantifiers taking into account some logical and linguistic. In section 3 I have shown that Basque strong quantifiers pose a problem to the standard compositional analysis proposed by GQT due to the element that must necessarily appear helping them: the article *-A/-AK*. Observing the ambiguity (section 3.2) shown by the article I have concluded that the Basque article can be interpreted in two ways: (i) direct functional application, (ii) variable over choice functions. In the last part of the paper (section 3.3) I have argued that Basque Generalized Quantifiers (both strong and strongly interpreted weak quantifiers) are composed in a two step process (relying on Matthewson’s 2001 work); weak quantifiers (which following Milsark (1974, 1977) do not have quantificational force) will combine directly with the common noun and will need of the existential closure in order to be interpreted.

References

Abney, S. P., 1987, *The English Noun Phrase in its Sentential Aspect*, Doctoral dissertation, MIT.

Artiagoitia, X., 2000, *Hatsarreak eta Parametroak lantzen*. Gasteiz: Arabako Foru Aldundia.

Bach, E., E. Jelinek, A. Kratzer & B. H. Partee, 1995, “Introduction”, in Bach, E., E. Jelinek, A. Kratzer & B. H. Partee (eds.), *Quantification in Natural Language*, Dordrecht: Kluwer, 1-11.

Barwise, J. & R. Cooper, 1981, “Generalized Quantifiers and Natural Language”, *Linguistics and Philosophy* 4. 159-219.

Chierchia, G., 1998, “Plurality of Mass Nouns and the Notion of ‘Semantic Parameter’”, in Rothstein, S. (ed.), *Events and Grammar*, Dordrecht: Kluwer, 53-103.

Diesing, M., 1992, *Indefinites*, Cambridge, MA: MIT Press.

Dowty, D., R. Wall & S. Peters, 1981, *Introduction to Montague Semantics*, Dordrecht: Reidel.

Elordieta, G., 1997, “Feature Licensing, Morphological Words, and Phonological Domains in Basque”, in Mendikoetxea, A. & M. Uribe-Etxebarria (eds.), *Theoretical Issues at the Morphology-Syntax Interface*, Vitoria-Gasteiz: ASJU (EHU-UPV), 171-202.

- Ettxeberria, U., 2001, "La interpretación de algunos determinantes cuantificadores del euskara", ms. EHU-U.Basque Country.
- , 2002a, "La interpretación de los cuantificadores del euskara", *Interlingüística II*, Alacant: AJL. 53-70.
- , 2002b, "Hacia la semántica de algunos determinantes cuantificacionales del euskara", ms. EHU-U.Basque Country.
- , under review, "Do all quantifiers behave the same way compositionally?", in Irurtzun, A. & U. Ettxeberria (eds.), *Selected Proceedings of 1st WoSS*, Vitoria-Gasteiz.
- , to appear, *Quantification in Basque*, Doctoral dissertation, EHU-U.Basque Country.
- Ettxepare, R., 2000, "Askoren zenbait alderdi", *Lapurdum* 5, 99-111.
- , 2003, "Negation", in Hualde & Ortiz de Urbina (eds.): 387-421.
- Euskaltzaindia, 1985, *Euskal Gramatika: Lehen Urratsak-I*, Bilbo: Euskaltzaindia.
- , 1993, *Euskal Gramatika Laburra: Perpaus Bakuna*, Bilbo: Euskaltzaindia.
- von Fintel, K., 1994, *Restrictions on Quantifier Domains*, Doctoral dissertation, UMass.
- Freeze, R., 1992, "Existentials and Other Locatives", *Lg* 68.3. 553-95.
- Heim, I. & A. Kratzer, 1998, *Semantics in Generative Grammar*, London: Blackwell Publishers.
- Herburger, E., 1997, "Focus and Weak Noun Phrases", *Natural Language Semantics* 5. 53-78.
- , 2000, *What Counts. Focus and Quantification*, Cambridge, MA: MIT Press.
- de Hoop, H., 1992, *Case Configuration and Noun Phrase Interpretation*, Doctoral dissertation, Univ. of Groningen. [Published 1997 by Garland Publishing Inc., New York.]
- Hualde, J. I. & J. Ortiz de Urbina (eds.), 2003, *A Grammar of Basque*, Berlin: Walter de Gruyter.
- Kamp, H., 1981, "A Theory of Truth and Semantic Representation", in Groenendijk, J. & M. Stokhof (eds.), *Truth, Interpretation and Information*, Dordrecht: Foris, 1-41.
- , & U. Reyle, 1993, *From Discourse to Logic*, Dordrecht: Kluwer.
- Keenan, E. L., 1987, "A Semantic Definition of 'Indefinite NP'", in Reuland, E. & A. G. B. ter Meulen (eds.), *In the Representation of (in)definiteness*, Cambridge, MA: MIT Press, 286-317.
- , 1996, "The Semantics of Determiners", in Lappin, S. (ed.), *The Handbook of Contemporary Semantic Theory*, Oxford: Blackwell Publishers. 41-63.
- , 2000, "Quantification in English is Inherently Sortal", in Shapiro, S. (ed.), *The History of Philosophy and Logic* 20. 251-65.
(URL: <http://www.linguistics.ucla.edu/people/keenan/Corcoran.pdf>)
- , 2002, "Some Properties of Natural Language Quantifiers", in G. Carlson, F. J. Pelletier, R. Thomason (eds.), Special Issue of *Linguistics and Philosophy* 25, 627-54.
- , & J. Stavi, 1986, "A Semantic Characterization of Natural Language Quantifiers", *Linguistics and Philosophy* 9, 253-326.
- Kratzer, A., 1998, "Scope or Pseudoscope? Are there Wide-Scope Indefinites?", in Rothstein, S. (ed.), *Events and Grammar*, Dordrecht: Kluwer, 163-96.
- Ladusaw, W. A., 1982, "Semantic Constraint on the English Partitive Construction", in Flickinger, D., et al. (eds.), *Proceedings of WCCFL* 1, 231-242.
- Link, G., 1983, "The Logical Analysis of Plurals and Mass Terms: A Lattice Theoretical Approach", in Bäuerle, R., C. Schwarze, & A. von Stechow (eds.), *Meaning, Use, and Interpretation of Language*, Berlin: de Gruyter, 302-23.
- Lumsden, M., 1988, *Existential Structures*, London: Croom Helm.

- Marlett, S. A., 2000, "Quantification with "all" in Seri", in *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*, Volume 44. 1-11. (URL: <http://www.und.edu/dept/linguistics/wp/2000Marlett.PDF>)
- Matthewson, L., 1999, "On the Interpretation of Wide-Scope Indefinites", *Natural Language Semantics* 7, 79-134.
- , 2001, "Quantification and the Nature of Crosslinguistic Variation", *Natural Language Semantics* 9, 145-79.
- McNally, L., 1992, *An Interpretation for the English Existential Construction*, Doctoral dissertation, University of California, Santa Cruz. [Published 1997 by Garland Publishing Inc., New York.]
- , 1998, "Existential Sentences without Existential Quantification", *Linguistics and Philosophy* 21, 353-92.
- Milsark, G. L., 1974, *Existential Sentences in English*, Doctoral Dissertation, MIT. [Published 1979 by Garland Publishing Inc., New York.]
- , 1977, "Toward an Explanation of Certain Peculiarities of the Existential Construction in English", *Linguistic Analysis* 3-1, 1-28.
- Montague, R., 1973, "The Proper Treatment of Quantification in Ordinary English", in Hintikka, J., J. Moravcsik, & P. Suppes (eds.), *Approaches to Natural Language*, Dordrecht: D. Reidel.
- Partee, B. H., 1982, "Compositionality", in Landman, F. & F. Veltman (eds.), *Varieties of Formal Semantics, Proceedings of the 4th Amsterdam Colloquium*, Dordrecht: Foris.
- , 1988, "Many Quantifiers", in Powers, J. & K. de Jong (eds.), *Proceedings of Fifth ESCOL*, The Ohio State University, Columbus: 383-402.
- , 1995, "Quantificational Structures and Compositionality", in Bach, E., E. Jelinek, A. Kratzer & B. H. Partee (eds.), *Quantification in Natural Language*, Kluwer: 541-601.
- Reinhart, T., 1997, "Quantifier Scope: How Labor is Divided Between QR and Choice Functions", *Linguistics and Philosophy* 20. 335-397.
- Rodríguez, S., 2003, "Euskal artikularen sintaxiaz", ms., EHU-U.Basque Country.
- Winter, Y., 1997, "Choice Functions and the Scopal Semantics of Indefinites", *Linguistics and Philosophy* 20, 399-467.
- Zucchi, A., 1995, "The Ingredients of Definiteness and the Definiteness Effect", *Natural Language Semantics* 3, 33-78.

HE HIMSELF: REDISCOVERING A NON-LOCAL ANAPHOR*

Gerardo Fernández-Salgueiro and Michael R. Marlo

University of Michigan, Ann Arbor

Abstract

In this paper, we discuss data that were first introduced in a brief but we believe unsettled controversy in Linguistic Inquiry in the late 1980s and early 1990s that has not received much attention since. We analyze the properties of elements like he himself in English, which has to be coreferent with a non-local c-commanding antecedent, provided there is one in the sentence. We take he himself to be the result of the adjunction of himself to he at a certain point in the derivation. After presenting our analysis we discuss its theoretical implications.

1. Introduction

In his squib, Bickerton presents novel data, such as the example in (1), in which the complex element *he himself* is obligatorily coreferent with a non-local, c-commanding antecedent.¹ Bickerton reports that coreference is impossible with the non-c-commanding element *John* and with a discourse element, such as *Bill*. Bickerton claims that *he himself* has some unusual anaphoric properties.

1. [John_i's father]_j thinks that [he himself]_{*_i,j,*_k} is smart.

In his response to Bickerton, McKay claims that Bickerton's judgments are "incorrect" (p. 370), and that Bickerton was "misled" (p. 369) by not considering his examples with sufficient context. McKay presents examples similar to those from Bickerton, but with some additional discourse context. In contrast with Bickerton, McKay claims that with the appropriate context, the *he himself* in a sentence like (2) can refer either to the non-c-commanding element *John* or to someone else in the discourse.

* We would like to thank the University of Michigan Syntax Support Group for comments on a preliminary presentation of this talk. We are particularly grateful for fruitful discussions with Sam Epstein, Cati Fortin, Dina Kapetangianni, Hamid Ouali, and Daniel Seely.

¹ In order to maintain consistency among the various examples presented here, we have abstracted away from many of the precise examples from Bickerton and McKay.

2. Unlike his father, John has gotten excellent grades throughout his educational career and has excelled in every academic pursuit that he has ever attempted. [John_i's father]_j thinks that [he himself]_{i,#j,*k} is smart.

McKay argues that the coreference possibilities of *he himself* are essentially no different from the pronominal *he*, and, as a result, pragmatic considerations determine how coreference is determined for *he himself* in any given context.

Example (3) shows the well-known fact that the pronominal *he* can refer either to *John*, *John's father*, or someone else, depending on the context.

3. [John_i's father]_j thinks that he_{i,j,k} is smart.

In this talk, we argue that McKay was wrong to simply dismiss Bickerton's judgments as "incorrect". For some speakers, even with extremely biased context, the possible antecedents for *he himself* seem to be restricted by the syntax. For example, the sentence in (4) shows that for some speakers, *he himself* can only corefer with *John's father*, even though it is dispreferred pragmatically.

4. Unlike his father, John has gotten excellent grades his entire academic career and has excelled in every academic pursuit that he has ever attempted. [John_i's father]_j thinks that [he himself]_{i,#j,*k} is smart.

Bickerton and McKay present data from separate dialects of English. In this talk, we would like to reconsider the facts from Bickerton's dialect and consider additional data as well. We offer an analysis that attempts to capture the distribution and interpretation of *he himself*, and we will discuss some potential implications for Binding Theory.

2. "Anaphoric" vs. "emphatic" himself

Before we present the core data in detail, we want to make it explicit what data we intend to account for in our analysis and what data we don't intend to account for. McKay claims that *himself* simply functions as an emphatic pronoun, as in (5) – (7), where *himself* modifies an NP/DP resulting in the meaning, 'NP_i, and no one but NP_i'.

5. John himself did it. (Bickerton 1987: 345)
 6. I gave it to Bill himself. (Bickerton 1987: 345)
 7. John gave it to Mary himself. (Bickerton 1987: 345)

For the dialects under consideration, the "anaphoric" use of *himself* are distinct from "emphatic" uses of *himself* (Bickerton 1987: 345). At least some speakers of Bickerton's dialect with "anaphoric" *himself* also have an "emphatic" *himself*. In these dialects, the emphatic use seems to require a special intonation contour, which we will represent with small caps *HE HIMSELF*. For these speakers, *HE HIMSELF* has the same coreference properties as those reported for *he himself* in McKay's dialect and, therefore, as for the pronominal *he*. With the appropriate context, emphatic *HE HIMSELF* can be coreferent with a matching non-c-commanding antecedent or with a matching antecedent that is not the closest c-commanding antecedent. For example,

in (8) (cf. (2) and (4) above), in which the context biases reference to *John*, *HE HIMSELF* can corefer with *John*.

8. Unlike his father, John has gotten excellent grades throughout his educational career and has excelled in every academic pursuit that he has ever attempted. [John_i's father]_j thinks that [HE himself]_{i,#j,#k} is smart.

At first glance, this suggests that perhaps McKay was “mised” by prosodic factors. However, McKay (1991: 370) explicitly addresses this concern: “the examples considered here do not need special stress.” Therefore, we conclude that we really are dealing with separate dialects. We also note that *he alone* has the same possible antecedents as “emphatic” *HE HIMSELF*, McKay's *he himself*, and the pronominal *he*, as shown in (9).

9. Unlike every other member of his family, John has gotten excellent grades throughout his educational career and has excelled in every academic pursuit that he has ever attempted. [John_i's father]_j thinks that [he alone]_{i,#j,#k} is smart.

In short, it appears that in the dialect under investigation there are two *he himself*s: one that is “anaphoric”, and one that is “emphatic”. In this talk, we will focus on the properties of the “anaphoric” *he himself*, not on the “emphatic” *he himself*. In addition to the distinction we make between “anaphoric” and “emphatic” *he himself*, we also assume that the anaphoric *he himself* is qualitatively different from examples where *he himself* functions as a VP-adjunct, as in (10).

10. John did the work himself.

3. The core data

As is well known, the sentence in (11) is ungrammatical because the anaphor *he himself* must be locally bound (Condition A), but it is not. Therefore, anaphors, such as *he himself*, are generally restricted from subject positions because they cannot have a local antecedent.

11. [John_i's father]_j thinks that himself_{*i,j,*k} is smart.

The example from (3) above, repeated here as (12), shows that pronominals, such as *he*, are locally free (Condition B), and coreference to non-local antecedents or discourse entities is possible.

12. [John_i's father]_j thinks that he_{i,j,k} is smart.

The example in (13), repeated from (1) above, shows that *he himself*, unlike *he himself* alone, is licit in subject position of the embedded clause, and it must be coreferent with the, non-local c-commanding antecedent *John's father*. As Bickerton (1987: 347) pointed out, *he himself* therefore seems to have hybrid properties of both pronominals and anaphors. On the one hand, *he himself* behaves as a pronoun in that it can appear in subject position, receiving an external theta role, checking Nominative Case, and corefering with a non-local antecedent. On the other hand,

though, *he himself* behaves as an anaphor since it is bound by a c-commanding antecedent. After discussing the relevant examples in detail in the following section, we will provide an analysis in section that attempts to deduce these hybrid properties from independent principles in the grammar.

13. [John_i's father]_j thinks that [he himself]_{*i,j,*k} is smart.

An interesting property of *he himself* is that it can skip over a c-commanding antecedent if it disagrees in phi-features to corefer with a higher c-commanding antecedent that does agree in phi-features. Example (14) shows that the closest c-commanding antecedent *Mary* disagrees in gender with *he himself*. However, the higher c-commanding antecedent *John's father* agrees in its phi-features with *he himself*. Coreference is possible only between *he himself* and *John's father*. As we saw before, *he himself* cannot corefer with the non-c-commanding antecedent *John*.

14. [John_i's father]_j said that [Mary]_k believes that [he himself]_{*i,j,*k,*l} is smart.

Although *he himself* can skip over a c-commanding antecedent that disagrees in phi-features to corefer with a higher c-commanding antecedent that does agree in phi-features, a c-commanding antecedent that agrees in phi-features with *he himself* cannot be skipped. As shown by (15), coreference is only possible between *he himself* and the closest matching c-commanding antecedent *Bill's brother*. The examples in (14) and (15) show that coreference is established between *he himself* and the closest c-commanding antecedent that agrees in phi-features.

15. [John_m's father]_n said that [Bill_i's brother]_j believes that [he himself]_{*i,j,*k,*m,*n} is smart.

The data become a bit more complex and interesting when considering wh-elements extracted from the subject position that c-commands *he himself*. As shown in (16), when the wh-element matches *he himself* in its phi-features, the wh-element is coreferent with *he himself*. In terms of overt elements, *Bill* is the closest matching c-commanding antecedents; the overt wh-element is higher in the structure. However, in this example, coreference can only occur between *he himself* and the wh-element; coreference with *Bill* is blocked. Note, though, that the wh-trace is the closest c-commanding antecedent of *he himself*; before wh-movement, the wh-element was the closest c-commanding antecedent of *he himself*.

16. Who_i did Bill_j say t_i believes that [he himself]_{i,*j} is smart?

The sentence in (17) shows what happens when the wh-element and therefore the wh-trace disagree with *he himself* in phi-features. As expected, the disagreeing wh-trace is skipped, and *he himself* corefers with the closest matching c-commanding antecedent, *Bill*.

17. [What girl]_i did Bill_j say t_i believes that [he himself]_{*i,j} is smart?

Given that *he himself* can only refer to the closest matching c-commanding antecedent, an interesting property of *he himself* emerges in its interaction with quantifiers. For comparison, consider the example in (18), in which the quantifier *everyone* is the subject of the matrix clause, while the pronominal *he* is the subject of the em-

bedded clause. The pronominal *he* can either refer to the subject of the matrix clause or to a discourse entity. As a result, there are two possible interpretations of the sentence in (18), as indicated by the two diagrams below the sentence. In one reading, in which *he* refers to a discourse entity, there is one person who everyone thinks is smart. In the other reading, in which *he* refers to the subject of the matrix clause, *everyone*, the interpretation is that for each person X, X thinks that X is smart.

18. Everyone_i thinks that he_{i,j} is smart.



As we noted above, the possible antecedents of *he himself* are more restricted than those of the pronominal *he*. Whereas the pronominal *he* can refer either to an antecedent within the same sentence or to a discourse entity, when there is a matching c-commanding antecedent, *he himself* can only refer to the closest matching c-commanding antecedent. As a result, in (19), *he himself* can only corefer to the subject of the matrix clause, *everyone*. Therefore, there is only a single possible interpretation of the sentence in (19): for each person X, X thinks that X is smart.

19. Everyone_i thinks that [he himself]_{i,j} is smart.



4. Discourse reference

In the previous section, we established that *he himself* corefers with the closest c-commanding antecedent that matches in its phi-features. The whole picture is somewhat more complicated in that in the absence of a c-commanding antecedent, coreference can be established between *he himself* and a discourse referent or with a non-c-commanding antecedent. For example, the sentence in (21), based on Bickerton's example in (20), has no-c-commanding antecedents, but there are two non-c-commanding antecedents, *Mary* and *Susan*, either of which can corefer with *she herself*.

20. The essays that Mary_i wrote were things that [she herself]_i attached little importance to. (Bickerton 1987: 347)
21. The essays that Mary_i wrote to Susan_j were things that [she herself]_{i,j} attached little importance to.

In addition, Bickerton gives examples, such as the one in (22), in which there is no c-commanding antecedent to *she herself*. As a result, *she herself* can corefer with the discourse entity, *Mary*.

22. A: How will Mary_i do in the exam? (Bickerton 1987: 346)
 B: I don't know, but [she herself]_i says she'll pass.

In these examples involving *he himself* and no matching c-commanding antecedent, syntactic factors cannot determine coreference. As McKay argues, it is reasonable for us to admit that coreference is determined pragmatically. It is important to keep in mind, however, that "The presence of a c-commanding antecedent guarantees that *he himself* will not corefer with a non-c-commanding antecedent even in the same sentence." Similarly, discourse reference is only possible when there is no c-commanding antecedent that agrees in phi-features (Bickerton 1987: 346). As shown in (23), which parallels (22) but adds a c-commanding antecedent, coreference between *she herself* and the discourse entity is impossible; *she herself* can only corefer with the matching c-commanding antecedent.

23. A: How will Mary_i do in the exam? (Bickerton 1987: 346)
 B: I don't know, but Susan_j says that [she herself]_{i,j} will pass.

5. Analysis

As pointed out in the introduction, the complex *he himself* has properties of both pronouns and anaphors combined. One possible analysis is that *he himself* is a lexical item with idiosyncratic properties, like being able to bind a c-commanding antecedent outside a local domain (roughly Bickerton's approach). However, we believe that the properties of *he himself* follow from more general properties of the grammar.

We propose that *he himself* is not a lexical item but actually a complex form resulting from an adjunction operation. Before we show our analysis, consider (11) again, repeated here as (24):

24. *[John_i's father]_i thinks that himself_i is smart.

There are two main reasons why this sentence is ungrammatical. On the one hand, *himself* is an accusative case DP in Spec-TP, which leads to a crash at LF because the Case-F on the DP and the phi-Fs on Tense have not checked/deleted. On the other hand, *himself* is an anaphor, and as such, requires an antecedent in its local domain. Imagine that English happened to have a nominative version of *himself*, something like *herself*. Would the sentence still be ungrammatical because of the binding theory violation? Consider (25) and (26) (discussed in Williams 1994):

25. [John and Mary]_i think that [each other]_{i,j} is smart
 26. Mary_i said [her own]_{i,j} mother would do it

In these examples, *each other* and *her own* bind an antecedent that is outside of their local domain, which means that binding theory also has to aim at explaining binding properties outside a local domain. It seems reasonable to suggest, then, that (24) above is ungrammatical because of Case, and it actually reflects a gap in the English anaphor paradigm, which has nominative, accusative, and possessive reciprocals, and accusative and possessive reflexives, but no nominative reflexive.

Why does *he himself* have hybrid properties? We propose here that the complex *he himself* is the result of an operation that adjoins himself to *he* at a certain point in the derivation. To be more precise, we propose that adjoining himself to *he* restricts the range of referents that *he* alone would have. Consider (27), for example.

27. John_i said that he_{i,j} is smart.

He can refer to any male person, including *John*, given that it has the feature [+pron]. After binding possibilities are evaluated, *himself* is acyclically adjoined to *he* and adds to it the property that it has to be bound by a matching c-commanding antecedent. For this sentence, the only matching c-commanding antecedent is *John*. As *John* is the only antecedent compatible with both *he* and *himself*, obligatory co-reference occurs.

Checking theory and Full Interpretation force us to assume this is an adjunction operation. If it were not so, *himself*, as a DP with accusative case, would need to check its Case-Fs, which obviously cannot be checked in the same domain as the nominative subject's domain. Adjuncts, even when they are DPs, do not have to check Case-Fs:

28. I saw the movie [the other day].

29. He prepared the food [himself].

The idea that the reflexive is inserted in order to restrict the range of possible referents provides us with an explanation of why (30) and (31) below are not acceptable:

30. ?*I think that [I myself] am smart.

31. ?*You think that [you yourself] are smart.

In these examples the reference of the pronouns *I* and *you* is already unique, speaker and hearer, respectively. Therefore, the insertion of *myself* and *yourself* results in an unacceptable sentence.

6. He himself in non-subject positions

Bickerton (1987: 347) argues that “he himself is confined to positions that are nominatively Case-marked.” This is not the case for all speakers, however. Some speakers do accept sentences like the ones in (32) and (33), where the anaphor is adjoined to an object (accusative) and to the object of a preposition (oblique):

32. [Mary_i's daughter]_j thinks that Sally likes [her herself]_{i,j}

33. [John_i's father]_j believes that Bill was thinking about [him himself]_{i,j}

The data seems to be clearer in Spanish, perhaps because in this language the pronoun and the reflexive are phonologically distinct (see Baker 1995):

34. María_i piensa que Luisa_j se lo dio a ella_{i,*j,k}
 Mary thinks that Luisa her it gave a her
 'Mary thinks that Luisa gave it to her.'
35. María_i piensa que Luisa_j se lo dio a [ella misma]_{i,*j,*k}
 Mary thinks that Luisa her it gave a her self
 'Mary_i thinks that Luisa gave it to her_i.'
36. Juan_i nos habló de él_{i,j} ayer
 John us spoke of him yesterday
 'John told us about him/himself yesterday.'
37. Juan_i nos habló de [él mismo]_{i,*j} ayer
 John us spoke of him self yesterday
 'John told us about himself yesterday.'

The analysis that we presented in the previous section can be extended to these examples as well. In (32), for example, *her* alone could refer to *Mary*, *Mary's daughter*, or some other woman. Adjoining *herself* to *her* forces coreference with the closest matching c-commanding antecedent that *her* alone can have. Notice that *her herself* cannot refer to *Sally* here, because that would violate Binding Theory Principle B, which we assume is evaluated before *herself* is adjoined. Actually, these sentences constitute evidence that it is the reflexive that is adjoined to the pronoun and not the other way around. If this were true, the reflexive would bind the local antecedent, and then adjoining the pronoun would contradict this, since the local antecedent is actually the only element the pronoun cannot bind.

7. Theoretical implications of our analysis

In our analysis, we have been tacitly assuming two stages in the process of (co)reference assignment. First, 'classic' Binding Theory evaluates/assigns the possible coreference possibilities for the elements that were cyclically inserted in the derivation, and the result of this evaluation cannot be contradicted, only restricted. There is actually independent evidence that supports this idea. Consider the examples in (38) and (39):

38. John_i took a picture of him_i/himself_i
 39. John_i read a book about him_i/himself_i

In (38), only *himself* can refer to *John*, because it's part of a complement. Both the pronoun and the anaphor are possible, however, when they are part of an adjunct, as in (39). This means that Binding Theory cannot make predictions about pronouns or reflexives inserted by adjunction.

But what regulates then (co)reference assignment for adjuncts (i.e., after the cycle)? What syntactic relation(s) is/are relevant at this second stage? We propose here that there must be another set of Binding principles to account for coreference assignment after the cycle. In (39), for example, *him*, apart from referring to *John*,

could refer to any other male referent in the discourse, or to a higher masculine singular DP if there were one. *Himself*_i however, can only refer to *John* in this sentence. Returning now to the phenomenon we are exploring, the [pronoun + reflexive] complex does not behave in the same way as a regular anaphor would do.

There are two main differences: First, a regular anaphor (i.e., an anaphor inserted cyclically in the derivation) cannot bind a DP if there is another DP that is closer. If the anaphor and the closest DP do not match then the result is an ungrammatical sentence. Conversely, the *he himself* complex binds the closest matching antecedent. In other words, there appears to be a kind of intervention effect in the cycle that disappears at this second stage. Examples (40) and (41) illustrate this contrast very clearly:

40. *[John and Mary]_i think that Bill_j believes that [each other]_{i,j} is smart
(cf. ok John and Mary think that each other is smart)
41. [John and Mary]_i think that Bill_j believes that [they themselves]_{i,*j} are smart

Second, a regular anaphor must always be bound. If there is no feature matching between the anaphor and its antecedent, the sentence is ungrammatical. In the case of *he himself*, if a suitable antecedent is not found, it can refer to a discourse entity. Examples (42) and (43) illustrate this. These sentences are intended to be the answer to something like *how will Mary do in her exam?*

42. *I don't know, but herself says she will pass.
43. I don't know, but [she herself] says she will pass.

8. Conclusion

In this paper, we have analyzed the interesting properties of elements like *he himself* in English. We have argued that the Bickerton-McKay controversy was unsettled for the simple reason that they were talking about two different dialects of English. We agreed with Bickerton in that *he himself* has to be coreferent with a non-local *c*-commanding antecedent but we offered an analysis which is not based on 'unusual' properties of the lexical item *he himself*. Instead, we proposed an adjunction operation and tried to link the properties of *he himself* to more general properties of adjuncts and their interaction with Binding Theory principles.

References

- Baker, C. L., 1995, "Contrast, Discourse Prominence, and Intensification, with Special Reference to Locally Free Reflexives in British English", *Lg* 71, 63-101.
- Bickerton, D., 1987, "*He himself*: Anaphor, Pronoun, or ... ?" *LI* 18, 345-48.
- McKay, T. J., 1991, "*He himself*: Undiscovering an anaphor", *LI* 22, 368-73.
- Williams, E., 1994, *Thematic Structure in Syntax*. Cambridge: MIT Press.

MINIMALIST EDGE COORDINATIONS*

Ángel J. Gallego

Universidad Autònoma de Barcelona

0. Introduction

In this paper I provide a preliminary analysis of some structures in which a complex correlative pattern obtains involving contrastive focal particles and polarity effects. In particular, I would like to assess the nature of structures such the ones in (1), which Bianchi & Zamparelli (2001) dub “edge coordinations”:

- (1) a. Juan talked not only to Lucy, but also to Mary.
- b. Juan didn't talk to Lucy, but to Mary.

Although I think that the basic analysis can carry over to some salient counterparts such as (*n*)*either X ... (n)or Y* and *both X ... and Y*, I will focus on the correlative pair *not (only) X ... but (also) Y*. It is furthermore appealing to extend a possible analysis to other kinds of correlative structures, mainly displayed in paratactic-like constructions (e.g., *when...then*, *as...as*, *more...than*, *if... then*, etc.):

- (2) a. *When* John arrives, *then* we will go out.
- b. John is *as* smart *as* Mary.
- c. John made *more* mistakes *than* her sister.

The analysis I concentrate on here must be taken as tentative, since it is being further developed in work in progress; I will, nonetheless, settle the scene for a minimalist approach not only to these constructions, but also to the more general phenomenon of “doubling” (or resumption) in structures that, at first glance, seem to impose some sort of parallelism requirement, strongly resembling the facts studied by Fox (2000), Belletti (2003), Torrego (1995; 1998), Uriagereka (1995a; 2001) and Boeckx (2003), among others.

The paper is divided as follows. Section 1 examines the data concerning edge coordinations, in particular two slightly (but crucially) different coordination patterns

* I would like to thank José M. Brucart, Urtzi Etxeberria, Ricardo Etxepare, Raquel González, Aritz Iruartzun and Juan Uriagereka for suggestions and helpful discussion on some of the issues that are assessed in this paper. I assume the errors. This research was supported in part by a grant of FEDER/Ministerio de Ciencia y Tecnología (BFF2003-08364-C02-02) - DURSI/Generalitat de Catalunya (2001SGR 00150).

are shown, each of them illustrating remarkable syntactic asymmetries. In the next section, I review Bianchi & Zamparelli's (2001) analysis. In section 3, I spell out my solution to these structures, which capitalizes on Brucart's (1987; 1999) treatment of «corrective negation». Section 4 concentrates on the leading role of focus in the derivation of edge coordinations. Section 5 contains a summary.

1. Edge coordinations: the asymmetries

Traditionally, “correlative structures” belong to a rather generous list of constructions that were placed somewhere between subordination and coordination, with some formal marks being used as syntactic cues to draw the line. As I said in the outset, the specific structures that are assessed here show different traits that must be highlighted; first of all, they all contain a conjunctive head endowed with a polarity nature that must establish a checking operation in its specifier; second, contrastive focus plays a leading role; finally, questions arise as whether these structures display either some sort of ellipsis process or just a “corrective coda”.

Yet, what I would like to concentrate on in this section is the sharp asymmetries that the examples in (1) illustrate (repeated below as (3)). Following Bianchi & Zamparelli (2001), they will be referred to as *adjacent* and *non-adjacent* orders, respectively;¹ as is clear, in the first one, the whole coordinate structure (whose parts I will be labelling here ‘head’ and ‘coda’, just for the sake of exposition) forms a continuous cluster-like string, with the negative particle introducing the first correlate (or ‘head’):

- (3) a. Juan talked **not** (*only*) to Lucy, **but** (*also*) to Mary. (ADJACENT)
 b. Juan didn't talk to Lucy, **but** to Mary. (NON-ADJACENT)

Note, moreover, that both patterns can be found either at the beginning or at the end of the sequence.

- (4) a. **Not** (*only*) MARY, **but** (*also*) LUCY he decided to invite. (ADJACENT INITIAL)
 b. **Not** (*only*) Mary did I invite, **but** (*also*) Lucy. (NON-ADJACENT INITIAL)
 (5) a. He invited **not** (*only*) Mary, **but** (*also*) Lucy. (ADJACENT FINAL)
 b. I didn't invite (*only*) Mary, **but** (*also*) Lucy. (NON-ADJACENT FINAL)

Let us now move to the asymmetries. In first place, it must be noted that, while in the non-adjacent order the ‘coda’ can be dropped, this possibility is ruled out in the adjacent order. So, truncation is possible in the first case, contrary to what happens when the adjacent order obtains.

- (6) a. John called **not** Mary *(, **but** Lucy). (ADJACENT)
 b. John didn't call Mary (, **but** Lucy). (NON-ADJACENT)

¹ Throughout the paper, the conjunctive heads will appear in boldface. At the same time, I will use brackets to indicate some optional focal particles that can show up in these structures (basically, *only* and *also*).

As noted by Bianchi & Zamparelli (2001), things get better when *only* appears in the head of the coordination, but not extremely so:

- (7) a. He saw **not** Mary *(**but** Lucy). (ADJACENT)
 b. He saw **not only** Mary ??(**but (also)** Lucy). (ADJACENT)

This contrast seems to be also tenable in other languages, like Catalan, where the contrast is still insufficient to yield a grammatical result.

- (8) a. En Joan va veure **no només** la Maria ??(, **sinó també** la Laura).
 (Catalan)
 The Joan AUX-3SG to-see not only the Maria, but also the Laura
 ‘Joan saw not only Maria (, but also Laura)’
 b. En Joan va veure **no** la Maria *(, **sinó** la Laura).
 (Catalan)
 The Joan AUX-3SG to-see not the Maria, but the Laura
 ‘Joan saw not Maria (, but Laura)’

The second asymmetry deals with agreement effects between verb and subject and has to assume that a process of ellipsis is at stake. Since the point to be made is more salient in languages that show overt agreement, I will illustrate it with Italian and Spanish; the important thing to notice here is that, again, only the non-adjacent order allows agreement, the adjacent one requiring strict identity between the two verbs (the overt and the assumed elliptical one):²

- (8) a. ?Ha hablado **no** Juan, **sino han hablado** sus primos.
 (ADJACENT-Spanish)
 Have-3SG talked not Juan, but have-3PL talked his cousins
 ‘Juan has not talked, but his cousins’
 b. **No** ha hablado Juan, **sino han hablado** sus primos.
 (NON-ADJACENT-Spanish)
 Not have-3SG talked Juan, but have-3PL talked his cousins
 ‘Juan hasn’t talked, but his cousins’
 (9) a. ??È arrivato **non** Gianni, **ma sono arrivati** i suoi genitori.
 (ADJACENT-Italian)
 Is arrived not Gianni, but are-3PL arrived the his parents
 ‘Gianni has arrived not, but his parents’
 b. **Non** è arrivato Gianni, **ma sono arrivati** i suoi genitori.
 (NON-ADJACENT-Italian)
 Not is arrived Gianni, but are-3Pl arrived the his parents
 ‘Gianni has not arrived, but his parents’

The third contrast is related to a parallelism requirement in both conjuncts: the adjacent order does not tolerate extraneous constituents, and, if accepted at all, they receive a parenthetic intonation. Interestingly, the non-adjacent order has no problem whatsoever when that situation arises.

² I use lines to indicate the alleged process of ellipsis, an issue I return to.

- (10) a. I have invited **not** your brother to the party, **but** your sister ??(to the cinema).
 b. I haven't invited your brother to the party, **but** your sister (to the cinema).
- (11) a. **No** le di los libros a María, **sino** las revistas (a Juan).
 (NON-ADJACENT-Spanish)
 Not CL-him the books to María, but the magazines (to Juan)
 'I didn't give the books to María, but the magazines to Juan'
- b. Le di **no** los libros a María, **sino** las revistas ??(a Juan).
 (ADJACENT-Spanish)
 CL-him not the books to María, but the magazines (to Juan)
 'I gave not the books to María, but the magazines to Juan'

In the adjacent order, plural agreement with two subjects can obtain with slight deviance (cf. (12a)); as far as the non-adjacent order is concerned, it is not possible for the conjoined subject to trigger plural agreement (cf. (12b)). That makes the final asymmetry.³

- (12) a. ??Hablaron con Juan **no sólo** María, **sino también** Laura.
 (ADJACENT-Spanish)
 Talked-3PL to Juan not only María, but also Laura
 'Talked to Juan not only María, but also Laura'
- b. ***No** hablaron con Juan **sólo** María, **sino también** Laura.
 (NON-ADJACENT-Spanish)
 Not talked-3PL to Juan only María, but also Laura
 'Didn't talk to Juan only María, but also Laura'

Once we have revisited the most intriguing asymmetries regarding edge coordination's patterns of (3), we are in a position to offer an analysis that can account for the data. In the next section, I offer the basics of the appealing approach by Zamparelli & Bianchi (2001).

2. Bianchi & Zamparelli's (2001) analysis

In Bianchi & Zamparelli (2001) two different analyses that try to capture the just observed facts are put forth. Let us have a look at the adjacent order first, which is the one in (13):

- (13) a. The assassin killed **not** (*only*) Smith, **but** (*also*) his dog.
 b. El asesino mató **no** (*sólo*) a Smith, **sino** (*también*) a su perro. (Spanish)

The analysis in Bianchi & Zamparelli (2001) goes like this: first of all, they assume some peripheral functional projections in order to derive the desired semantic effects (*pace* Rizzi 1997 and much related work); to be more concrete, they hold that focus particles (in the case at hand, *not only...but also*) are generated as directly attached constituents to the material that acts as the sentence's focus, which in a subsequent derivational step move to the specifier of a Focus Phrase. Furthermore, in the top of the

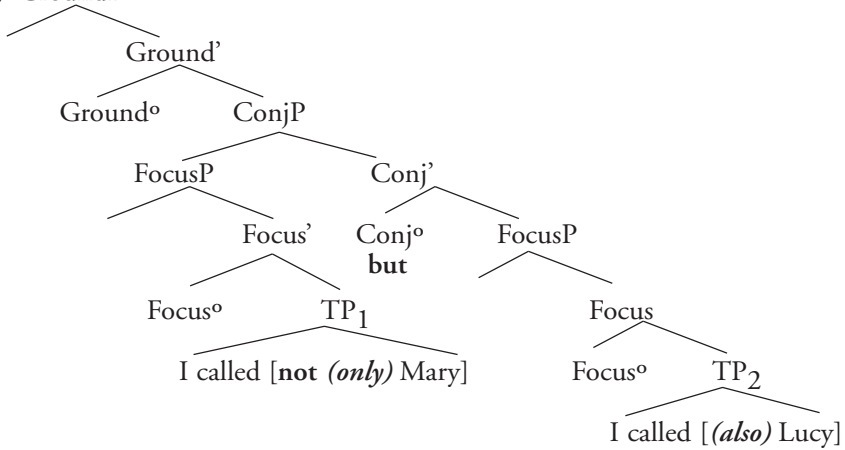
³ Bianchi & Zamparelli (2001) point out a fifth asymmetry dealing with presupposition. I will put it aside for the time being, since I investigate it in work in progress.

structure we find a Ground Phrase, whose specifier is the target for material containing backgrounded information (or the ‘aboutness’, in the sense of Herburger 2000). Bianchi & Zamparelli (2001) defend the idea that there is only one GP per speech act: “the rationale is that this projection should host material which is factored across all conjuncts, becoming background for the whole current speech act.” (p. 5)

The proposed structure, thus, would be as in (15), for a sentence like the one in (14):

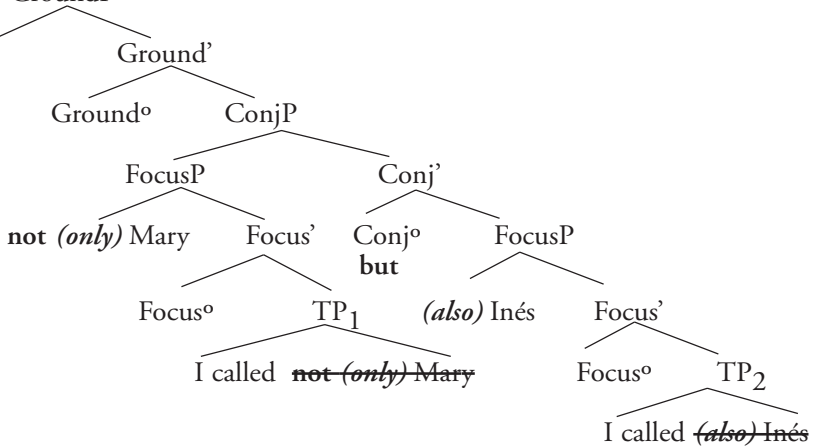
(14) I called **not** (*only*) Mary, **but** (*also*) Lucy.

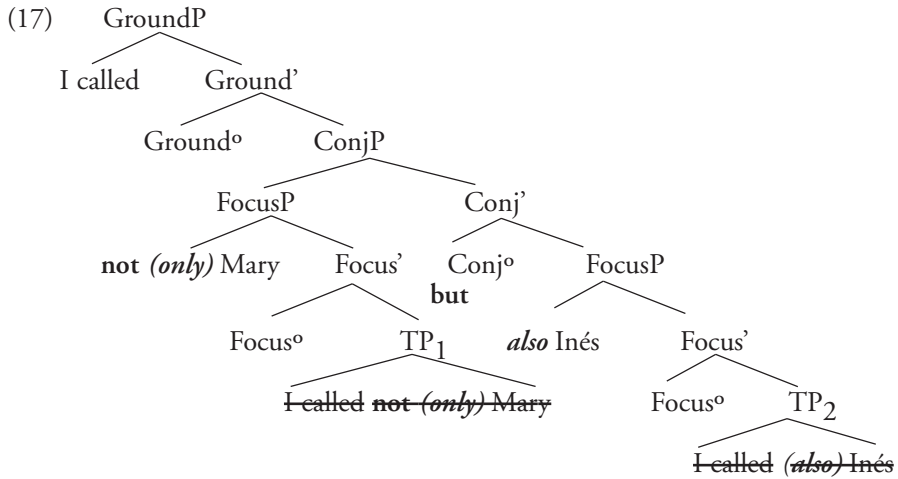
(15) GroundP



As (15) shows, the basic skeleton of the analysis assumes, below the GP, the existence of a Conjunction Phrase (cf. Munn 1993, Progovac 2003, *inter alia*) that takes as its arguments two Focus Phrases, which, in turn, dominate two clauses. In order to derive the adjacent order, two steps are needed: first, the head and the coda (the ‘correlates’, in Bianchi & Zamparelli’s 2001 terms) move to the specifiers of both FocPs (cf. (16)); second, a remnant movement operation of the TPs takes place in an ATB-fashion targeting the [Spec, GP] (cf. (17)).

(16) GroundP



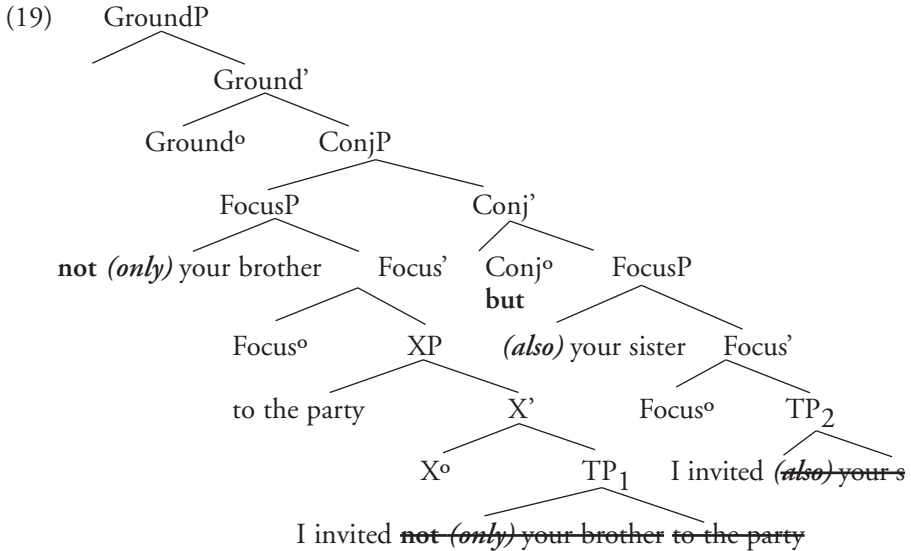


Now, let us see how such a derivation of the adjacent order faces the asymmetries of section 1. Consider first the truncation phenomenon, illustrated in (6); in Bianchi & Zamparelli's (2001) account, it has its source in the fact that the string consisting of the ATB-raised and the first correlate do not correspond to a syntactic constituent.

The agreement facts that constitute the second difference between both orders follow from (17): the verb in the ATB-raised TP would have to spell-out two inconsistent ϕ -features (e.g., singular vs. plural), as we could see in (8) (repeated here as (18)):

- (18) a. ?Ha hablado **no** Juan, **sino** ~~han hablado~~ sus primos. ADJACENT-Spanish)
 Has-3SG talked not Juan, but have-3PL talked his cousins
 'Juan has not talked, but his cousins'
- b. **No** ha hablado Juan, **sino** ~~han hablado~~ sus primos. NON-ADJACENT-Spanish)
 Not has-3SG talked Juan, but have-3PL talked his cousins
 'Juan hasn't talked, but his cousins'

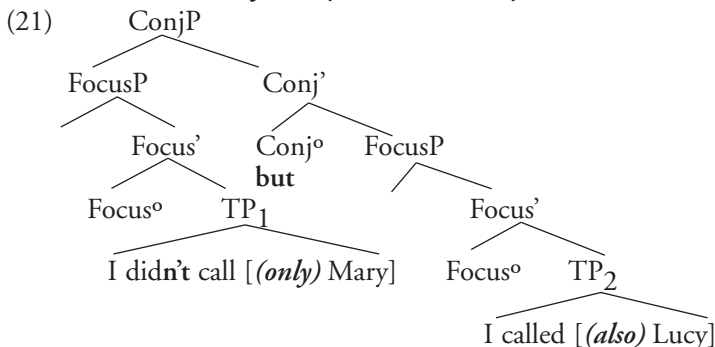
As for the parallelism requirement, it is derived if we assume that introducing any extraneous constituent would invoke an additional 'scrambling' operation out of the TP, which would render the remnant TPs not identical, hence barring the ATB process. Therefore, something along the lines of (19) would have to be at stake:



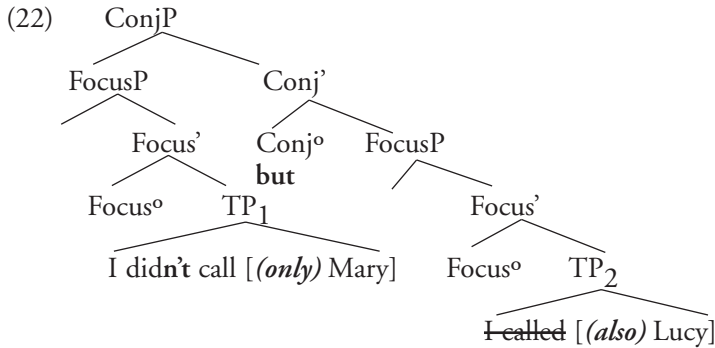
Finally, plural agreement between two conjoined subjects is allowed under this analysis by if the [Spec, TP] position of both correlates shares the same referential index, which predicts that the same inflectional agreeing head will be spelled-out with plural features.

What about the non-adjacent order? Bianchi & Zamparelli (2001) claim that some modifications to the starting structure in (15) are needed so that we can explain the full range of properties. In their implementation, (20) starts its derivational life as in (21):

(20) I didn't call *(only)* Mary, **but** *(also)* Lucy.



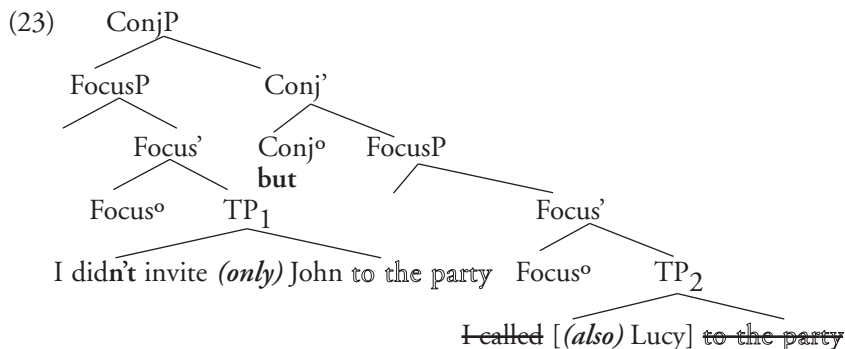
As the reader may have already noted, the two structures differ in non-trivial respects: the non-adjacent order dispenses with the GP, the first of the two correlates is not launched to the specifier of the FocP, and, this time, ellipsis is obtained by some kind of PF process subject to a parallelism requirement, not ATB movement, as indicated in (22):



In Bianchi & Zamparelli's (2001) words:

We remain agnostic as to the exact nature of this ellipsis process. The only property that is crucial for our argument is that this process does not require strict identity of the antecedent IP and the elliptical IP (contrary to the ATB derivation); in particular, it seems to be insensitive to differences in functional features. Bianchi & Zamparelli (2001: 9).

Going back to the asymmetries, given the structure in (21), they can receive an explanation. First, the possibility of truncating the string is a result of the whole structure corresponding to the first correlate (i.e., the first FocP). Second, the ϕ -features mismatch between the two verbs are a direct consequence of the PF ellipsis, which, contrary to ATB movement, does not require strict identity in overt ϕ -features. Third, the appearance of an extraneous constituent showing up between the head and the coda is faced by assuming some kind of 'reconstruction' of an identical counterpart in the elliptical TP, as roughly illustrated in (23):



Finally, plural agreement is not possible in the non-adjacent order due to the fact that the verb that survives the ellipsis process is just the one belonging to the first TP, being just coindexed with the subject of that very TP.

This section has summarized the theoretical assumptions made by Bianchi & Zamparelli (2001) in order to provide an analysis that can account for the interest-

ing asymmetries that the adjacent and non-adjacent orders of edge coordinations exhibit, and, as we have just seen, the results seem to fit with the proposal. However, it is worth wondering whether a more unitary and elegant analysis is possible, and, in this respect, whether some aspects should be looked at more carefully, specially the two different assumed base structures, the ellipsis/ATB processes, and the use of so much functional structure to make the semantics be transparent at Logical Form.

3. A Minimalist Analysis

In the following pages I offer an alternative analysis for edge coordinations dispensing with much of the machinery in Bianchi & Zamparelli (2001). The aspects of their analysis that I will be concentrating on are two: whether ellipsis should be invoked and the corrective nature of such constructions. I will leave the issues concerning focus for the next section.

3.1. Ellipsis

The first aspect of Bianchi & Zamparelli (2001) I am going to dwell on is the ellipsis processes they postulate for edge coordinations.⁴ It is interesting to compare these structures with some that can be taken to imply ellipsis as well; a good example is the coordinated structure in (24a). The question can more generally be stated as follows: do we have to assume ellipsis every time we have a distributive interpretation of an event? Do we have to assume, for instance, that (24a) —in one of its readings— has been reduced by ellipsis, as indicated in (24b)?

- (24) a. Mary and John went to the cinema.
 b. [_{ConjP} [_{CP} Mary ~~went to the cinema~~] [_{Conj'} [_{Conj} and] [_{CP} John went to the cinema]]]

By parity of reasoning, whenever we have two objects, a similar derivation should be assumed.

⁴ The reduction process outlined in Bianchi & Zamparelli (2001) is an extension of Zamparelli (2000), where the same kind of mechanism is used to derive the “distributive coordination” in (i), whose initial structure is as in (ii):

- (i) **Both** John and Mary went to the cinema.
 (ii) [_{XP} **Both** [_{CP} [_{FocP} John went to the cinema] [_C [_C and] [_{FocP} Mary went to the cinema]]]

As Zamparelli (2000) points out:

“The idea I want to pursue to explain this complex pattern is that distributive coordination is the result of the coordination of two full sentences —in most cases, root sentences— which are ‘reduced’ by meaning-preserving syntactic operations. The distributive semantics falls out automatically from this structures without having to stipulate the existence of a special, distributive type of conjunction” (Zamparelli 2000: 9).

Due to space limitations, I cannot fully review Zamparelli (2000) here. I come back to the whole issue in work in progress.

- (25) a. John sent Mary flowers and books.
 b. John [_{ConjP} [_{vP} sent Mary flowers] [_{Conj} and [_{vP} ~~sent Mary~~ books]]]

The point I want to make, at any rate, is whether we have to assume the analysis in (25b) for the structure in (26), which is a *bona fide* edge coordination (cf. (26b)), or else a non-elliptical counterpart can be defended (cf. (26c)):

- (26) a. I didn't buy a book, **but** a magazine.
 b. [_{ConjP} [_{CP} I didn't buy a book] [_{Conj} but [_{CP} ~~I bought~~ a magazine]]]
 (ELLIPSIS)
 c. I didn't buy [_{ConjP} [_{DP} a book] [_{Conj} but [_{DP} a magazine]]]
 (NO ELLIPSIS)

On such cases, one could perfectly assume that the event quantifier is the only element that gets affected by the distributive reading, resorting to a neo-Davidsonian approach along the lines of Herburger (2000) or Beghelli & Stowell (1997), whereby all sentences (and not only those containing stage-level predicates, in the sense of Kratzer 1995) contain an existential quantification over events which can give rise either to collective or distributive interpretations. In particular, following Beghelli & Stowell (1997), one could argue that distributive readings arise by keeping the existential quantifier (which is generated within the *vP*, and later on moved to a particular functional projection in the CP-field: a Share Phrase) under the scope of the subject; if the covert existential quantifier takes wide scope, then distribution fails, and a collective reading obtains instead.

Nevertheless, it seems that in certain circumstances, a process of ellipsis must be postulated, like in the example (27b), where, besides the distributive reading, two adverbs show up: since it is impossible to introduce an adjunct in a simple PP coordination, a supporting verb must have been deleted, as indicated by Nunes (2001):

- (27) a. Eu conversei com o João e a Maria. (Portuguese)
 I talked-1SG with the João and the Maria
 'I talked to João and Maria'
 b. Eu conversei com o João (sábado) e com a Maria (domingo). (Portuguese)
 I talked-1SG with the João (saturday) and with the Maria (sunday)
 'I talked to João (on Saturday) and to Maria (on Sunday)'
 [from Nunes (2001: 339)]

Here I will assume a non elliptical process for edge coordinations, *pace* Beghelli & Stowell (1997). It is important, however, to distinguish edge coordinations proper (cf. (28)) from a very similar —although crucially different— pattern, namely, *verbal ellipsis* (cf. Brucart 1987; 1999) or, according to Depiante (2004), *pseudostripping* (cf. (29)):

- (28) a. Ana vio a María, pero **no** a Susana. (Spanish)
 Ana saw-3SG to María, but not to Susana
 'Ana saw María, not Susana'
 b. Ana no vio a María, pero **sí** a Susana. (Spanish)
 Ana not saw-3SG to María, but yer to Susana
 'Ana didn't see María, but Susana'

- (29) a. Ana vio a María, pero a Susana **no**. (Spanish)
 Ana saw-3SG to María, but to Susana not.
 ‘Ana saw María, but not Susana’
- b. Ana no vio a María, pero a Susana **sí**. (Spanish)
 Ana not saw-3SG to María, but to Susana yes
 Ana didn’t see María, but Susana’

[from Depiante (2004: 62)]

According to Depiante (2004), the structures in (28) should be treated together with *gapping* and *stripping*. As for the examples in (29), they belong to the same group *ellipsis* and *sluicing* do. I will put aside here whether the examples of (29) are real instances of VP (or TP, in Depiante’s (2004) analysis) ellipsis, what I want to discuss in what follows is the exact status of the coordinated structures in (28).

As I just said, Depiante (2004), just like Bianchi & Zamparelli (2001), assimilates edge coordinations to ellipsis, and, more precisely, to *stripping*. On the face of it, Brucart (1999) argues that there is no such process at all in the structures under inspection; in Brucart’s (1999) words:

“The negation that appears in these constructions, which we will call “corrective”, adopting the proposal by Bosque (1984), is not the remnant of an elliptical VP, but a partial negation that only affects the phrase to its right [...] with which it forms a non-sentential syntactic projection. Moreover, the negative constituent acts as a parenthetical adjunct of an element in the main clause with which it holds a polarity contrast relation” (Brucart 1999: § 43.2.3.4.) [my translation]

Depiante (2004) offers three arguments against Brucart’s (1999) proposal: let us quickly review them. The first one has to do with the licensing conditions of the alleged ‘remnants’, and, more precisely, with case a θ -role assignment. The second problem is related to structures such the one in (30), where an anaphoric dependency can be established despite there not being any overt antecedent (arguably, the material undergoing ellipsis contains it):

- (30) Juan **no** tiene un auto, **pero sí** Pedro, y está en muy buenas condiciones. (Spanish)
 Juan not have-3SG a car, but yes Pedro, and be-3SG in very bad conditions
 ‘Juan does not have a car, but Pedro does, and it looks very good’

[from Depiante (2004: 65)]

The third argument has to do with the fact that there is a list of well-known properties of what Depiante (2004) dubs “local ellipsis” (which includes *gapping* and *stripping*) with which edge coordinations pattern: it is restricted to coordinated TPs (cf. (31)), it is locally bounded (cf. (32)), it is sensible to islands (cf. (33)), and it can operate on non-syntactic constituents (cf. (34)).

- (31) a. Susan didn’t read a book although Mary did. Non-local ellipsis
 b. *Susan read a book although not a magazine. Local ellipsis

[from Depiante (2004: 58-59)]

- (32) a. I play tennis every weekend and I think that Susan said that Peter claimed that Max does too. Non-local ellipsis.
 b. *I read a book and I think that Susan said that Peter claimed that Max a magazine. Local ellipsis
 [from Depiante (2004: 59-60)]
- (33) a. John parked his car where Mary did. Non-local ellipsis
 b. *John parked his car where Mary her van. Local ellipsis
 [from Depiante (2004: 60)]
- (34) a. I read a book and [_{TP} Mary did [_{vp} ~~read a book~~ too]] Non-local ellipsis
 b. Peter caught an eel for Mary in the Charles River and [_{TP} John [_{vp} ~~caught a flounder for Mary in the Charles River~~]] Local ellipsis
 [from Depiante (2004: 61)]

Let us address each argument of Depiante (2004). The first one can be dismissed right from the beginning, since, if real, it could also be raised in many more cases, as the ones in (35), for which it is far from obvious that a process of ellipsis should be invoked:

- (35) a. Mary and John are brothers.
 b. I called Mary and John.

Note that the coordinated DPs do not trigger a distributive reading, but, what really matters here is whether they (actually, one of them) can or cannot receive case and θ -role. If there were just one verb and Depiante (2004) were correct, the derivation of both examples in (35) should *crash*, contrary to fact. A reasonable move would be to assume that both DPs receive the same case (Nominative and Accusative) and θ -role.

As for Depiante's (2004) second argument, it should be noticed that ellipsis does exist here, but because of the conjunction that is being used: Spanish "pero" can only head clauses (that is, propositional entities), contrary to English "but", which can correspond not only to "pero", but also to "sino" (and "excepto"), which is the one we are interested in here. In this vein, note that (36a) must be translated in Spanish as (36b), not as (36c) or (36d):

- (36) a. There is not one people, **but** two.
 b. *Hay no una persona, **pero** dos. (Spanish)
 c. Hay no una persona, **sino** dos. (Spanish)
 d. Vinieron todos excepto Juan. (Spanish)
 CAME-3PL all except Juan
 'All (of them) came but Juan'

When clausal structures are considered, "pero" is fine:

- (37) a. John will come, **but** he says he is tired.
 b. Juan vendrá, **pero** dice que está cansado. (Spanish)

More importantly for my purposes, note that the kind of anaphoric dependency pointed out in Depiante (2004) does not arise in the case of true edge coordinations (which, to repeat, involve the coordinating conjunction "sino", not "pero"):

- (38) *Juan **no** tiene un auto, **sino** Pedro, y está en muy buenas condiciones. (Spanish)

Juan not have-3SG a car, but Pedro, and be-3SG in very bad conditions
 ‘Juan does not have a car, but Pedro, and it looks very good’

Moving on to the third argument, I think that it just shows that local and non-local ellipses are different, but nothing deeper.

What I would like to defend here, much in the sense of both Bosque (1984) and Brucart (1987; 1999) is that sentences like (39b) do not involve ellipsis at all, but just a conjunction phrase that contains a “corrective negation/affirmation”; (39a), however apparently identical, does involve ellipsis.

- (39) a. Juan ha cantado, **pero** Pedro **no**. (Spanish)
 Juan have-3SG sung, but Pedro not
 ‘Juan has sung, but Pedro has not’
 b. Juan ha cantado, (*y*) **no** Pedro. (Spanish)
 Juan have-3SG sung, but not Pedro
 ‘Juan has sung, and not Pedro’

The next data, taken from Brucart (1999), provide evidence supporting a non-elliptical analysis of edge coordinations. (40) illustrates that true ellipsis, but not edge coordinations, requires anaphoric dependencies, making it impossible for the elliptical chunk to precede the structure ellipsis relies on (cf. (40c)):

- (40) a. Juan, (*y*) **no** Pedro, es el verdadero asesino. Edge Coordination (Spanish)
 Juan, (and) not Pedro, is the true guilty
 ‘Juan, (and) not Pedro, is the true guilty’
 b. Juan trabaja los lunes, **pero** Pedro **no** ~~trabaja los lunes~~. True ellipsis (Spanish)
 Juan work-3SG the Mondays, but Pedro not
 ‘Juan works on Monday, but Pedro does not’
 c. *Juan, (*y*) Pedro **no** ~~es el verdadero asesino~~, es el verdadero asesino. True ellipsis (Spanish)
 Juan, (and) Pedro not is the true murderer, is the true murderer
 ‘Juan, and Pedro is not, is the true murderer’

In (41), we can see that aspectual adverbs like “todavía” (Eng. *yet*) are allowed in edge coordinations, but not in *bona fide* elliptical contexts:

- (41) a. Juan ha llamado a Inés, y Pedro todavía **no** ~~ha llamado a Inés~~. True ellipsis (Spanish)
 Juan have-3SG called to Inés, and Pedro yet not
 ‘Juan has called Inés, and Pedro hasn’t (done it) yet.’
 b. Juan ha llamado a Inés, (*y*) **no** Miguel. Edge coordination (Spanish)
 Juan have-3SG called to Inés, (and) not Miguel
 ‘Juan has called Inés, (and) not Miguel’
 c. *Juan ha llamado a Inés, (*y*) **no** todavía César. Edge coordination (Spanish)
 Juan have-3SG called to Inés, (and) not yet César
 ‘Juan has called Inés, (and) not yet César’

Another argument against ellipsis is offered in (42), which proves that only edge coordinations allow negation when the second correlate is an instance of what I am calling “corrective negation”; ellipsis does not.

- (42) a. ?Inés **no** hizo los deberes, **no** Ana. Edge coordination (Spanish)
 Inés not make-PAST-3SG the homework, not Ana
 ‘Inés didn’t make her homework, not Ana’
- b. Fue Inés, (y) **no** María, la que **no** hizo los deberes.
 Edge coordination (Spanish)
 Be-PAST-3SG Inés, (and) not María, the that not make-PAST-3SG the homework
 ‘It was Inés, (and) not María, the one that didn’t make her homework’
- c. *Luisa **no** hizo los deberes, y María ~~no hizo los deberes~~.
 Ellipsis (Spanish)
 Luisa not make-PAST-3SG the homework, and María not
 ‘Luisa didn’t make her homework, and María didn’t’

For the punch line, witness that in (43), edge coordinations cannot be followed by another clause, unless it is an appositive relative, a fact that supports the non-clausal status of these structures:

- (43) a. María suspenderá un examen, (y) **no** Luis, pero ya lo recuperará en septiembre.
 Edge coordination
 María fail-FUT-3SG an exam, (and) not Luis, but already CL-it pass-FUT-3SG in September
 ‘María will not pass an exam, (and) not Luis, but he will pass it in September’
- b. María suspenderá un examen, (y) **no** Luis, quien ya lo recuperará en septiembre. Edge coordination
 María fail-FUT-3SG an exam, (and) not Luis, who already CL-it pass-FUT-3SG in September
 ‘María will not pass and exam, (and) not Luis, who will pass it in September’
- c. *María suspenderá un examen, y Luis ~~no suspenderá un examen~~, pero ya lo recuperará en septiembre. Ellipsis
 María fail-FUT-3SG an exam, and Luis not, but already CL-it pass-FUT-3SG in September
 ‘María will fail in an exam, and Luis won’t, but he will pass it in September’

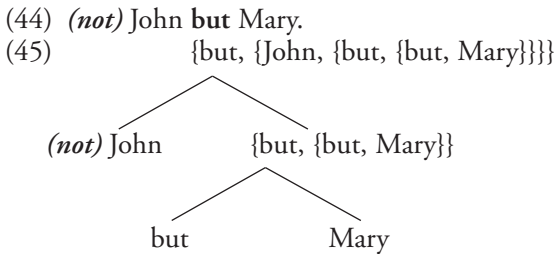
3.2. A Polarity Correction

At this point, we need to clarify what it means to be a “corrective negation/affirmation”. Technically, I argue that it involves a conjunction phrase whose head may or may

not be overtly filled, and a species of ‘contrastive polarity’ checking. When considering structures of a similar type, Herburger (2000) makes the following reasoning:⁵

Finally [...] this type of contour always suggests that there is a “polarity-reversing” continuation. If the sentence is negated, having a bound reading, the fall-rise contour signals a positive continuation, along the lines of “not X, but Y”; and if the sentence is affirmative, a fall-rise contour signals “X, but not Y.” Fall-rise contour thus seems like a tonal way of saying *but*. (Herburger 2000: 54)

Before concentrating on the role played by focus in edge coordinations, we have to address two issues: the structure to be adopted in edge coordinations and the nature of the polarity checking. As for the structure, I will adopt the mainstream analysis for coordination (cf. Munn 1993, Kayne 1994, Larson 1991, Progovac 2003, *inter alia*), in which the coordinating conjunction heads a syntactic projection that takes the coda as its ‘complement’ (i.e., its sister) and the other coordinated elements as its specifiers, a position that is sound under the *Bare Phrase Structure* proposal outlined by Chomsky (1995). So, a string like (44) should be represented as in (45):



Edge coordinations can show not only two arguments, but actually multiple “correlates”, as in (46). The good news of the analysis of (45) is that this case of ‘conjunction doubling’ can be seen as an agreement mark, following the treatment of Spanish “ni” (Eng. *neither*) put forward by Bosque (1994):

⁵ In Herburger (2000: ch. 2, *Appendix*), it is pointed out that the interaction between focus and negation can provoke a *bound reading*, whereby negation only affects focus; in the *free reading* negation affects the verb, as indicated in (i) and (ii):

- (i) Sascha didn’t visit MONTMARTRE.
- (ii) “What Sascha visited wasn’t Montmartre.” [bound reading]
- (iii) “What Sascha didn’t visit was Montmartre.” [free reading]

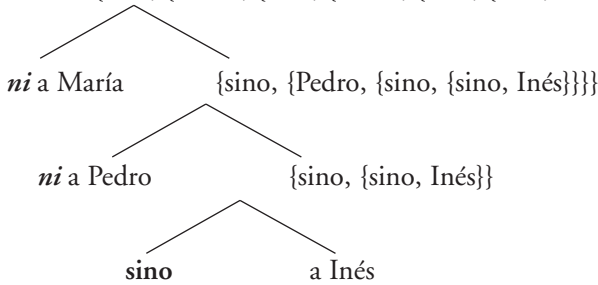
[from Herburger (2000: 29)]

It is interesting to note that edge coordinations seems to provide a test to differentiate free and bound readings. In (iv), continuing with *but*... forces the *bound reading*; continuing it with *and not*... or *but not*... forces the *free reading*. That is, if the coda is negative, *free reading* emerges, and vice-versa.

- (iv) Sascha didn’t visit MONTMARTRE, **but** THE LOUVRE. [bound reading]
- (v) Sascha didn’t visit MONTMARTRE, **and not** the LOUVRE. (He DID in fact visit the Louvre) [free reading]

[from Herburger (2000: 30)]

- (46) Luis **no** vio a Juan, **ni** a María, **ni** a Pedro, **sino** a Inés. (Spanish)
 Luis not see-PAST-3SG to Juan, neither to a María, neither to Pedro, but
 to Inés
 ‘Luis didn’t see Juan, nor María, nor Pedro, but Inés’
- (47) {sino, {María, {sino, {Pedro, {sino, {sino, Inés}}}}}



Bosque (1994) explores cases like those in (48), and argues that the doubling of “ni” is an agreement marker between the specifier and the Conj head in order for the whole projection to be properly identified as a NPI. If that situation fails, preverbal ConjPs headed by “ni” are ruled out, just like any other non negative element merging with Σ .⁶

- (48) a. No cantaron (ni) Juan ni Pedro. (Spanish)
 Not sing-PAST (neither) Juan neither Pedro
 ‘Neither Juan nor Pedro sang’
- b. *(Ni) Juan ni Pedro cantaron. (Spanish)
 (Neither) Juan nor Pedro sing-PAST-3PL
 ‘Neither Juan nor Pedro sung’

[from Bosque (1994: 191)]

As Bosque (1994) points out:

It is reasonable to think that what happens here is that preverbal conjunctive phrases cannot be recognized as NPIs because *ni*’s position does not allow to identify as such the whole constituent, but only the conjunction’s complement. We can thus suppose that for the specifier of a conjunctive phrase with *ni* to be licensed as negative it is necessary that it agrees with its head. In order to get that, it is necessary for the specifier to contain either a negative quantifier, like in *Ningún libro ni ningún artículo le han gustado* [Eng. He didn’t like any book or magazine], or else a syntactic agreement mark (the first *ni* in [(48)]) with the head of the ConjP, or, finally, what seems coherent, a negative operator properly governed containing the appropriate agreement features (Bosque 1994: 191) [my translation].

⁶ In his analysis, Bosque (1994) argues against the deep structure of (ii) for (i), which is coherent with what we have been assuming all along:

- (i) No veo a tu padre ni a tu madre. (Spanish)
 Not see-1SG to your father nor to your mother
 ‘I don’t see your father nor your mother’
- (ii) No [veo a tu madre] ni [~~veo~~ a tu madre]. (Spanish)

[from Bosque (1994: 193)]

Now, recall that edge coordinations can come in different disguises, depending on whether the first sentence contains negation or not, and, at the same time—and this is what crucially differentiates *adjacent* and *non-adjacent* orders—, this negation can be a normal case of clausal negation (which we assume to be placed in Laka's 1990 Σ ; cf. (49a)) or an instance of the so-called “constituent negation” (CN) (cf. Klima 1964, Lasnik 1972 and Horn 1989; cf. (49b)):

- (49) a. I didn't read the books, but the magazines. Clausal negation
 b. I read **not** the books, but the magazines. Constituent negation

The same pattern can arise in other environments, like (50), where the standard view is that negation directly merges with the QP “pocos” (Eng. *few people*), as Ricardo Etxepare (p.c.) has informed me:

- (50) a. [Pocos **no** han venido a la fiesta]. Clausal negation (Spanish)
 Few not have-3PL come to the party
 ‘Few people have come to the party’
 b. [[**No** pocos] han venido a la fiesta]. Constituent negation (Spanish)
 Not few have-3PL come to the party
 ‘Not few people have come to the party’

The non-easy part of the story is that not any kind of XP (nor any position, for that matter) allows CN. So, for instance, postverbal QP rejects CN:

- (51) a. **No** todos han venido. Preverbal CN (Spanish)
 Not all have-3PL come
 ‘Not everyone has come’
 b. *Han venido **no** todos. Postverbal CN (Spanish)
 Have-3PL come not all
 ‘Not everyone has come’

Obviously, any attempt to clarify the *adjacent* order of edge coordinations must have something to say about CN. In Etxepare (in progress), it is argued that CN does not form a constituent when it appears with strong QQ, the alleged adjacency being a by-product of “association with focus”, in the sense of Rooth (1985) and Herburger (2000): negation selects a focus projection which contains the QP, having propositional scope and forcing the implication that some element in the set of propositional alternatives induced by the focus is true. In other words, in (52a) we assert that John did drink something, although, whatever it is, it was not beer; in Rooth's (1985) terms, its ‘focus semantic value’ $[[\phi]]^F$ would be the one in (52b), which spells out the set of alternative values for the focus variable:

- (52) a. John drank **not** [beer]_{FOCUS}.
 b. $[[\phi]]^F = \{[[\text{John did not drink [beer]}^F]], [[\text{John did not drink [vodka]}^F]],$
 $[[\text{John did not drink [wine]}^F]], [[\text{John did not drink [bourbon]}^F]],$
 $[[\text{John did not drink [cognac]}^F]], \dots\}$

The structure proposed by Etxepare (in progress) to obtain (50b), repeated as (53), is (54):

- (53) [No pocos] han venido a la fiesta. (Spanish)
 Not few have-3PL come to the party
 ‘Not few people have come to the party’

(54) [_{ΣP} no [_{TP} ... [_{XP} pocos [_X X° [. . .]]]]]

Restricting the range of data to DPs, it is obvious that, whatever the order, the result is fully out (I use Spanish here, since this language allows postverbal subject DPs):

- (55) a. *Han llegado [no los niños]. (Spanish)
 Have-3PL arrived not the kids
 ‘There have arrived not the kids’
 b. *[No los niños] han llegado. (Spanish)
 Not the kids have-3PL arrived
 ‘There have arrived not the kids’

A plausible source of the ungrammaticality of (55) is the definite article, which, although can receive a strong Q analysis, as shown by its incompatibility in existential contexts (cf. (56b)), is able to appear in exemplary or presentational ones (cf. (56a)), as noted by Hornstein & Uriagereka (2002).

- (56) a. What can we use for a prop? There’s always the table, ... Presentational context
 b. #There’s the table you got me for a prop on stage. Existential context
 [from Hornstein & Uriagereka (2002: 117)]

Hornstein & Uriagereka (2002) account for those facts by arguing that definite descriptions (including proper names) are intrinsically presuppositional. If this is all on track, it could be the case that presuppositional elements (i.e., definite descriptions), despite behaving like strong QQ, cannot associate with the alleged cases of CN we have been considering so far; but one still wonders why. At present I have no principled explanation for why NC cannot associate with definite descriptions, but only with *bona fide* quantifiers (e.g., *all*, *few*, *many*, etc.).⁷

The bottom line, anyway, is why should a derivation like (57a) be bad?

- (57) a. *Ha hablado no el profesor. (Spanish)
 Have-3SG talked not the teacher
 ‘The teacher has not talked’
 b. [_{CP} [_{ΣP} no [_{TP} ha hablado el profesor]]]

Going back to Hornstein & Uriagereka’s (2002) point: if definite descriptions are presuppositional, it would make perfect sense for them to resist being in focal structures. As a result, their natural *locus* should be the restrictive clause of the existential quantifier (or somewhere outside its scope, for that matter), as indicated in (58b).⁸

⁷ One possibility would be to blame them for being referential (i.e., presuppositional): things being so, these elements must occupy a wide-scope position at LF, one crucially *c*-commanding negation (in Beghelli & Stowell’s 1997 system, they would be launched to [Spec, RefP], the highest position in the CP), and this would conflict with CN.

⁸ Following Irurtzun (2003), I have underlined the focus to make it more salient.

The problem with this approach is that the proper noun is actually the focus (and, more generally, definite descriptions can constitute the focus).

- (58) a. A boy saw John.
 b. [*the* x: John x] [\exists e: C (e) & see (e) & [\exists y: boy y] Experiencer (e,y)]
Theme (e,x)

Note further that the problem cannot be solved as in Etxepare (in progress), since, even if CN involves movement to Uriagereka's (1995b) [Spec, FP], hence requiring a preverbal position, this strategy does not help much in the case of definite descriptions (cf. (59b)), although it does with quantifiers (cf. (59a)):

- (59) a. [**No** todos] han venido. (Spanish)
 Not all have-3PL come
 'Not everyone has come'
 b. ***[No los niños]** han venido. (Spanish)
 Not the kids have-3PL come
 'Not the kids have come'

An additional drawback is that focused elements tend to be postverbal in Spanish (they are preverbal just in 'contrastive focus' environments).

My answer to the observed facts runs as follows: only proportional or universal QQ can associate with CN so that no negative interpretation is obtained. In plain terms, whenever negation combines with universal and proportional QQ, the result is non-negative, as the paraphrase in (60b) indicates:

- (60) a. [[**Not** many] students] came.
 b. [Few students] came. (not many = few)

Existential quantifiers do get a negative interpretation; so, (61a) has the rough meaning of (61b); interestingly, existential quantifiers do not allow CN either, as (61c) shows:

- (61) a. The president didn't answer any question.
 b. There is **no** question such that the president answered it.
 c. ***[Not some]** question is annoying.

An important trait of edge coordinations is that, when apparent CN shows up (in the adjacent order), the corrective coda can save the sequence:

- (62) a. John drank not beer *(, **but** tequila).
 b. John drank beer (, (*and*) **not** tequila).

I assume that the facts in (62) are to be captured by a polarity checking operation between the correlates, as is clear in (63), where we can see that the same polar value in these two constituents yields ungrammaticality:

- (63) a. *John drank **not** beer, (*and*) **not** tequila.
 b. *John drank beer, **but** tequila.

The checking I am assuming, would, then, be a local one, but locality does not obtain in many cases, given that the corrective coda can appear in different positions:

- (64) a. John, [**not** Peter], said that.
 b. John said that, [**not** Peter].

To complicate matters even more, the head of the construction (the specifier of the ConjP) can dispense with the appropriate polar element when it appears in Σ , as happens in (65):

- (65) a. Mary **didn't** call his brother, **but** Peter.
 b. Mary called his brother, **not** Peter.

The first problem goes away if some process of direct generation and subsequent stranding is assumed (cf. Boeckx 2003). Regarding the second one, it could undergo the same fate if we take the specifiers of ConjPs to be able to carry the polar null operators, much in the lines of Brucart's (1995) analysis of Spanish NPIs:

- (66) a. Juan no leyó [**OP**_{NEG} libro alguno] (Spanish)
 Juan not read-PAST-3SG book any
 'Juan didn't read any book'
 b. Juan no leyó [**ningún** libro] (Spanish)
 Juan not read-PAST-3SG [any book]
 'Juan didn't read any book'

Consequently, what we would have is as depicted in (67):

- (67) a. John called [**OP**_{POS} Mary, **not** John].
 b. John **didn't** call [**OP**_{NEG} Mary, **but** John].

In this section I have addressed the issue of ellipsis in edge coordinations. I have argued, *contra* Bianchi & Zamparelli (2001) and Depiante (2004), that edge coordinations do not involve ellipsis, but just a process of corrective negation/affirmation that has no propositional nature. The process of correction that these structures involve, however, can be of different types, requiring the coda to be either positive or negative, as noted by Herburger (2000). The trickier part of the analysis is related to some apparent cases of constituent negation that crucially bear on the *adjacent* order. I have assumed that these cases involve association with focus: negation generates in Σ —not directly merged with the constituent at hand—, and then movement operations apply to obtain the final word order. It is important to highlight that the analysis I have put forth does not explain why (68a) is fine and (68b) out:

- (68) a. Mary **didn't** eat the peanuts.
 b. *Mary ate **not** the peanuts.

In Bianchi & Zamparelli's (2001) analysis, (68) follows from the fact that negation and direct object form no constituent; but, even if so, that does not explain why the continuation is needed. Actually, these authors add the following piece to the puzzle:

Put differently, we have to account for the intrinsic "binary" nature of the adjacent order. At present it is not entirely clear how this constraint should be best captured, and to what extent it can follow from other modules of the grammar. The generalization we need to express is that (i) "edge coordinations" always trigger overt raising of the first correlate to a Focus position, and (ii) once an opera-

tor appears in this position a second correlate must also be present” (Bianchi & Zamparelli 2001: 6).

I have nothing specially deep to add to this quote, apart from noting that (68b) could be explained if that instance of negation, not being in Σ (assume, that in this case, “not” is generated as an agreement mark in [Spec, ConjP], just like Bosque’s 1994 treatment of Spanish “ni”), can only be licensed by a corrective ConjP. Obviously, this raises many questions, and many more so if we restrict ourselves to the streamlined operations within the Minimalist Program. Being in its natural place (i.e., Σ), (68a) poses no problems; now, if negation in (68b) is not in Σ , then it would be logical to expect some additional mechanism to locally license it, namely, a corrective continuation. Note that the same problem seems to be at stake in other doubling structures involving conjunctions, as (69) suggests:

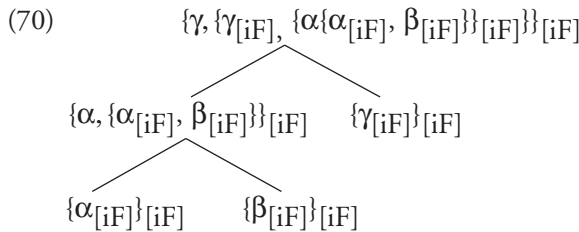
- (69) a. I will go either to Rome *(or to Berlin)
- b. I want both the milk *(and the biscuits)

4. The Role of Focus

So far, nothing has been said about the focal nature of edge coordinations (cf. den Dikken 2003, Han & Romero 2004, Hendriks 2001, Herburger 2000, *inter alia*). I assume here Irurtzun’s 2003 analysis of focus, whereby focus features (i.e., [iF]) are a kind of formal (and interpretable) feature that are assigned in the *Numeration* (cf. Chomsky 2000), as detailed in (70b), assuming that *Mary* is the focus:

- (70) a. I love Mary.
- b. {I, love, *v*, T, C, Mary_{[iF]}}}

I also follow Irurtzun (2003) in taking focus projection to preserve the *command units* created by the monotonic application of *Merge*:



As far as edge coordinations go, the analysis for them would be as in (71):

- (71) a. I love Mary, not Susan.
- b. {I, love, *v*, T, C, Mary_[iF], not, Susan_[iF]}

There is one remarkable difference between the adjacent and the non-adjacent order: in the latter no focus association (Herburger’s 2000 *bound reading*) is necessary, while it is in the former. So, in a sentence like (72a), any of the elements in the c-command domain can be the focus, but there is no need for that; in (73), only

the constituent(s) following the negation is the focus. Note, for one thing, that in the adjacent order, negation has to negate the element that immediately precedes; that is to say, we cannot have long-distance focus marking (cf. (73c) vs. (73d)):

- (72) a. I didn't give_[iF] the books to John, (**but** take_[iF] them from him).
 b. I didn't give the_[iF] books_[iF] to John, (**but** the_[iF] magazines_[iF]).
 c. I didn't give the books to_[iF] John_[iF], (**but** to_[iF] Mary_[iF]).
 d. I didn't give the books to John; I called him.
- (73) a. I gave **not** the_[iF] books_[iF] to John, (**but** the_[iF] magazines_[iF]).
 b. I gave **not** the_[iF] books_[iF] to_[iF] John_[iF], (**but** the_[iF] magazines_[iF] to_[iF] Mary_[iF]).
 c. *I gave **not** the books to_[iF] John_[iF], (**but** to_[iF] Mary_[iF]).
 d. I gave the books **not** to_[iF] John_[iF], (**but** to_[iF] Mary_[iF]).

5. Back to the asymmetries

We have arrived at the critical point of the argument. In this section I would like to lay out how the analysis I have put forward can account for the data of Bianchi & Zamparelli (2001). As we saw, their analysis can explain the noted asymmetries, but in so doing, they increase the number of technical assumptions: functional projections, ellipsis, ATB movement, etc.

Truncation, as I have just said (cf. section 3), could be the consequence of not obtaining the right licensing mechanism: if “not” does not head its projection in these cases (which does not force us to assume that it is CN either, but just some sort of agreement mark), then it is not unlikely that it must undergo a special kind of local licensing, which I take to be incarnated by the ConjP. In fact, remember that this mechanism would be working in other conjunction doubling structures.

The facts about identity, which are explained in Bianchi & Zamparelli (2001) by means of ATB-movement (for it to take place, the remnants must be identical). Note first, that this effect does not yield total ungrammaticality, contrary to truncation. Under the focus analysis I am assuming, the key to the problem has to do with the very nature of these structures: they focus some constituents, and then the ‘alternatives’ that must occupy the focus variable that are offered in the coda. Recall also that the adjacent order works in a rather peculiar way: in its more neutral reading, it marks the whole subsequent string as the focus (it can also mark the first constituent as the focus; cf. (73a)), which is tantamount to saying that the exact number of elements will be needed to satisfy the alternative offer.

The third and fourth asymmetries had to do with agreement: first, only the non-adjacent order allowed for the two verbs (assuming ellipsis) to be different. The relevant data were in (8), which is repeated here as (74):

- (74) a. ?Ha hablado **no** Juan, **sino** ~~han hablado~~ sus primos. ADJACENT (Spanish)
 Has-3SG talked not Juan, but have-3PL talked his cousins
 ‘Has talked not Juan, but his cousins’

- b. No ha hablado Juan, ~~sino han hablado~~ sus primos.
NON-ADJACENT (Spanish)

Not has-3SG talked Juan, but have-3PL talked his cousins
'Hasn't talked Juan, but his cousins'

Second, only the adjacent order allowed two coordinated subjects to trigger plural agreement in the verb, as indicated in (75):

- (75) a. ?Hablaron con Juan **no** (*sólo*) María, **sino** (*también*) Laura.
ADJACENT (Spanish)

Talked-3PL to Juan not only María, but also Laura
'Talked to Juan not only María, but also Laura'

- b. *No hablaron con Juan (*sólo*) María, **sino** (*también*) Laura.
NON-ADJACENT (Spanish)

Not talked-3PL to Juan only María, but also Laura
'Didn't talk to Juan only María, but also Laura'

I argue that both facts can receive a natural explanation if, in the non-adjacent order, the ConjP enters *Agree* as a more compact unit; what does that mean? I will take assume that this structure behaves just like partitive phrases do in languages like Spanish, as noted by Brucart (1997). To be precise, Brucart (1997) shows that partitive phrases containing differently inflected DPs can trigger either singular or plural agreement (which has interpretive consequences, as Brucart 1997 notes, a matter I put aside here):

- (76) a. La mayoría de los estudiantes {aprobó/aprobaron}. (Spanish)

The most of the students-3PL {passed-3SG/passed-3PL}
"Most of the students passed"

- b. El diez por ciento de los soldados {regresó/regresaron}. (Spanish)

The ten per cent of the soldiers {came-back-3SG/came-back-3PL}
"The 10% of the soldiers came back"

The logic I am suggesting is that, whenever the adjacent order obtains, the DPs count as a complex unit whose ϕ -features can be counted.

6. Concluding remarks

The aim of this paper has been to defend the idea that the structures in (77), labelled "edge coordinations" Bianchi & Zamparelli (2001), do not invoke a complex derivation involving ellipsis:

- (77) a. John, (*and*) **not** Peter, accepted the job.
b. I didn't drink wine, **but** beer.

I have tried to show, following Brucart (1999), that all we need is a Conjunction Phrase and a polarity checking between its two arguments. The proposal leaves, nevertheless, many questions without a principled explanation (what is the correct analysis of CN in Spanish, what is the source for the binary nature of these and

other structures, which is the exact status of the polarity checking, etc.), but I turn to them in work in progress.

References

- Beghelli, F. & T. Stowell, 1997, "Distributivity and Negation: The Syntax of *Each* and *Every*", in A. Szabolcsi (ed.), *Ways of Scope Taking*, Dordrecht/Boston/London: Kluwer, 71-107.
- Belletti, A., 2002, "Aspects of the Low IP Area", ms. U.Siena.
- Bianchi, V. & R. Zamparelli, 2001, "Edge Coordinations", ms. SNS di Pisa/Università di Bergamo. [Printed in D. Adger, C. De Cat & G. Tsoulas (eds.), 2004, *Peripheries: Syntactic edges and their effects*, Dordrecht: Kluwer].
- Boeckx, C., 2003, *Islands and Chains. Resumption as stranding*, Amsterdam: John Benjamins.
- Bosque, I., 1984, "Negación y elipsis", *Estudios de Lingüística* (Alicante), 2, 171-199.
- , 1994, "La negación y el principio de las categorías vacías", in Demonte (ed.), 167-99.
- , & P. J. Masullo, 1998, "On Verbal Quantification in Spanish", in O. Fullana & F. Roca (eds.), *Studies on the Syntax of Central Romance Languages*, Girona: U. Girona, 9-64.
- , & V. Demonte (eds.), 1999, *Gramática descriptiva de la lengua española*. Madrid: Espasa.
- Brucart, J. M., 1987, *La elisión sintáctica en español*, Doctoral dissertation, UAB.
- , 1995, "Sobre *más* aditivo y los términos de polaridad negativa en español", ms. UAB.
- , 1997, "Concordancia *ad sensum* y partitividad en español", in M. Almeida & J. Dorta (eds.), *Contribuciones al estudio de la lingüística hispánica. Homenaje al profesor Ramón Trujillo*, Tenerife: Montesinos. Vol. 1, 157-83.
- , 1999, "La elipsis", in Bosque & Demonte (eds.), 2787-863.
- Camacho, J., 1999, "La coordinación", in Bosque & Demonte (eds.): 2635-2694.
- Chomsky, N., 1970, "Remarks on Nominalization", in R. A. Jacobs & P. S. Rosebaum (eds.), *Readings in English Transformational Grammar*, Washington: Georgetown U.P., 184-221.
- , 1986, *Barriers*, Cambridge (Mass.): The MIT Press.
- , 1995, "Bare Phrase Structure", in H. Campos & P. Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory: Studies in Honor of Carlos P. Otero*, Washington: Georgetown U.P., 51-109.
- , 2000, "Minimalist Inquiries: The Framework", in R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step. Essays on Minimalist Syntax in Honour of Howard Lasnik*, Cambridge (Mass.): The MIT Press, 89-155.
- Cinque, G., 1990, *Types of A'-Dependencies*, Cambridge (Mass.): The MIT Press.
- Demonte, V. (ed.), 1997, *Gramática española*. México: Publicaciones de la *Nueva Revista de Filología Hispánica*.
- Den Dikken, M., 2003, "The Syntax of *either... or... one more*", ms. CUNY.
- Depiante, M., 2004, "Dos casos de elipsis con partícula de polaridad en español. Evidencia a favor de una visión no uniforme de la elipsis", *Rasal* 1, 53-69.

- Etxepare, R. (in progress), "On some Cases of Constituent Negation in Spanish", ms. CNRS (IKER).
- Fox, D., 2000, *Economy and Semantic Interpretation*, Cambridge (Mass.): The MIT Press.
- Gallego, Á. J. (in progress), "A Derivational Approach to Correlative Structures", ms. UAB.
- Han, Chung-hye & M. Romero, 2004, "Disjunction, Focus, and Scope", *LI* 35.2, 179-217.
- Herburger, E., 2000, *What Counts: Focus and Quantification*, Cambridge (Mass.): The MIT Press.
- , 2003, "A note on Spanish *ni siquiera*, *even*, and the analysis of NPIs", *Probus* 15, 237-56.
- Horn, L., 1989, *A Natural History of Negation*, University of Chicago Press: Chicago.
- Hornstein, N. & S. Franks, 2000, "Across-The-Board Movement and LF", *Syntax* 3.2, 107-28.
- , & J. Nunes, 2002, "On Asymmetries between parasitic gap and Across-The-Board constructions", *Syntax* 5.1, 26-54.
- , & J. Uriagereka, 2002, "Reprojections", in S. Epstein & T. Seely (eds.), *Derivation and Explanation in the Minimalist Program*, Malden (Mass.): Blackwell, 106-32.
- Irurtzun, A., 2003, "The Locus of [+F]ocus", ms. EHU/UPV.
- Klima, E., 1964, "Negation in English" in J. Fodor and J. Katz (eds.), *The Structure of Language*, Englewood Cliffs (New Jersey): Prentice-Hall.
- Kratzer, A., 1995, "Stage-Level and Individual-Level Predicates", in G. Pelletier (ed.), *The Generic Book*, Chicago: University of Chicago Press, 125-75.
- Laka, I., 1990, *Negation in Syntax: On the Nature of Functional Categories and Projections*, Doctoral dissertation, MIT.
- Larson, R., 1991, "The Projection of DP and DegP", ms. SUNY - Stony Brook.
- Lasnik, H., 1972, *Analyses of Negation in English*, Doctoral Dissertation, MIT.
- López, L., 1999, "The Syntax of Contrastive Ellipsis: Evidence from VP-Ellipsis", in J. Gutiérrez-Rexach & F. Martínez Gil (eds.), *Advances in Hispanic Linguistics*, Somerville (Mass.): Cascadilla Press, 412-27.
- Lorenzo, G., 1999, "Negación sin NegP", *Verba* 26, 193-223.
- Merchant, J., 2001, *The syntax of silence: Sluicing, islands, and the theory of ellipsis*, Oxford: Oxford U.P.
- , 2003, "Fragments and Ellipsis", ms. University of Chicago.
- Munn, A., 1993, *Topics in the syntax and semantics of coordinate structures*, Doctoral dissertation, University of Maryland at College Park.
- , 1998, "ATB Movement without identity", in J. Austin & A. Lawson (eds.), *Proceedings of ESCOL 97*.
- Neale, S., 1990, *Descriptions*, Cambridge (Mass.): The MIT Press.
- Nunes, J., 2001, "Sideward Movement", *LI* 32.2, 304-44.
- , 2004, *Linearization of Chains and Sideward Movement*, Cambridge (Mass.): The MIT Press.
- Progovac, L., 1999, "Events and economy of coordination", *Syntax* 2.2., 141-59.
- , 2003, "Structure for coordination", in L. Cheng & R. Sybesma (eds.), *Second Glot International State-of-Article Book*, Berlin: Mouton de Gruyter, 241-87.
- Rooth, M., 1985, *Association with focus*, Doctoral dissertation, UMass (Amherst).
- Sánchez, C., 1999, "La negación", in Bosque & Demonte (eds.), 2561-2634.

- Schein, B., 1997, "Conjunction Reduction Redux", ms. USC.
- Schwarz, B., 1999, "On the syntax of *either ... or*", *NLLT* 17, 339-70.
- Torrego, E., 1995, "On the Nature of Clitic Doubling", in H. Campos & P. Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory: Studies in Honor of Carlos P. Otero*, Washington: Georgetown U.P., 399-418.
- , 1998, *The Dependencies of Objects*, Cambridge (Mass.): MIT Press.
- Uriagereka, J., 1995a, "Aspects of the Syntax of Clitic Placement in Western Romance", *LI* 26.1, 79-123.
- , 1995b, "An F Position in Western Romance", in K. Kiss (ed.), *Discourse configurational languages*. Oxford / New York: Oxford U.P., 153-75.
- , 2001, "Doubling and Possession", in B. Gerlach & J. Grijzenhout (eds.), *Clitics in Phonology, Morphology and Syntax*, Amsterdam: John Benjamins, 405-31.
- , 2002, "Multiple Spell-Out", in J. Uriagereka, *Derivations. Exploring the Dynamics of Syntax*, London/New York: Routledge, 45-65.
- , & J. Nunes, 2002, "Cyclicality and Extraction Domains", in J. Uriagereka, *Derivations. Exploring the Dynamics of Syntax*, London/New York: Routledge, 66-85.
- Zamparelli, R., 2000, "Distributive Conjunction and Sentence Reduction", ms. U. Bergamo.

ON VERB FOCALIZATION IN CENTRAL AND WESTERN BASQUE¹

Bill Haddican
New York University

Central and Western Basque dialects have a verb focalization strategy involving the dummy verb *egin*, which as a lexical verb is akin to English ‘make’ or ‘do.’ (Rebuschi 1984, Ortiz de Urbina 1989, Zuazo 1998, Etxepare and Ortiz de Urbina 2003). An example of this construction is given in (1) which Ortiz de Urbina (1989) gives as a felicitous answer to the question, “What happened to your father?”

- (1) Hil egin da gure aita
die do AUX our father
‘Our father has DIED.’

This paper makes the following two main claims about focalized verbs in sentences such as (1). First, this paper develops in greater empirical detail Rebuschi’s (1984) proposal that the focalized verb in constructions such as (1) moves (as an XP) to the same left-peripheral focus position —Spec, FocP— targeted by other kinds of foci. Intonation data presented in connection with this claim, moreover, lends support to Ortiz de Urbina’s (2002) remnant-movement approach to right-peripheral focalizations in Basque. Second, the focalized main verb is argued to be an infinitival, on a par with verbs under modals. These constructions, moreover, are argued to be monoclausal (Cinque 2000, Wurmbrand 2001, 2004). Evidence supporting this claim comes the behavior of the negative morpheme, *ez* and from transparency in agreement marking on the auxiliary reminiscent of clitic-climbing with Romance restructuring verbs. The Basque data presented here, then, lend support to recent work treating restructuring as a monoclausal phenomenon. These two proposals are developed in turn below.

¹ I am deeply grateful to the residents of Oiartzun for their hospitality and support during the fieldwork portion of this study. I am grateful to audience members at BIDE04, the HIM taldea and other members of the Basque linguistics community for helpful comments and suggestions about these data. Thanks in particular to Xabier Artiagoitia, Jabi Elizasu, Arantzazu Elordieta, Urtzi Etxeberria, Ricardo Etxepare, Aritz Irurtzun, Richard Kayne, Itziar Laka, Maider Lekuona, Javier Ormazabal, Jon Ortiz de Urbina, Juan Uriagereka and Koldo Zuazo. I am responsible for all remaining shortcomings. This work is supported by a Fulbright grant and NSF dissertation improvement grant number 0317842.

1. On the position of the main verb in *egin* focalization constructions

The first half of this paper develops in greater empirical detail Rebuschi's (1984) suggestion that focalized verbs in constructions such as (1) target the same left-peripheral focus position as focalized arguments and adjuncts. Evidence in favor of this claim comes from the fact that focalized main verbs behave like other kinds of foci in terms of word order, extraction from complement clauses, clausal pied-piping and intonation, which are discussed in turn below.

1.1. Word order

1.1.1. *Left-peripheral foci*

The positioning of arguments in Basque is discourse-sensitive. Foci and *wh*-phrases canonically must appear left-adjacent to the negative morpheme *ez* in negative sentences, and left adjacent to the main (aspect-bearing) verb in affirmative constructions.

- (2) a. Nor-k/JON-EK (*Miren) ikus-i du (√Miren)
 Who-ERG/JON-ERG (Mary) see-Asp(perfect) AUX (Mary)
 'Who/JOHN saw Mary.'
- b. Nor-k/JON-EK (*Miren) ez du ikus-i (√Miren)
 Who-ERG/JON-ERG (Mary) NEG AUX see-Asp(perfect) (Mary)
 'Who/JOHN didn't see Mary.'

Focalized main verbs must also be left-adjacent to the main (aspect-bearing) verb in affirmatives.

- (3) Hil (*aurten/*gure aita) egin-Ø da aurten gure aita
 die egin-Asp(perfect) AUX this year our father
 'Our father has DIED this year.'

Similarly, in negative sentences focalized verbs behave like other kinds of foci in requiring left-adjacency to the negative morpheme, *ez*.

- (4) etorri (*Jon) ez da egin (Jon).²
 come NEG AUX egin (Jon).
 'Jon didn't COME.'

In such constructions, the focalized verb must scope over negation as shown by the infelicity of continuations with concessive 'but' (*baizik*) in (5b).

- (5) a. etorri ez da egin, eta ez joan.
 come NEG AUX egin and NEG go
 'It was to come that she didn't do and not to go'
- b. etorri ez da egin, # joan baizik.
 come NEG AUX egin go but
 'It was to come that she didn't do #but to go'

² In affirmative contexts, focalized verbs are interpretable as both contrastive/corrective foci and information foci (i.e. as an answer to a *wh*-question, questioning the focalized element). For foci in negative sentences, however, a contrastive/corrective interpretation is preferred.

In this respect as well, focalized main verbs behave like other kinds of foci: left-peripheral, focalized arguments and adjuncts also obligatorily scope above negation (Etxepare and Ortiz de Urbina 2003).

- (6) (Etxepare and Ortiz de Urbina 2003)
 HORREGATIK ez nien lagun-ei arrapostu, #beste arrazoi bategatik baizik.
 That.because NEG AUX friends-DAT reply other reason one.because but
 ‘THAT is why I did not reply to my friends, #but because of another reason.’

1.1.2. *Right-peripheral foci*

A more marked and less-well studied focalization strategy is also available for some speakers in which focalized constituents appear right-peripherally,³ as in (7).

- (7) A. Elordieta (2001)
 Ardoa ekarri diot (#) ANDONI-RI
 wine brought AUX Andoni-to
 ‘I brought the wine to ANDONI.’

There appears to be significant cross-dialectal variation in the availability of this phenomenon (Etxepare and Ortiz de Urbina 2003). In some dialects this construction is marginal and requires a heavy intonational break between the right-peripheral focalized constituent and the rest of the sentence. Hualde Elordieta and Elordieta (1994) and Elordieta (2001), for example, report that in the Bizkaian dialect of Lekeitio right-peripheral focalization is extremely marked except with copulative verbs and requires a significant intonational break. In Oiartzun and in neighboring central dialects, however, this phenomenon seems to be more robust. It is not restricted to copulative environments and need not always include a heavy intonational break. (Intonation is discussed in detail below).

(8) shows that in Oiartzun Basque and neighboring dialects, main verbs in *egin*-constructions may also appear right-peripherally.

- (8) (From interview data, A1)
 Horrek egi-ten du ZUZENDU.
 That egin-IMP AUX correct
 ‘The latter corrects it.’

Crucially, this strategy seems to be most marked precisely in those dialects in which other kinds of right-peripheral foci are marked. In the dialect of Lekeitio, for example, which is otherwise conservative with respect to post-verbal foci, right-peripheral verb focalizations such as that in (8) are also marginal (A. Elordieta, p.c.).

³ In fact, for some speakers, right-peripheral foci need not be strictly right peripheral (cf. Ortiz de Urbina 2001). In particular the “right-peripheral” focalized constituent can be followed by a topic if it is set off by a pause as in (i), below.

- (i) Jonek eman dio BIZIKLETA BAT # Miren-i.
 jon give AUX bicycle one Miren-to.
 ‘Jon has given a BICYCLE to Miren.’

The most thorough generative treatment of postverbal foci in Basque is by Ortiz de Urbina (2002), who argues that in both preverbal and postverbal focus constructions, the focalized constituent moves to the same position—spec, FocP.^{4,5} The two constructions differ minimally in that postverbal focalization constructions involve an additional movement step in which the remnant constituent below FocP raises to the left of FocP, leaving the focalized constituent as the most deeply embedded material in the tree. This movement step is illustrated in (9). (See also Uribe-Etxebarria 2003).

- (9) (Ortiz de Urbina, 2002)
- $$\text{TopP}[\text{CP}_i [\text{Top}_{\text{FocP}} [\text{XP} [\text{Foc } t_i]]]]$$
-

The present proposal that the verb in *egin*-focalization constructions is an XP (in spec, FocP) seems to predict that other VP material should be able to raise with the verb. From the perspective of Ortiz de Urbina's remnant movement proposal, this predicts the availability of focalized VPs right-peripherally in those dialects with the *egin*-construction, and which are tolerant of right-peripheral focalization. Indeed, the following examples in which verbal complements may appear to the right of *egin* (but to the left of the main verb) as in (10)-(12) seem to bear out this prediction. In these examples, the most natural reading is one in which the entire VP (in brackets) or a verbal complement receives focus interpretation.

- (10) (Interview data, P1)
 Monjak egin zigun [barruan utzi.]
 Nuns egin AUX inside leave
 'The nuns LEFT US INSIDE.'
- (11) (Interview data, P1)
 Berak egin behar zituen [bi txiki jarri.]
 He/she egin need AUX two small put
 'He/she had to PUT TWO SMALL ONES.'
- (12) (Interview data, P1)
 Egin behar duzu hurrengo egun-ean [dena enboteilatu.]
 egin need AUX next day-on all bottle
 'The next day you have to BOTTLE IT ALL.'

1.2. Extraction from complement clauses and clausal pied-piping

Another well-documented property of *wh*-phrases and foci in Basque is that they may extract from complement clauses, especially under verbs of saying (Ortiz de Urbina 1989).

⁴ Ortiz de Urbina limits his proposal to corrective focalization: "In this article, I will concentrate on the 'corrective' type of contrastive focus, which finds its way more easily onto monitored registers than other types of final emphasis" (2002: 514). In this paper, I will extend Ortiz de Urbina's proposal to focus in the sense of "answer to a *wh*-question."

⁵ Ortiz de Urbina does not discuss verb focalization in this paper.

- (13) Etxepare and Ortiz de Urbina (2003)
 Nola esan du Jon-ek uste du-ela Peru-k egin behar-ko litzatekeela?
 how say AUX Jon-ERG think AUX-COMP Peru-ERG make need-FUT AUX
 'How did Jon say Peru thinks it should be made?'
- (14) Etxepare and Ortiz de Urbina (2003)
 HORRELA uste dut egin behar-ko litzatekeela aukerarena.
 this way think AUX make need-FUT AUX COMP choice
 'IN THIS WAY do I think the choice should be made.'

Predictably, at least for some speaker, focalized verbs in *egin* constructions may also extract from complement clauses, as shown in (15). The availability of extraction in such cases is further evidence that verb raising in *egin*-constructions is A'-movement.

- (15) ? ETORRI_i esan didate [t_i egin zinela].
 come say AUX egin AUX-COMP
 'They have told me that you CAME.'

Wh-phrases and foci may also pied-pipe entire clauses to the front of the matrix clause as in (16) and (17) (Ortiz de Urbina 1993).

- (16) Ortiz de Urbina (1993)
 [Nor etorri-ko d-ela bihar] esan diozu Miren-i?
 Who come-FUT AUX.COMP tomorrow say AUX Miren-DAT.
 'That who will come tomorrow have you told Mary?'
- (17) Ortiz de Urbina (1993)
 [JON etorri-ko d-ela bihar] esan diot Miren-i.
 Jon come-FUT AUX-COMP tomorrow say AUX Miren-DAT.
 'That it is Jon that will come tomorrow I have told Mary.'

Again, as expected from the standpoint of the present proposal, clausal pied-piping is also available with verb focalizations with *egin*.⁶

- (18) [Etorri egin zine-la] esan didate.
 come egin AUX.COMP say AUX
 'They say you CAME.'

⁶ Strikingly, clausal pied-piping is not available with right-peripheral foci in the lower clause.

- (i) Jonek esan dit [atzo erosi zuela BIZIKLETA BAT].
 Jon.ERG say AUX yesterday buy AUX.COMP bicycle one
 'Jon has told me that yesterday he bought a bicycle.'
- (ii) *[atzo erosi zuela BIZIKLETA BAT] esan dit Jonek.
 yesterday buy AUX bicycle one say AUX Jon.
 'Jon has told me that yesterday he bought a bicycle.'
 The same holds true for focalized verbs in *egin*-constructions.
- (iii) Esan didate [egin zin-ela han-dik etorri].
 say AUX egin AUX.COMP there.ABL come
 'They say that you CAME FROM THERE.'
- (iv) ?? [Egin zin-ela han-dik etorri] esan didate.
 egin AUX.comp there.ABL come say AUX
 'They say that you CAME FROM THERE.'

No account of these facts can be offered here.

1.3. Intonation

The following discussion presents data showing that the intonational properties of focalized verbs (and VPs) are similar to those for focalized arguments and adjuncts in both preverbal and postverbal position. These data, then, provide additional evidence that the main verb in *egin* focalization constructions occupies the same position as other kinds of foci.

1.3.1. Argument and adjunct focalization

Sentence stress in Oiartzun is similar to that described by Elordieta (2003) for the dialect of Tolosa (four towns away to the southwest).⁷ Main prominence canonically falls on a word in the syntactic phrase immediately preceding the main verb. This prominence is characterized by: (i) a pitch (F_0) peak, followed by (ii) a sharp fall in pitch, and (iii) a reduced pitch range for clausal material following the pitch peak (Hualde, Elordieta and Elordieta 1994, Elordieta 2003). An example of this pattern is given in Figure (1), the gloss for which appears in (19).⁸

(19) (Interview data, P1)

Ne(re) aurre-ku-k geio izango dute, baino bai
 my before-of-PL more have-FUT AUX but yes
 ‘Those older than me must have more, but yes.’

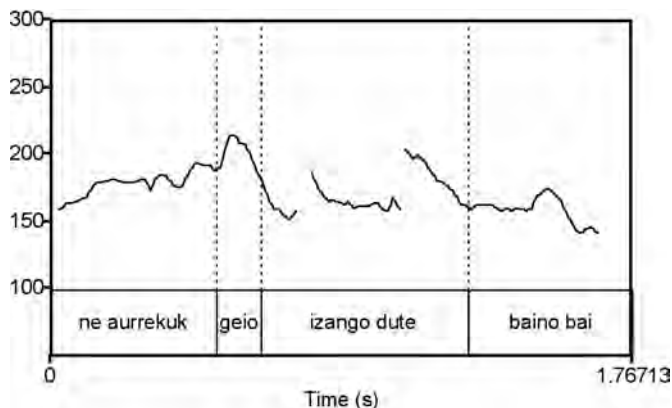


FIGURE 1

⁷ Word stress in Oiartzun Basque is similar to that in the neighboring dialect of Bortziria as described by Hualde (1991). Unlike in western, Bizkaian varieties, Oiartzun Basque has stress accent rather than pitch accent. Stress typically falls on the peninitial syllable.

(i) baséria 'farm(house)'
 Astígarraga place name
 ardóa 'wine'

There are two kinds of exceptions to this pattern: monosyllabic items, in which stress is realized on the root, and lexically marked exceptions in which stress typically falls on the initial syllable, e.g. *tálua*, 'corn tortilla'.

⁸ The following intonational data were analyzed using PRAAT (Boersma and Weenink 2003). The data were collected in sociolinguistic interviews by the author (a non-native speaker) and a native-speaking assistant in the Fall of 2003 and by the author in the Summer of 2001. The examples are given in standard Basque orthography, adjusted to reflect local phonological features.

The intonational contour in Figure 1 shows the principal properties observed by Hualde, Elordieta and Elordieta (1994) and Elordieta (2003) for foci in other Basque dialects. In particular, the focalized element in Figure 1, *geio* ‘more’ is marked by: (i) a pitch peak; (ii) followed by a sharp fall in pitch; and (iii) a reduced pitch range for the material following the focus. Note, also that the pre-focal topic phrase, *ne aurrekuk*, ‘those before me’ has a rising pitch contour. This intonational property of topics —also described by Elordieta (2003) in data from the dialect of Tolosa— will bear crucially on subsequent discussion of right-peripheral foci.

Figure 2 shows that the pitch contour of postverbal focus phrases is very similar to that for preverbal focus phrases (cf. Ortiz de Urbina 2002, Elordieta 2001). (The gloss and discourse context for the sentence is given in (20)).

(20) (Participant data, P1, describing school desks)

ta ordun genun maia # bikuak, ez? #zaten zien # eta mahai bat zegon HAUTSIYA.

and so have table of two no be ALX and table one was break-ABS.

‘And so we had tables for two, no?...they were, and one of the tables was broken.’

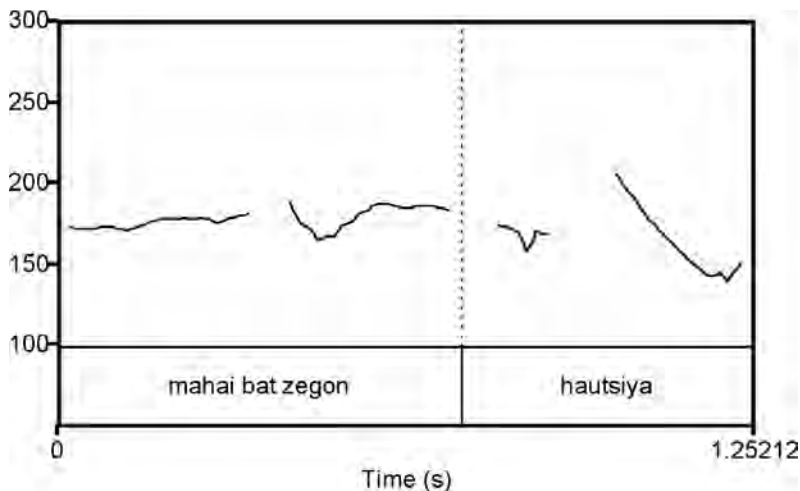


FIGURE 2

In Figure 2, the focalized element *hautsiya* ‘broken’ is marked by (i) a pitch peak followed by (ii) a fall in pitch and (iii) a greater overall pitch range than the rest of the sentence. Note also that the non-focalized material in the first part of the sentence, *mahai bat zegon*, ‘a table was’ has a slightly rising pitch contour, reminiscent of that for topic phrases (as in Figure 1). This pattern suggests additional evidence in favor of Ortiz de Urbina’s (2002) proposal that right-peripheral focus constructions are derived by moving the remnant material below the focalized constituent (in FocP) to a higher, topic position. In particular, the remnant-moved material to the left of the focus seems to have precisely the same pitch properties

otherwise evidenced by pre-focal topics. In addition, Ortiz de Urbina (2002) points out that the material preceding right-peripheral foci has another key intonational property of topics. As discussed in 2.1.2, some speakers require an intonational break between the postverbal focalized phrase and the preceding material. This break is similar to the pause often required between topics and the focus phrase (Ortiz de Urbina 1989). From the perspective of Ortiz de Urbina's proposal, then, these topic-like intonational properties of the material to the left of the focalized constituent in Figure 2 are accounted for straightforwardly, since by this approach, they *are* in fact topics.

1.3.2. Verb focalization

In discussing the Bizkaian dialect of Lekeitio, Elordieta (2003) reports that in left-peripheral verb focalization constructions with *egin*, main prominence is assigned to the main verb. Again, from the standpoint of the present proposal in which the main verb moves to the same position as other kinds of foci, this is precisely the pattern expected.

In the dialect of Oiartzun, the verb also receives main prosodic prominence. Figure 3, shows the F_0 contour for an *egin*-focalization construction in Oiartzun Basque. The gloss for this example is given in (21).

(21) (Interview data F1)

Ordun ya, gerra ondo-an hori itxi (eg)in zen.
 so then, war after-LOC that close egin AUX.
 'So then, after the war the latter closed.'

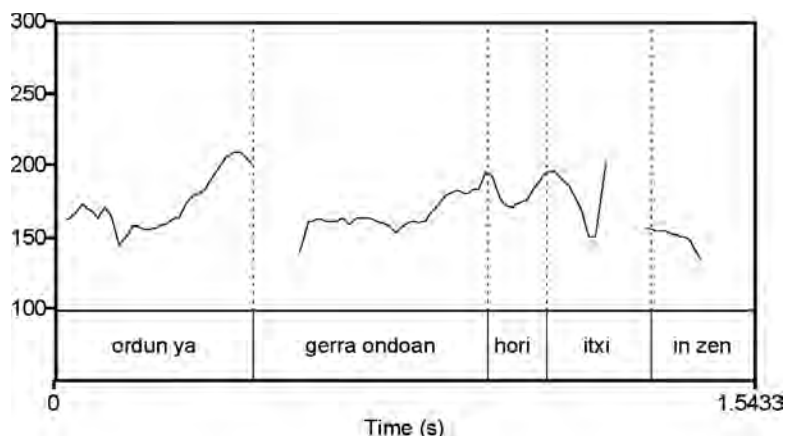


FIGURE 3

Figure 3 shows that the main pitch peak is realized on the main verb, *itxi*, 'close' (setting aside the pre-sentential tag) followed by a sharp fall in pitch, characteristic of focalized constituents. Figure 3 also shows that the preverbal topic phrases have a

rising intonation, each higher than the previous one, as described by Elordieta (2003) for the dialect of Tolosa.

Figure 4 shows a pitch contour for a right-peripheral verb focalization with *egin*. The corresponding context and gloss is given in (22).

- (22) (Interview data, P1)
 ordun altxa zen-in **nik (eg)in nion sila kendu #** eta lurr-e(r)a
 so get up AUX-COMP I-ERG *egin* AUX chair take away and floor.ABL
 ‘So when he got up I pulled out the chair, and he fell down.’

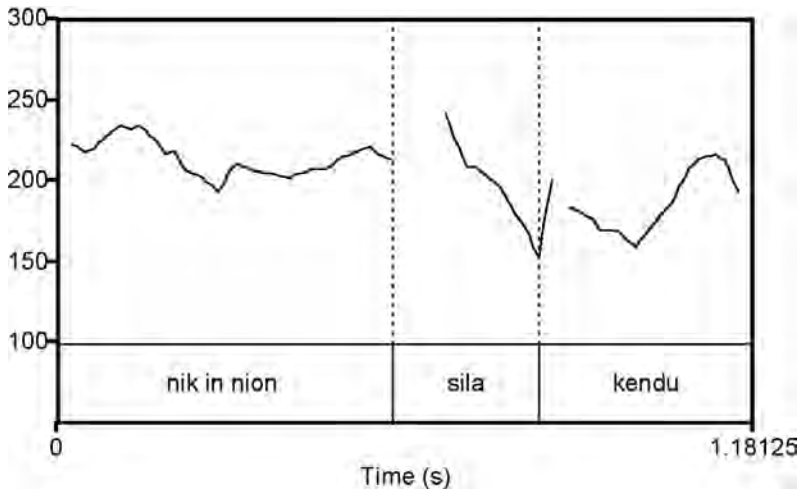


FIGURE 4

In Figure 4, the main pitch peak is realized on the object *sila*, ‘chair’ followed by a sharp fall in pitch. The focalized VP is also marked by a greater pitch range than the rest of the sentence.

Figure 5 —corresponding to the gloss in (23)— shows a slightly different intonational pattern for right-peripheral, focalized VPs.

- (23) (Interview data, P1)
 Monjak egin zigun [barruan utzi.]
 Nuns egin AUX inside leave
 ‘The nuns left us inside.’

The pitch contour in Figure 5 is similar to that in Figure 4 in certain respects. The right-peripheral VP has a pitch peak on the stressed syllable of the locative complement *barrúan*, ‘inside,’ followed by a sharp pitch fall. The pitch range for the focalized VP is also greater than the preceding non-focalized material. In Figure 5, however, the pitch peak on the stressed syllable in *barrúan* is lower than the preceding peak marking the right edge of the topicalized remnant material, *monjak in zigun*. It appears then, that in such cases, it is primarily the pitch fall and overall pitch range that does the work of marking prominence.

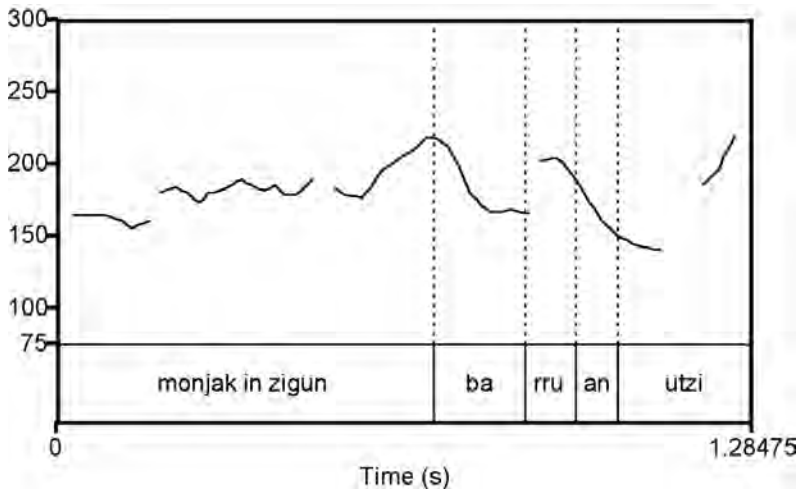


FIGURE 5

To review, the foregoing data on word order, intonation, extraction and clausal pied-piping strongly recommend a unified analysis for focalized main verbs in *egin*-constructions, on one hand, and focalized arguments and adjuncts on the other. In particular, I follow Ortiz de Urbina (1998, 2002) in assuming movement of both types of foci to a left-peripheral focus position, FocP. This approach allows for a unified analysis of right-peripheral and left-peripheral focalizations in Basque (Ortiz de Urbina 2002): in both cases, the focalized constituent moves to FocP, however right-peripheral constructions involve an extra movement step in which the remnant material below FocP raises to a higher topic position. An additional virtue of this analysis is that it also accounts for the topic-like intonational properties of remnant-moved constituents to the left of right-peripheral foci.

2. The focalized main verb as an infinitive

The second main claim to be made in this paper concerning verb focalization constructions with *egin* is that the focalized verb is an infinitival, on a par with verbs under modals. These infinitivals, moreover, are argued to be merged in the same extended functional sequence as the main verb (Cinque 2000, Wurmbrand 2001, 2004). Evidence supporting this claim comes from transparency in agreement marking on the auxiliary and the behavior of the negative morpheme, *ez*. The Basque data presented here, then support recent work treating restructuring as a monoclausal phenomenon.

2.1. Modals and infinitivals

The dummy verb, *egin*, in verb focalization constructions such as (1) bears one of three aspectual markers —perfect $-\emptyset$, imperfect $-t(z)en$ or future $-ko$ — normally realized on the main verb as shown in (24) and (25).

(24) **verb focalization**

erori (egin-go/egi-ten) du etxea.
 fall egin.FUT/egin.IMP AUX house
 'The house is going to fall.'/'The house falls.'

(25) **non-verb focalization**

etxe-a (erori-ko/eror-tzen) da
 house-the fall.FUT/fall.IMP AUX
 'The house will fall down.'/'The house has fallen down.'

In verb focalization environments, the main verb bears one of three affixes, *-tu*, *-i* or $-\emptyset$, depending on the verb class.⁹ These affixes are standardly analyzed as perfective markers, in view of the fact that they co-occur with perfective interpretations in non-focalization environments like (26)-(28) and are in complementary distribution with the imperfect marker *-t(z)en*, as shown in (29) (Laka 1989, Zabala and Odriozola 1996). The verb root + *-tul-il- \emptyset* is also the citation form of the verb.

(26) opera-tu didate
 operate-PERF. AUX
 'They operated on me.'

(27) etor(r)-i da
 come-PERF. AUX
 'She has come.'

(28) eman- \emptyset didate
 give-PERF. AUX
 'They gave it to me.'

(29) funtziona-tzen du.
 function-IMP. AUX
 'It works.'

In view of these facts, Laka (1989, 1990) proposes that perfective *-tul-il- \emptyset* and imperfective *-tzen* are alternate values of a single aspectual head, Asp⁰ (cf. Zabala and Odriozola 1996). Nevertheless, two aspects of the behavior of these affixes in *egin*-focalization constructions are surprising from the perspective of this analysis. First, as discussed above, other aspectual markers such as imperfective affix *-t(z)en* and future *-ko* are realized on the dummy verb, *egin*, as shown in (30) repeated below.

(30) eror-i (egin-go/egi-ten) du etxea.
 fall egin.FUT/egin.IMP AUX house
 'The house will fall.'/'The house falls.'

The main (focalized) verb however obligatorily bears *-tul-il- \emptyset* as in (30) and (31).

(31) (From interview data, A1)
 Horrek egi-ten du zuzen-du.
 That egin-IMP AUX correct
 'The latter corrects it.'

From the standpoint of Laka's AspP proposal and the assumption that *-tul-il- \emptyset* are always perfective markers, the data in (31) are perplexing since they seem to require the realization of different values of a single aspectual head — *-t(z)en* and *-tu* — on different items in a single clause. (Evidence is provided below that these

⁹ The open class affix is *-tu* as shown in (26). A smaller, older class of verbs takes *-i* as shown in (27), and an even smaller class of verbs ending in /n/ takes $-\emptyset$ as shown in (28).

constructions are in fact monoclausal rather than biclausal.) Second, this analysis seems to predict conflicting aspectual interpretations for examples like (30) and (31). In these cases, however, the aspectual reading is invariably determined by the affix on the dummy verb, as reflected in the glosses.

The behavior of *-tul-il-Ø* on verbs selected by modals provides additional reason for skepticism with regard to the traditional analysis of these elements. In particular, verbs under *behar* ‘need’, *nahi* ‘want’, and *ahal* ‘can’ all obligatorily take *-tul-il-Ø* regardless of the perfectiveness of the action, as shown in (32). In other words, when suffixed to verbs selected by modals, *-i*, *-Ø* and *-tu* do not always mark perfective aspect (Ortiz de Urbina 1992, cited in Zabala and Odriozola 1996: 238, fn.2).

- (32) Maiz etorri nahi dute
 often come-I want AUX-3PLE
 ‘They want to come often.’ (want>often)

I would like to propose that these problems with the standard analysis of *-tul-il-Ø* can be solved by positing a dual identity for these morphemes. In examples such as (26)-(28), these morphemes are true perfective markers. With modals, and in *egin*-focalization constructions however, these morphemes do not mark perfective aspect but rather are infinitival markers. This approach explains: (i) the fact that *-tul-il-Ø*, unlike other aspectual affixes may appear on the main verb in these environments; (ii) their compatibility with other aspectual heads in *egin*-focalization constructions; and (iii) the availability of imperfective aspectual readings with modals. Moreover, the fact that the verb root + *-tul-il-Ø* is also the citation form of the verb suggests additional support for this analysis over the competing, standard approach to these elements as (always and everywhere) aspectual heads. That is, as citation forms, verbs with infinitival markers are commonplace while verbs with perfective morphology as citation forms are more marked.

2.2. Evidence for a monoclausal approach to infinitivals

It might be objected that both *egin*-focalization constructions and modal constructions are plausibly biclausal. From this perspective, *-tul-il-Ø* on the main verb might be understood as the realization of *Asp⁰* in the lower, non-finite clause. Evidence against this approach, however, comes from the fact that both *egin* and modals participate in agreement phenomena reminiscent of clitic-climbing with Romance “restructuring” verbs (Rizzi 1978). On the assumption that restructuring is not possible across CP boundaries (Cinque 2000, Wurmbbrand 2001, 2004), then such facts suggest that these infinitival constructions are monoclausal.

Like other transitive verbs, *egin* as a lexical verb requires that person and number agreement with the object(s) be marked on the auxiliary.

- (33) etxe-ko lanak egin ditut
 house-of work-PL-A do AUX-3Pl.Abs.-1Erg.
 ‘I’ve done my homework.’

In *egin*-focalization constructions, however, the agreement marking is determined by the argument structure of the focalized main verb: unaccusative verbs require intransitive agreement and transitive verbs require transitive agreement.

- (34) a. Joan egin naiz (unaccusative)
 go do AUX-1Abs.
 'I have GONE.'
- b. Torrea ikusi egin dut (monotransitive)
 tower see do AUX-3Abs.-1Erg.
 'I've SEEN THE TOWER.'
- c. Jon-i liburua eman egin diot (ditransitive)
 Jon-D. book give do AUX-3Abs.-3Dat.-1Erg.
 'I've GIVEN Jon the book.'

Similar facts obtain for verbs under the modal *ahal* 'can'. (35a)-(35c), for example, show that agreement marking on the auxiliary is a function of the argument structure of the main verb.

- (35) a. Joan ahal naiz (unaccusative)
 go MOD AUX-1Abs.
 'I can go.'
- b. Torrea ikusi ahal dut (monotransitive)
 tower see MOD AUX-3Abs.-1Erg.
 'I can see the tower.'
- c. Jon-i liburua eman ahal diot (ditransitive)
 Jon-D. book give MOD AUX-3Abs.-3Dat.-1Erg.
 'I can give Jon the book.'

The modals *behar* 'need' and *nahi* 'want' behave somewhat differently. These verbs require ergative agreement morphology on the auxiliary, even with unaccusative main verbs. The difference between these two classes of modals is illustrated by the contrast between (35a) and (36a).

- (36) a. Joan nahi dut (unaccusative)
 go MOD AUX-1Erg.-1Abs.
 'I want to go.'
- b. Torrea ikusi nahi dut (monotransitive)
 tower see MOD AUX-3Abs.-1Erg.
 'I want to see the tower.'
- c. Jon-i liburua eman nahi diot (ditransitive)
 Jon-D. book give MOD AUX-3Abs.-3Dat.-1Erg.
 'I want to give Jon the book.'

These three modals (and *egin* as a dummy verb) differ from other subject control verbs such as *saiatu*, 'try' in two key respects. First, verbs under subject control verbs like *saiatu*, 'try,' do not bear *-tul-il-Ø* but rather *-t(z)en* as shown in (37)-(40) below.¹⁰ Second, other kinds of subject control verbs, do not exhibit transparency in agreement marking. Rather, agreement marking is exhaustively determined by argument structure of the higher verb.

¹⁰ See Zabala and Odriozola (1996) for a discussion of this affix including an argument that they are the same affix as the imperfective morpheme.

(37)
 Joa-ten saia-tu naiz
 go-*t(z)en* try-PERF AUX-1Abs.
 'I have tried to go.'

(38)
 Torrea ikus-ten saia-tu naiz
 Tower see-*ten* try-PERF AUX-1Abs.
 'I have tried to see the tower.'

(39)
 Ona iza-ten ikas-i dut
 good be- *t(z)en* learn-PERF AUX-3Erg-1Abs.
 'I have learned to be good.'

(40)
 Gitarra jo-tzen ikas-i dut
 guitar play-*t(z)en* learn-PERF AUX-3Erg-1Abs.
 'I have learned to play guitar.'

The transparency in agreement marking with verbs under *egin* and modals *ahal*, *behar* and *nahi* as in (34)-(36) suggest that these constructions are monoclausal (Cinque 1999, 2000, Wurmbrand 2001, 2004). Verbs such as *saiatu*, 'try' and *ikasi*, 'learn,' on the other hand, are presumably main verbs, which optionally select for non-finite clauses in a familiar way.

Additional evidence that modal constructions are monoclausal, unlike constructions with *saiatu*, 'try' and *ikasi*, 'learn,' comes from the behavior of negation. With the latter class of verbs, the negative morpheme, *ez*, can appear before the lower verb. The interpretation of these examples is one in which negation scopes between the lower verb and the upper verb as reflected in the glosses.

(41) Saiatuko naiz ez pentsatzen hor-(r)etan
 try-FUT AUX NEG think-*tzen* that-in
 'I'm going to try not to think about that.'

(42) Ikas-i behar duzu ez iza-ten hor-(r)ela
 learn-INF need AUX NEG be-*tzen* that-like
 'You have to learn not to be like that.'

However, negation is not possible with verbs under the modals *ahal*, *nahi* and *behar*.

(43) *Nahi dut ez jun (44) *Behar duzu ez izan hor-(r)ela
 want AUX NEG go need AUX NEG be that-like
 'I want not to go.' 'You need not to do that.'

These facts follow readily from the monoclausal approach adopted here. The negative morpheme *ez*, in (41) and (42) is plausibly merged in the lower, non-finite CP. In contrast, no negation is possible in (43) and (44) because such constructions are monoclausal and no NEG position is available that low in the functional sequence in which to merge these elements.

Finally, this monoclausal vs. biclausal distinction suggests an explanation of the differences in morphology on the lower verb in these two environments. Main verbs below subject control, non-modals such as *saiatu* obligatorily bear a *-tzen* affix, (which is standardly described as an imperfect aspectual marker.) In contrast, main verbs under modals take a different morpheme, *-il/-tu/-Ø*, plausibly akin to infinitival markers like *-ar/-er/-ir* in Spanish. From this monoclausal perspective, then, the *-tzen* affix in (41) and (42) might plausibly be taken to reflect a non-finite CP boundary.

3. Conclusions

This paper discusses verb focalization constructions in Central and Western Basque dialects. Some specific theoretical consequences of this proposal are summarized below.

1. This paper provides support for Rebuschi's (1984) proposal that the main verb in verb focalization constructions targets the same position as other kinds of foci. Evidence for this claim comes from the behavior of these elements in terms of word order, extraction from complement clauses, clausal pied-piping and intonation.
2. The intonational data presented here support Ortiz de Urbina's (2002) remnant movement approach to right-peripheral foci in Basque. The latter proposes that right-peripheral focus constructions are derived by (leftward) remnant movement of non-focalized material to a topic position above the focalized constituent in FocP.
3. The data presented here support recent approaches to "restructuring" as a monoclausal phenomenon (Cinque 2000, Wurmbrand 2001, 2004). In particular, restructuring in Basque is available in precisely the same class of non-finite constructions in which sentential negation above a lower, non-finite verb is unavailable. The Basque data are striking in that "restructuring" constructions exhibit different morphological properties from other kinds of subject control constructions. This difference plausibly reflects the existence of a non-finite CP boundary in one case but not the other.

References

- Boersma, P. and D. Weenink, 2003, PRAAT, www.praat.org.
- Cinque, G., 1999, *Adverbs and Functional Heads*. New York: CUP.
- , 2000, "Restructuring' and Functional Structure". Ms. University of Venice.
- Elordieta, A., 2001, *Verb Movement and Constituent Permutation in Basque*. Utrecht: LOT Dissertation Series.
- Elordieta, G., 1997, *Morphosyntactic Feature Chains and Phonological Domains*. Doctoral dissertation, USC.
- , 2003, "Intonation". In Hualde and Ortiz de Urbina (eds.), pp. 72-112.
- Ettxepare, R. and J. Ortiz de Urbina, 2003, "Focalization". In Hualde and Ortiz de Urbina (eds.), pp. 460-516.
- Hualde, J. I., 1991, *Basque Phonology*. London: Routledge.
- , G. Elordieta and A. Elordieta, 1994, *The Basque dialect of Lekeitio*. Supplements of the International Journal of Basque Linguistics and Philology XXXIV. Bilbo/Donostia: University of the Basque Country & Gipuzkoako Foru Aldundia.
- , and J. Ortiz de Urbina (eds.), 2003, *A Grammar of Basque*, Berlin: Mouton of Gruyter.
- Laka, I., 1989, "Constraints on Sentence Negation: The Case of Basque". In *Functional Heads and Clause Structure*, I. Laka and A. Mahajan (eds.), *MIT Working Papers in Linguistics* 10: 199-216.
- , 1990, *Negation in Syntax. On the Nature of Functional Categories and Projections*, Ph.D. dissertation, MIT.

- Ortiz de Urbina, J., 1989, *Parameters in the Grammar of Basque*. Dordrecht: Foris.
- , 1992, “Tenseless Verbal Forms in Basque”. Ms. Universidad de Deusto.
- , 1993, “Feature Percolation and Clausal Pied-Piping”. In *Generative Studies in Basque Linguistics*, J. I. Hualde and J. Ortiz de Urbina (eds.), pp. 189-219. Amsterdam/Philadelphia: John Benjamins.
- , 1998, “Force Phrases, Focus Phrases and Left Heads in Basque”. In *Grammatical Analyses in Basque and Romance Linguistics*, J. Franco, A. Landa and J. Martín (eds.), pp. 179-94. Philadelphia: Benjamins.
- , 2002, “Focus of Correction and Remnant Movement in Basque”. In *Erramu Boneta: Festschrift for Rudolf P.G. de Rijk*, X. Artiagoitia, P. Goenaga and J. A. Lakarra (eds.), pp. 512-524. Bilbo: University of the Basque Country.
- Rebuschi, G., 1984, *Structure de l'énoncé en Basque*. Paris: SELAF.
- Rizzi, L., 1978, “A Restructuring Rule in Italian Syntax”. In *Recent Transformational Studies in European Languages*, S. J. Keyser (ed.), pp. 113-58. Cambridge: MIT Press.
- Wurmbrand, S., 2001, *Infinitives: Restructuring and Clause Structure*. Berlin/New York: Mouton de Gruyter.
- , 2004, “Two types of restructuring-Lexical vs. functional”, *Lingua* 114,8: 991-1014.
- Zabala, I. and J. C. Odriozola, 1996, “On the relation between DP and TP. The Structure of Basque Infinitivals”, *Catalan Working Papers in Linguistics* 5: 231-81.
- Zuazo, K., 1998, “Euskalkiak gaur”, *FLV* 30: 191-233.

WHAT HAPPENS WHEN PHASES GET INDIVIDUALISTIC: ON LF-ONLY & PF-ONLY PHASES*

Franc Marušič
SUNY at Stony Brook

Introduction

The notion of cycle has played an important role in linguistic theory from the beginning, when it was defined on completed clauses as a point at which all the operations and rules were applied, to the latest developments of the Minimalist Program and the Phase theory (Chomsky 2001). In phase theory, a cycle/phase is a complete stage in the derivation from the numeration to the units at the two interfaces. A phase starts with its own numeration of lexical items, which are merged together to build structure. Lexical items in this view are of type {S,P,F}-bundles of semantic, phonological and formal features. All parts of the initial structure therefore begin the derivation at the same time, but since some items might move out of their source phase, it is obvious that not all are spelled-out (sent to the two interfaces) at the same time—in the same phase. Since we already have a phase mismatch between what enters the derivation and what exits when a phase is completed, we can expect the mismatch will be actually even bigger (either we have full matching, or else we have no matching). So the question is, where else can we find mismatch in terms of things that entered the derivation simultaneously not getting spelled-out in the same phase.

This paper looks at another possible mismatch, a mismatch of phase completion for the two interfaces. The standard assumption is that when a phase is completed features participating in the derivation get shipped to both interfaces at the same time, so that vP and CP would both be a PF and an LF phase. This seems to be the easiest way in which a derivation could go, but it might not conform to the minimal design requirement. In addition, as pointed out above, we already have a mismatch. Suppose that at the point of Spell-Out a phase is either only a PF phase or only an LF phase — that at the point of Spell-Out features can get only shipped to a single interface. Such theory seems to involve less rules and could be therefore preferred.

* Work on this topic was partially supported under the NSF Grant BCS-0236952 (Richard Larson, PI). I'd like to thank the organizers of BiDe'04 for the invitation and the audience for comments. I also thank Richard Larson, Dan Finer, and Rok Zaucer for helpful discussions and Diane Abraham, Sandra Brennan, Carlos de Cuba, and Jon MacDonald for help with data.

Allowing phases to be interface specific we can make several interesting predictions. Since pronunciation and interpretation of a particular item do not have to occur at the same time, an item can be interpreted higher than it surfaces or it can get pronounced higher than it is interpreted. Both cases seem to correspond to phenomena in natural language. When something is interpreted lower in the structure than its surface position, it is said to have reconstructed. On the other hand, when, for example, a quantifier phrase gets interpreted higher than its surface position, it is said to have covertly moved.

- (1) interpretation > pronunciation → instance of **Covert movement**
 pronunciation > interpretation → instance of **Reconstruction**

These ideas have been anticipated to some extent in the literature. For example Nissenbaum (2000) describes the difference between *covert* and *overt movement* as a difference in timing between Spell-Out and move. If movement to the edge applies prior to Spell-Out, movement is overt, but if Spell-Out applies prior to movement to the edge, movement is covert. With the standard assumption that Spell-Out is simultaneous, movement to the edge should not be possible after Spell-Out.¹

As for the other case, Aoun & Benmamoun (1998) treat total reconstruction to be the result of PF movement. Sauerland & Elbourne (2002) extend this proposal and claim that only things that are already shipped to LF at some earlier point, at some intervening phase, reconstruct.

In this paper I look at the latter phenomenon and show that is not really a case of PF movement, as claimed by Aoun & Benmamoun (1998), but rather a case of an LF phase being inside a wider PF phase (an LF phase being smaller than a PF phase). In section 1 I describe the problem, and give the proposal. In section 2, I discuss two controversial claims made in the proposal — the existence of split phases and the agreeing PF features. Section 3 discusses some potential problems and section 4 concludes the paper.

1. Total reconstruction

As it is well known, the examples in (2) are ambiguous. That is, both indefinite subjects in (2) can be interpreted non-specifically, in the scope of *likely*. To be more specific, there need not be any particular Basque in (2a) that has the property of being likely to win the Tour, nor need there be anyone specific from Xabier in (2b) who has the property of being likely to win the lottery. It could be that it is just likely that some Basque cyclist wins the tour or that a resident from Xabier will win the lottery.

- (2) a. *A Basque is likely to win the Tour (in the coming years)* likely > a Basque
 b. *Someone from Xabier is likely to win the lottery* likely > someone from Xabier

¹ The truth is, he wasn't talking about simultaneous shipment, since, at the time, Spell-Out was just a point where phonological features got separated from the derivation, which continued its way to LF.

Since the surface position of *a Basque* is higher than the position of *likely* in (2a), the DP is said to have reconstructed. The DP is interpreted in the lower clause from which it originates, but it does not surface there, therefore some operation had to either move it up for pronunciation or move it down for interpretation.

Both of these possibilities have been explored. One of the earliest analysis of reconstruction took it to be a lowering operation at the LF, after syntax has done all its upward movements (May 1985). In particular, the cases in (2) were analyzed to involve total reconstruction, lowering of the entire DP constituent to a lower scopal position, something like what is depicted in (3), where *A Basque* first raises over *likely*, and finds itself in the highest surface position (at the point of Spell-Out), but is later lowered to the clausal boundary where it takes lower scope. With this kind of lowering operation, the derivation returns to a previous stage of the derivation. The lowering operation is thus an undoing operation and as such unwanted. In an ideal language design we would not want to do something just so that we can undo it later.

- (3) a. $[A \text{ Basque}]_i$ is likely to t_i win ... (in syntax proper)
 └──────────────────┘
 b. └── is likely [a Basque] $_i$ to t_i win ... (at LF)

Chomsky (1993) proposed a different approach to reconstruction which does not involve undoing operations. According to the *copy theory* of movement, movement leaves a copy of the moved material rather than a trace in every position it moves from. At the two interfaces one of the two copies must be deleted but that is not necessarily the same copy at both interfaces. In case of total reconstruction as in (2), the first-merged constituent is not pronounced at PF, but it gets interpreted at LF, while the remerged higher copy doesn't get interpreted at LF, but it is interpreted at PF — pronounced.

Although we don't have any unwanted undoing operations, we are still left without an answer to an important question: "How do we know/determine which copy is pronounced and which interpreted?" There is a further problem since this approach leaves the interpretation at the two interfaces as optional. But having things as optional is not an optimal design feature.

It should be added that the kind of reconstruction we are interested in, total reconstruction, is different from the better known and more widely discussed *binding reconstruction* or *partial reconstruction*, as in (4) (this short discussion is a summary of Sauerland & Elbourne's).

- (4) [*Which article about himself*] $_i$ did every politician $_k$ read t_i ?

As evident from the indexing, part of the moved *wh*-constituent must reconstruct in order for the universal quantifier to c-command the reflexive at LF. But the question is what part reconstructs. As Saito (1989) pointed out, reconstruction found in (4) is not comparable to the one in (2) for the simple reason that in (4) it is not the whole *wh*-constituent that reconstructs. This can be most clearly seen in (5). If the whole *wh*-constituent reconstructs leaving in the upper most scope position only the Q marker, we would expect (5a) to be the LF representation of (5b).

- (5) a. *Did Maggie ask [which cousin]_i to call t_i?*
 b. *[Which cousin]_i did Maggie ask whether to call t_i?*

But it is not just the Q marker that is interpreted high. (5a) and (5b) are two different questions. In these cases only parts of the moved phrase occupy a lower position at LF, (6) being an LF representation of the question in (4).

- (6) *Which_i did every politician_k read [article about himself]_j_i*

Whatever the best analysis of these cases turns out to be, they are crucially different from the phenomena discussed here — total reconstruction, where the entire moved phrase occupies a lower position at LF. Total reconstruction is not available with *wh*-movement.

1.1. Total Reconstruction as PF Movement

All proposals so far analyzed total reconstruction as involving the initial overt movement followed by an optional undoing operation, either lowering or deletion of the remerged element. To avoid the undoing operation, Sauerland & Elbourne (2002) defend a proposal by Aoun & Benmamoun (1998) that total reconstruction comes as a result of PF movement. Aoun & Benmamoun show that PF movement is involved in certain Clitic left dislocated phrases in Lebanese Arabic, which are also subject to total reconstruction. As they explain, since these dislocated phrases only move in the PF component of the derivation, they do not affect their LF structure, which remains as it is at the end of the common syntactic derivation at the point of Spell-Out. Sauerland and Elbourne (2002) extend and strengthen this claim by claiming that total reconstruction is available only as a result of PF movement.

The subjects in (1) are part of the common syntactic derivation, which they call *stem derivation* to the point of the first TP phase. When the derivation reaches TP, Spell-Out occurs, the subject is frozen in its position, and later sent to the interfaces. After the stem syntactic derivation, the subject moves higher in the PF component, to satisfy a PF interface condition.

Since all operations occurring at PF must follow Spell-Out, at which point material is shipped to the interfaces, they also follow the stem derivation. Since at the point of Spell-Out the derived structure is also sent to the LF interface, all subsequent PF only operations fail to have any effect on the LF. There is no path from the PF interface back to LF, therefore PF movement cannot affect interpretation. All PF moved constituents get interpreted at the point where they were located at Spell-Out (unless there are some further LF operations transposing them).

1.2. Why is total reconstruction not just PF movement

In order to get their analysis going, Sauerland & Elbourne (2002) have to make several controversial assumptions. They have to argue that the (standardly syntactic) need to have a filled SpecTP —the EPP— is satisfied with a PF movement, that EPP is in effect a PF condition. In addition, it is a bit strange, that PF movement can target specific syntactic positions, especially if PF consists of no more than

phonological features, but views of the structure available in the PF component differ and this should not be taken as an objection.

Sauerland & Elbourne's analysis of (2) also makes one crucial wrong prediction. If at the point of TP the derivation reaches a phase and all the material is frozen or shipped to the interfaces, we would predict that the DP later PF-moved to a higher position does not have any effect on the higher portion of the sentence, that it does not participate in the later syntactic derivation. In particular, the low-interpreted DP—with narrow scope interpretation—should not trigger agreement on the verb/T from the matrix clause, since its *phi*-features are already spelled out and have already left the syntactic derivation. But if nothing moves to check the features on T, they could only get default values (if any at all), but this is not what we find. The plural DP in (7) is subject to total reconstruction and it also agrees with the upper T.

- (7) a. *5 Basques are likely to win all the jerseys* likely > 5 Basques
 b. *Scissors are likely to be in the drawer* likely > scissors²

Unless we put agreement in the PF side of derivation, we would not be able to derive sentences in (7). But having agreement in PF is also not permitted by Sauerland & Elbourne (2002). They need agreement in the stem derivation to explain facts like (8) from British English. As seen in (8), even without overt plural marking, collective names can trigger plural agreement (supposedly with [Mereology: plural]). Interestingly, when they do trigger plural agreement in the sentences under discussion, the subject is not subject to total reconstruction (the indefinite only gets the specific reading), which suggests it was LF interpreted in its surface position. Since the agreement on the verb is forced by [Mereology: plural], which is a semantic feature that never spells-out to PF, it could not have been sent to LF at the lower TP phase, otherwise there would be nothing to interpret in the matrix clause and no features to trigger agreement.

- (8) a. *A northern team is likely to be in the final.* \exists > likely, likely > \exists
 b. *A northern team are likely to be in the final.* \exists > likely, *likely > \exists

With this in mind, a PF moved DP should not be able to trigger any agreement in the matrix clause, yet as we see in (7) it does.

1.3. A different approach to PF movement

The proposal made here that avoids these problems is already hinted at by Sauerland & Elbourne (2002: 315):

Slightly extending Chomsky's idea, we propose that actually the edge of a phase can be distinct for LF and PF and that a phrase is only the LF or PF edge of a phase is accessible only for LF or PF movement, respectively, in a later phase.

² More about the plural agreement and pluralia tantum nouns in section 2.2.

I want to elaborate this line of thinking and show how it can explain the data correctly. If we accept that phases can spell-out/ship features to PF or LF alone, we can retain all movements in (2) in syntax proper. Syntactic elements/objects would move in syntax proper, but whatever moves, would not be a standard syntactic object/item anymore. The item moved after such a partial Spell-Out would only have specific PF-only characteristics, with all the relevant LF-only features spelled-out.

When the derivation of (2) reaches the TP phase the Spell-Out doesn't spell-out to both interfaces, but only to LF. TP is not a standard phase boundary therefore we would not expect it to behave like other phase boundaries. Sauerland & Elbourne (2002) suggest that, in addition to vP and CP, TP should be considered a phase, but for them, every phase is a bi-interface phase and as we have seen, their approach runs into problems. Rather than taking TP as a usual/standard phase boundary, I suggest it is an LF-only phase induced by the scopal predicate *likely* in the sense of Wurmbrand & Bobaljik (to appear). Since whatever would normally be sent to PF in a complete phase (CP and vP) stays in the derivation, it can participate in further syntactic derivation (of course with some limitations). Accepting this kind of approach, we retain all the movements in syntax proper. The derivation for the particular item (parts of which were sent to LF) would not be a typical *stem derivation* anymore, since not everything participating in it would be sent to both interfaces when the next higher phase is reached.

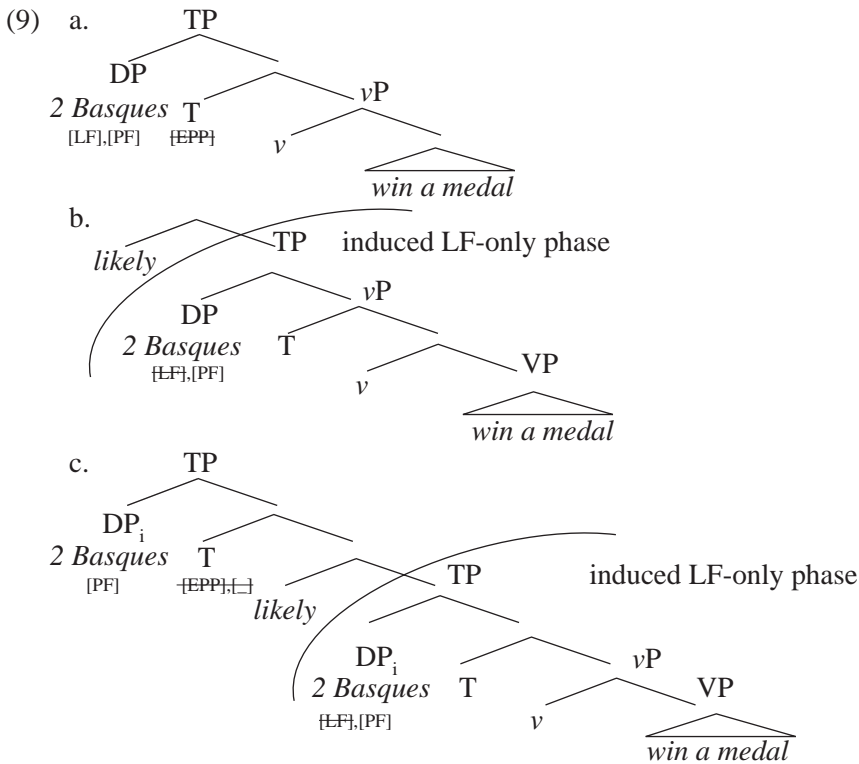
The lower clause is derived in the usual way by stem derivation. The feature bundle of the subject checks the EPP of the lower TP as it moves from Spec- vP to Spec-TP, (9a). As said, when *likely* is merged into the structure, it induces an LF phase, (9b). When the projection of *likely* (whatever it is) is complete, the LF related features ($[_{LF} x]$) of the complement of *likely* are frozen in its place/sent to LF, where the complement of *likely* creates a semantic unit, a proposition. But since *likely* only induces an LF phase, all the PF related features ($[_{PF} x]$) are left untouched.

A lexical item is standardly viewed as having three types of features, semantic, phonological and formal. Chomsky (1995b) claims each lexical entry is of the form $\{P, S, F\}$, where P (phonological features) yield π (pronunciation), S (semantic features) yield λ (interpretation), and F (formal features) participate in the derivation, but must be eliminated for convergence. Since the feature we are interested in at the moment, [Plural] on the DP, is interpretable at LF, it seems natural to treat it as a semantic feature. But if it is a semantic feature, then it should get spelled-out/frozen when the derivation reaches the induced phase. Yet as shown in (7), plurals can get non-specific interpretation and trigger plural agreement. Phonological features are the kind of features that have no influence on interpretation (they are sent to PF). But if plural has overt morphology on nominals, than [Plural] also has to reach the PF interface (or more precisely the Morphological component on the way from Spell-Out to actual pronunciation). In addition, *Pluralia tantum* nouns are not necessarily interpreted as plural entities, yet they have overt plural marking and in addition trigger plural agreement, suggesting, that the phonological [Plural] ($[_{PF} \text{Plural}]$) can trigger plural agreement just as well as purely semantic $[_{LF} \text{Plural}]$ does.

Since $[_{PF} \text{Plural}]$ is a phonological feature, one could suspect the whole verbal agreement occurs in PF, thus saving Sauerland & Elbourne's (2002) analysis. But notice that agreement on the verb can be triggered both by purely semantic features

like $[_{LF} \text{Mereology}]$ as claimed by Sauerland & Elbourne (2002) and by purely phonological features like $[_{PF} \text{Plural}]$. This shows that agreement cannot happen in only one part of the derivation (either only in PF or only in LF). Further, the fact that in our case plural agreement is derived with non-specific interpretation, suggest that agreement is also not restricted to stem derivations. We will return to the issue of what exactly $[_{LF} \text{x}]$ features and $[_{PF} \text{x}]$ features are in section 2.2.

When the rest of the upper clause is constructed, the subject (actually just its PF part - lacking $[_{LF} \text{x}]$) can move up to Spec-TP, to check the EPP and the phi features ($[\Phi]$) of the upper T. Since $[_{PF} \text{x}]$ of the subject include $[_{PF} \text{Plural}]$, agreement on the upper verb is not surprising.



At the end, when the entire sentence is constructed and the top CP closes the sentence, everything is spelled-out to both interfaces. But since the phases were misaligned the DP *2 Basques* is pronounced in the upper subject position, where its $[_{PF} \text{x}]$ features end up, and interpreted in the lower subject position (within the scope of *likely*), from where it was sent to LF. Thus we end up with sentence (10a), interpreted as (10b).

- (10) a. *2 Basques are likely to win a medal.*
 b. *It is likely that 2 Basques win a medal.*

(10) is underivable for Sauerland & Elbourne (2002), for whom the *phi*-feature [Plural] spells-out at the lower TP phase where both interfaces are fed. A purely PF movement that follows should not target a specific syntactic position and check the features of the upper T thus triggering plural verbal agreement. Agreement comes with the subject's checking the uninterpretable *phi*-features on T (Chomsky 1995a, 2000). Only a syntactic movement can trigger agreement with the required features. I proposed there are [_{PF} *phi*] features that are not affected by the LF-only phase and can trigger agreement, since a DP with phonologically overt Plural, e.g. in (10), has its [Plural] features visible both at LF and at PF. For non-specific reading, only the DP's [_{LF} x] features get frozen in the lower clause, while the rest of the DP continues the derivation and checks *phi*-features on T in the upper clause with its [_{PF} *phi*] features.

One obvious problem comes to mind. A phase boundary is not only the point where the completed phase is Spelled-Out but also the point where the new phase is started. Thus if the endpoints of the LF and PF phase differ, does it mean the starting points of new phases differ also? Since a phase is defined as a complete cycle including the numeration and derivation, having completely independent phases would suggest completely independent numerations consisting of only PF or only LF related material/features. But that would suggest matching of PF and LF is a result of pure coincidence and seems plainly wrong.

It seems natural to say that at the point of Spell-Out the numeration has to be emptied, but if at the point of Spell-Out derivation is not shipped to both interfaces, the phase is not really completed and maybe the lower numeration could still give items to the derivation, but than it also wouldn't be emptied. Thus, it seems, even at the point where only a partial phase is completed, the numeration has to be completely empty. But if it is empty, than of course the new phase has to bring in items relevant for both interfaces. Thus any partial phase, acts as a starting point for both phases. I see this as a welcome result. Notice that it doesn't matter how much material is being shipped to the interfaces at the point of Spell-Out since a phase regularly accepts items that joined the derivation in a previous phase and moved out. In this particular case the next PF phase would spell-out structure constructed from two enumerations—the lower TP (that was already shipped to LF) and whatever would get constructed on top of it. In a way thus phases do remain parallel and have one to one corresponding, is just that in some cases they don't Spell-Out to both interfaces simultaneously.

2. On two non-standard claims

Two non-standard assumptions were used in order to derive the desired non-specific reading of (10). First, I claimed that phases —stages in the derivation— do not have to be the same for the two interfaces, that a phase in the derivation of the LF representation, does not necessarily have a correspondent PF phase. Second, I claim there are [_{PF} *phi*] features, independent of their LF correspondents, that can also check T's *phi*-features and thus trigger plural agreement on the verb. In this section, we will look at both claims/assumptions in more detail in try to provide independent motivation for them.

2.1. Phases in the derivation

As proposed by Chomsky (2000, 2001), derivation proceeds in stages at the completion of which the constructed structure is no longer accessible for any further operations, which follows from the Phase impenetrability condition (cf. Epstein *et al.* 1998, Chomsky 2000, 2001, Uriagereka 1999). At the next higher phase, the lower phase is sent to the two interfaces — PF and LF. Phase is a self-contained subsection of the derivation, from Numeration → ... → Spell-Out → PF/LF. Since, at the point of Spell-Out, the derived structure is divided into what goes to LF and what goes to PF, we can talk about the structure being composed of these two parts even before Spell-Out. Therefore PIC actually talks about specific kinds of features being unavailable for specific operations. Namely, $[_{LF} x]$ being unavailable for all the operations LF features are involved in and $[_{PF} x]$ being unavailable for the kind of operations they would usually participate in.

Chomsky (2000, 2001) further proposes there are two strong phases in the main structure of the sentence — νP and CP — where a strong phase means a point of Spell-Out — point at which structure is interpreted and features sent to the interfaces. What exactly is sent to the interfaces and how the structure is read or what its importance is not entirely clear. But it is definitely true for the LF that it interprets the whole syntactic structure together with its leaf nodes and their semantic features, since LF interpretation is compositional, meaning, structure is essential. Less clear is this for PF, although it is still accepted that PF phrasing depends on syntactic structure (cf. Chomsky and Halle 1968, Cinque 1993, Truckenbrodt 1999, Kahnemuyipour 2003, Wagner 2004). I leave the question of what exactly participates in the derivation (either actual phonological features that direct our pronunciation or features that are then substituted with lexical items as in Late Insertion, e.g. Marantz 1997) since whatever we decide to adopt we still need some sort of features that are spelled out to PF and ultimately determine what we say.

Phases should also have interface realities — there should be a reflex of phases on the interfaces. According to Chomsky (2001), phases are propositional elements, suggesting that whatever is shipped to LF when a phase is completed forms a proposition at LF (or after the LF goes through all the required LF operations) — proposition can also be seen as a unit of information. Phases also offer a natural point for the calculation of Interpreted Logical Forms (Larson and Ludlow 1993). Interpreted logical forms are units of information and thus nicely correspond to propositional phases. Butler (2003), working in a different semantic framework, gives an account of syntax-semantics interface where every phase corresponds to a quantificationally closed situation. On the PF side, phases are reflected in as phonological units — i.e. prosodic word, prosodic phrase, intonational phrase etc. — they have some level of phonetic independence (cf. Marvin 2002, Marušič 2001).³

³ It should be mentioned here, that it is not the identity of the phase that determines what kind of semantic or phonological unit a particular spelled-out constituent is. Rather, it is the amount and type of the material inside this unspecified chunks of structure that are thrown into the computation at the interfaces that determines whether something is an event, proposition ... or, on the other side, an intonational or a prosodic phrase.

Standardly a phase is said to be both the point of PF and LF Spell-Out — freezing and shipping of the features to the two interfaces is said to happen simultaneously (Chomsky 2001, Legate 2001, 2003). But since phases are reflected as units at the two interfaces and if this is the only way interfaces units can be created, if every PF phase has a correspondent LF phase than every PF unit should have a correspondent LF unit and vice versa. Intuitively, this is not the situation in natural languages. We have both complex semantic constituents that form a single PF unit like compounds and complex PF constituents that form a single simple LF unit — idioms. In addition, it is not clear why we would want to restrict this possibility in a minimally designed grammar if it doesn't follow from anything else.

I will call non-simultaneous phases *Split phases* (shipping material to either only PF or LF interface). Split phases have been proposed and were discussed also by Felser (2004), Wurmbrand and Bobaljik (2003) and Marušič and Zaucer (2004). We will now look at some further evidence for the existence of split phases.

2.1.2. The Slavic FEEL-LIKE construction

Marušič and Žaucer (2004a, 2004b) give a thorough description of the Slovenian FEEL-LIKE construction, in which a single verb (root + affixes) is composed of parts belonging to two different clauses. Following Marušič and Zaucer, this apparently monoclausal construction, (11), is given a biclausal analysis with a hidden matrix predicate, (12). The hidden predicate corresponds to the overt verb that appears in the paraphrase.

- (11) *Gabru se je pilo koktejlje.* (Slovenian)
 Gaber_{MASC,DAT} SE AUX_{Sg} drink_{3,Sg,Neu,Past} cocktail_{MASC,Pl,ACC}
 ‘Gaber felt like drinking cocktails.’

- (12) [_{TP} *Gabru* NON-ACT [_{VP} FEEL-LIKE [_{FP} PRO [_{IP} [_{VP} *drink* [_{DP} *cocktails*]]]]]]]]

The construction is apparently monoclausal, since it only has one overt verb, but as suggested already in the glosses, it is interpreted with two distinct predicates, the pronounced verb and the non-pronounced dispositional element. Based on a variety of arguments, including double temporal adverbials referring to two distinct events denoted by the two predicates, two opposing depictive predicates, suggesting again two events occurring at two different times etc. Marušič and Zaucer conclude that the sentence is covertly composed of two clauses with their own predicates denoting separate events possibly occurring at different times. The whole issue of timing of the two events is important because it suggests two separate LF units and thus two phases. Two LF phases are also suggested by the intensionality of the construction, non-specific reading of the object, possible use of non-referring names etc. Last but not least, the construction has been standardly analyzed with a covert modal element, taking a proposition as a complement (cf. Franks 1995, Rivero & Milojević-Sheppard 2003 among others). If phases create propositions, than the complement of the ‘modal’ is an LF phase even in these monoclausal analyses.

What is important for the present discussion is that unlike its apparent LF structure, the construction's PF structure is quite simple. First, there is no apparent clausal boundary that would prevent scrambling and clitic climbing, as shown in (13).

- (13) a. *Televizijo se je Vidu [gledalo t_i že včeraj]*
 TV_{ACC} SE AUX_{3PSg} Vid_{DAT} watch_{Past,Sg,Neut} already yesterday
 “Vid felt like watching the television already yesterday.”
- b. *Včeraj se jo_i je Vidu [gledalo t_i]*
 Yesterday SE her_{CL,ACC} AUX_{3PSg} Vid_{DAT} watch_{Past,Sg,Neut}
 “Yesterday, Vid felt like watching her/it.” (e.g. television)

Even more revealingly, tense inflection on the only overt verb actually does not belong to this verb since it modifies the disposition rather than the event denoted by the verb. Thus tense morphology belongs to the hidden FEEL-LIKE predicate. Since the FEEL-LIKE predicate is the matrix predicate in this construction, the tense morphology modifying its event also originates in the matrix tense projection. Therefore the tense morphology from the matrix tense, ends up being attached to the lower verb, forming a single word composed of elements from two distinct clauses.

To show this with actual examples, (14), with future tense morphology, signifies a future disposition towards *sitting outside*, not a present disposition towards a future event of *sitting outside*. To express a present disposition, the verb has to be in the present tense, (15).

- (14) *Filipu se ne bo sedelo jutri odzuni.* (Slovenian)
 Filip_{DAT} SE NEG AUX-FUT_{Neut} sit_{Neut} tomorrow outside
 “Filip won’t feel like tomorrow sitting outside.”
 **“Filip doesn’t feel like tomorrow sitting outside in the future.”*
- (15) *Filipu se jutri sedi odzuni.* (Slovenian)
 Filip_{DAT} SE tomorrow sit_{Neut,Pres} outside
 **“Filip won’t feel like tomorrow sitting outside.”*
 “Filip doesn’t feel like sitting outside tomorrow.”

(15) additionally shows that although the verb is given in the present tense it can still appear with a future adverbial. Not surprisingly this is OK. The adverbial modifies the time of the sitting and is located inside the lower clause, which explains why there is no conflict between the present tense on the verb and the future adverbial.⁴

Verb and its tense inflection make up a single word. Note that the verb did not raise out of its position since it is interpreted inside the lower clause, inside the scope of the feel-like predicate (additionally, the verb itself is opaque, for example, it need not refer to an actual event, and one surely can feel-like levitating). With its temporal inflection clearly belonging to the matrix predicate, we have an example of a single word — a single phonological unit (created in a single PF phase) — that is composed of parts belonging to two different semantic units/LF-phases.

⁴ For a discussion on the futurate readings of the present tense and some facts showing this is not related to futurate, see Marušič and Zaucer (2004).

2.1.4. *Infinitives and restructuring*

Non-finite clauses show transparency for clitic climbing and scrambling, but can still induce scopal effects such as non-specific reading of the embedded object. On one hand they show evidence of lesser structure—they lack the CP projection as argued for by Marušič (2003)—while on the other they get interpretations parallel to those of other clausal complements - the complement is a proposition (cf. Wurmbrand 2001, Wurmbrand and Bobaljik 2003, Marušič 2003).

Scrambling from Slovenian finite clausal complements shows A-bar scrambling properties, while scrambling from non-finite complements shows A-properties: it does not trigger WCO, (16b), and it does not trigger total-reconstruction of the scrambled element, (16a).

- (16) a. *Vse punce se je nekdo odločil poklicati po telefonu* __ $\exists > \forall, \forall > \exists$
 all girls REFL AUX someone decided to call_{INF} over phone
 “Someone decided to call all girls”
- b. *Janeza_i je njegov_i oče sklenil poslati* __ *v semenišču*.
 J-ACC_i AUX his_i father decided send_{INF} to theological seminary
 “His_i father decided to send John_i to the theological seminary.”

Clitics can easily climb from Slovenian non-finite complements, (17). Assuming Slovenian second position clitics are positioned in PF (Marušič, in prep), this not only shows there is no CP boundary, but that there is in fact no phase between the matrix clause and the embedded non-finite complement. In particular, it shows there is no PF phase (or at least no more PF phases than there are in a non-embedded sentence).

- (17) *Res sem ji ga sklenil* [PRO opisati __ __]
 really AUX her him decide describe_{INF}
 “I really decided to try to describe him to her.”

Non-finite complementation creates opaque contexts and the clausal complement denotes a proposition, therefore the clausal boundary obviously shows properties of an LF phase. Thus we have a conflict of phases, what appears to be an LF phase, is not a PF phase.

As I showed, we get clear mismatch between phases in other constructions as well. Therefore positing split phases in cases of total reconstruction discussed in this paper doesn't sound as an unsupported assumption anymore.

2.2. On features

2.2.1. *PF plural features*

According to the proposal advanced in this paper, there are [_{PF} *phi*] features that can trigger verbal agreement. Although this is not an entirely controversial claim, it would be still nice to have some independent evidence for it. First note that Sauerland & Elbourne (2002) discuss a kind of plurality found in British English that has no morphological exponence but can nevertheless trigger plural agreement on the verb, (18).

- (18) *The government are ruining this country.* (Sauerland & Elbourne 2002: (13b))

For them, this is a case of the purely semantic plural feature [_{LF} Mereology] triggering verbal agreement. In a similar way as there are purely semantic [_{LF} *phi*] features, one can also think of purely phonological ones. I claim there are (gender and) number features that have morphological reflex but cannot or may not be interpreted at LF, and that in addition, they are part of the syntactic derivation and have the same effect on verbal agreement as the more common LF interpretable *phi*-features.

In Slovenian, the verb has to agree with the nominative subject in person, number, and gender. Since person is not a feature of the nominals (apart from personal pronouns) and gender is a bit tangential to the main point, they will be left out.

First note that there are a lot of plural place names in Slovenian. In such cases we clearly refer to a single individual —the town or village having the particular name— but the nominal morphology and the verbal agreement it triggers are both plural. Such names are for example: *Helsinki, Abitanti, Banjsce, Baske, Bate, Benetke, Brezje, Firenze, Gorje, Jesenice, Lohke* etc. Although, they are plural only phonologically, only plural personal pronouns can be used to refer to them.

- (20) *Lohke so majhne. Ampak večjega kraja od njih na Banjscah preprosto ni.*
 Lohke_{PL} are small. But there is no place bigger than them on Banjsce.

Apart from these peculiar place names, there are also plenty of pluralia tantum nouns, (21), that trigger plural agreement on the verb, (22), but refer to a single entity/item and fail to show any signs of semantic plurality. These kind of nouns can be used also with the numeral *one* (interpreted not as an indefinite but as a true numeral counting the number of items referred to), (22).

- (21) *hlače, očala, skarje, jetra, sanke* (Slovenian)
 trousers, glasses, scissors, liver, sled
- (22) *Razbila so se mi samo ena očala.* (Slovenian)
 broken AUX_{PI} REFL me only one glasses_{PI}
 “Only one pair of glasses have fallen apart.”

In addition, they cannot be used with a floating quantifier, which can only be used with semantically plural arguments, (23). Thus we can safely conclude these nouns are, despite their plural morphology, semantically simply not plural. Or at least not plural in the usual sense.

- (23) **The trousers have all been very dirty, since M^a José doesn't want to take them off*

Semantic plurality should license also the use of reciprocals. But again, as we see in (24), such use is not grammatical with the pluralia tantum nouns, when they refer to a single item. The plural that is realized with plural morphology is thus not spelled-out to the LF interface, suggesting we are dealing with a PF-only plural feature — [_{PF} Plural].

- (24) **Svoje edine hlače sem drgnil ene ob druge.*
 *I rubbed my only trousers against each other.

To confirm that we are really dealing with the same kind of plural features here and in the non-specific plural examples in (7), we have to show these same nouns with [_{PF} Plural] triggering plural agreement can also trigger plural agreement when they are

To license *all*, semantic plural features have to be present in the movement, which means the DP cannot be sent to LF inside the lower phase. (28) is comparable in its effect to the British English facts reported by Sauerland & Elbourne (2002), given in (8).

3. Further Issues

3.1. On apparent optionality

Both sentences in (1) are actually ambiguous between the specific and non-specific reading of the subject. The apparent problem for the argued optimal design of the theory is optionality of the specific reading. Sauerland & Elbourne (2002) claim the specific reading of (29), the one presented as $\exists > \textit{likely}$, comes from stem-movement of the DP to the upper position. But how can such stem-movement be allowed, how can the DP move after it was spelled out? Obviously it has to move to the edge of the phase to avoid that, but how can it move if it doesn't have to? How can such movement be optional?

(29) *A Bull is likely to run over a tourist in Iruña.*

This problem is actually shared by all approaches taking specific reading as a result of the indefinite moving out of the scope of *likely*. I do not want to go into this discussion too far. The same “optionality” is present in all cases of the indefinite taking wide scope discussed by Fodor and Sag (1982). I do not give any definite answer but only try to avoid stipulating any kind of optionality. Having or not having a phase cannot be optional. Nor does it sound acceptable to treat as optional the moving of the DP to the edge of the LF phase.

Specific interpretation can be seen also as a special case of the non-specific one. This is either done by saying indefinites are ambiguous between a true quantifier and a choice function (Kratzer 1997) or that their restriction, being a set, can be a singleton resulting in a so-called *singleton indefinite* that behaves just like a referential noun phrase (Schwarzschild 2002, von Stechow 2000). In both cases the specific/referential reading is not derived by movement, but is a result of some property of the indefinite quantifier and the consequences this property has on the form of the LF representation.⁵

3.2. Universal quantifiers and complex likely predicates

Sentences with a universal quantifier in subject position, (30), are a bit tricky — judgements are not really clear, but some sort of agreement can nevertheless be established. The universal quantifier can definitely be interpreted outside the scope of *likely*, but the reading with the quantifier taking narrow scope is less clear.

(30) *Every Basque cyclist is likely to be among the top 10* $\forall > \textit{likely}$, $\textit{likely} > \forall$

⁵ Note that in some cases, like the one where [mereology] of the collective DP triggers plural agreement on the matrix T/verb, the subject does appear to move to the matrix clause in (stem) syntax and not get shipped to LF in the lower clause. So there need to be some optionality allowed here, and if it is allowed in this case, we could also let it deal with other specific readings. A case of subject moving to the matrix clause (in that case obligatorily) might be also required for cases with more than one raising verb in a sequence (cf. section 3.5).

The confusion comes from two sources. One is the meaning of the predicate *likely*, the other is the entailing relations among relevant situations. *Likely* is commonly interpreted to mean something like “with the likelihood (much) higher than 50%”. But this kind of interpretation only applies to a situation where we are comparing two possible situations (e.g. in coin tossing, the result is either head or tails). In such a case, the situation that is likely has a higher probability of occurring than the other one —is more likely to occur. But in case there are more than just two possible outcome situations, likely can either mean “more likely than not” or “more likely than any other single situation”. In the first case, the actual probability would still need to be higher than 50%, but in the latter case the actual probability could be smaller than 50% as long as it is higher than any other probability of a single situation. With this in mind we can return to the interpretation of (30).

There seem to be two possible interpretations of the narrow scope universal quantifier. We are either comparing probabilities of single situations (situation with all Basques among the top ten vs. situation 1 with a particular (non-Basque) cyclist in place of Iban Mayo, situation 2 with a particular (non-Basque) cyclist in place of Haimar Zubeldia, etc.) or else we are comparing the likelihood of occurrence of a situation with all Basques among the top ten versus the probability of its non-occurrence (which is the sum of probabilities of all situations where it is not the case that all Basques made it to the top 10). The thing is that the first interpretation follows from the wide scope interpretation of the universal quantifier, while the second one entails the wide scope interpretation of the universal quantifier. Let me explain this a bit further.

We'll start with the narrow scope reading of the universal quantifier and >50% chance interpretation of *likely*. If it is true that it is likely that all Basques finish the race among the top 10, then it is also true for every Basque that he is likely to end up among the top 10. The individual probabilities would get very high and would be definitely different from the probability of the situation where all get placed among the top 10, but since *likely* doesn't specify the degree of likelihood, such entailments are allowed. Thus, the narrow scope reading is just a special case of the wide scope. Notice that under such interpretation of the sentence, the sentence obviously cannot be true if there are more than 10 Basques competing, since 11 participants cannot have a chance higher than 50% to be among the top 10 (just like the two sides of a coin cannot be both likely for the coin to land on).

And if it is true for every Basque that he has the property of being likely to be among the top ten, then we can say that the situation where all of them are placed among the top 10 is more likely than any other single situation in which an outsider occupies a position where a favorite could be. Thus the two possible narrow scope readings seem to be reducible to the wide scope interpretation of the universal quantifier. Since the second possibility is possible but not necessarily actual, I conclude the universal quantifier in (30) only has wide scope.

The same apparent ambiguity, as observed in (30), seems to be also available for other quantifiers, e.g. *most* in (31). Here again, under the >50% chance interpretation of *likely*, if it is true that it is likely that most Basques will finish among the top 10, then it is also true for most Basques that they have the property of being likely to finish among the top 10 (and there are at most 19 Basques participating). Thus the narrow reading entails the wide scope reading suggesting the narrow scope reading is again just a special case of the wide scope one.

(31) *Most Basques cyclists are likely to be among the top 10* most > likely, likely > most

The situation changes with more complex *likely* predicates that specify the degree of likelihood. Total reconstruction of an indefinite is possible over a more complex 3% *likely*, (32). Of course, both sentences in (32) have also the specific reading, which can be said to have either the wide scope indefinite or is triggered by singleton indefinites.

(32) a. *A Basque is 3% likely to win a gold medal* 3% likely > \exists
 b. *A Basque is somewhat likely to win a gold medal* s/w likely > \exists

But when a complex *likely* predicate is combined with a universal (or any other quantifier) the ambiguity of (31) disappears.⁶ (33) is unambiguous with the universal quantifier taking wide scope. The “reconstructed” interpretation from (31) is impossible.

(33) a. *Every Basque is 3% likely to be among the top 10* \forall >3% likely, *3% likely > \forall
 b. *Every Basque is s/w likely to be among the top 10* \forall >s/w likely, *s/w likely > \forall

Firstly, we have to note that the cases with the universal quantifier are genuinely different. As we have seen for (30), unlike with indefinites that get reconstructed, the universal quantifier is interpreted outside of scope of *likely*. We saw earlier how the narrow scope reading of the universal was just a special case of the wide scope reading and noted that this kind of entailment from narrow scope to the wide scope was possible because *likely* alone is a predicate of unspecified probability (as long as it is (much) bigger than 50%). Notice that the same entailment/inference is not available with a specific degree of likelihood. If it is 3% likely that everybody will be among the top 10, that doesn't entail mean that everybody has the property of being 3% likely to be among the top 10. If the probability for the situation in which everybody is positioned among the top 10 is 3%, then the probabilities for the individuals to get among the top 10 are different, but regardless of the actual number, they are not 3%, they should be much higher.

Although ambiguity was observed in (30), it is not surprising we don't find it in (33), since, as explained, (30) is ambiguous because of the undetermined *likely*.

3.3. Elided facts

Ellipsis is standardly taken to be licensed by some form of LF sameness of the antecedent and the elided part (Merchant 1999). A combination of LF sameness and the analysis presented here, where the high surfacing subject gets interpreted low because that is where it is located at LF, predicts ellipsis of a conjunct, should not really be possible, (34a) gets interpreted as (34b).

(34) a. *A Swiss is likely to be among the top 10 and a Czech is likely to be among the top 10 too.*
 b. It is likely for a Swiss to be in the top 10 and it is likely for a Czech to be in the top 10.

But as shown in (35), ellipsis in such cases is possible. The two indefinites are both read non-specifically, which according to the proposal advanced here is a result

⁶ This was pointed out to me by Andrew Nevins.

of the early LF spell-out of the lower clause. But if the subject is really interpreted in the lower clause, the elided clause is not LF identical to its antecedent.⁷

(35) *A Swiss is likely to be among the top 10 and a Czech is too.*

But we should ask ourselves if ellipsis is really conditioned solely by LF identity? I do not offer an answer, I simply want to point out to another case of ellipsis where LF identity seems to be violated. Consider the sentences in (36). Since the reflexive and the possessive in the subject are co-indexed with the subject of the embedded clause, the subject seems to be interpreted in some lower position inside the embedded clause.

(36) [*Three pictures of himself*]_{k,i} *is a lot for Peter_k to take t_i*
 [*Three pictures of himself*]_{k,i} *are a lot for Peter_k to take t_i*
 [*Three pictures of his teacher*]_{k,i} *are a lot for anybody_k to take t_i*

Since reflexives and quantifiers co-indexed with pronouns are commonly used as diagnostics for reconstruction and as a general interpretation location search, I take the claim that (36) involves reconstruction to be correct. Note that this is not a *though* construction, which is standardly taken to lack reconstruction properties (*Everyone is thought to please* ≠ *It is thought to please everyone*). Without going any further into this construction, let me just point out that this construction also allows ellipsis, presumably without LF identity.

(37) [*3 chickens is a lot to eat*] and [*3 melons is too*].
 [*3 chickens are a lot to eat*] and [*3 melons are too*].
 [*3 pictures of himself is a lot for Peter to take*] and [*3 pictures of his mother is too*].

Since this construction exhibits reconstruction properties and allows ellipsis of a constituent that is not LF identical with its antecedent, I conclude (35) does not represent a counterexample for the analysis presented in this paper.

3.4. A sequence of raising verbs

Aoun (1985) gives examples with two raising verbs one of which is *likely* and claims the indefinite subject cannot get the lower (“reconstructed”) interpretation, (38). This appears to be problematic for the approach advanced here. If *likely* induces an LF phase, the indefinite should get interpreted in the lower clause regardless of how many additional raising verbs are merged into the structure.⁸

(38) *Some politician seems to be likely to address John’s constituency.* (Aoun 1985: 84, (12))

Although this objection seems valid, it appears that the problem comes in with the raising verb *seem*, which appears to behave strangely, rather from our understanding of *likely*. First notice that such a sentence is impossible with the indefinite in its supposed base position, (39a). There is nothing wrong with expletive occupying the topmost subject position when there is just one rising verb in the sentence, (39b,c).

⁷ This was pointed out to me by Thomas Leu.

⁸ This was pointed out to me by Richard Larson.

- (39) a. ?* *There seems to be likely to be someone in the room.* (Aoun 1985: 81, (1b))
 b. *There is likely to be someone in the room.*
 c. *There seems to be someone in the room.*

Although (39a) isn't completely bad, it is still significantly worse than other comparable examples. Regardless of (39a), what is really revealing about the weirdness of *seem* is that it is not the case with all raising verbs that they block reconstruction (or trigger antireconstruction a term used by Wurmbrand and Bobaljik 2004). Actually even if we put more than two in a sequence it seems that the subject indefinite can still get the non-specific/narrow scope interpretation, as is the case in (40).

- (40) *A politician is expected [to give a speech at the convention]*
A politician appears [to be likely [to give a speech at the convention]]
A politician is expected [to appear [to be likely [to give a speech at the convention]]]
A politician was believed [to be likely [to give a speech at the convention]]

Without any answer on the account of (38), I conclude that since some raising verbs follow predictions, there is something about *seem* we are yet to understand.

4. Conclusion

When the derivation reaches a phase (or the next higher phase) features do not necessarily get shipped to both interfaces (PF and LF), since a phase can be an exclusively PF or exclusively LF phase. Features not spelled out at partial/split phases continue the derivation and can check the uninterpretable *phi*-features of T. A DP has both LF and PF *phi*-features semantic/interpretable at LF, like the number feature of a DP, should be part of what is shipped to PF. Allowing split phases opens the door also for a syntactic derivation of quantifier raising and other covert movements.

References

- Aoun, J., 1985, *The Grammar of Anaphora*. Cambridge, Mass.: MIT Press.
 Butler, J., 2003, "Phase structure is event structure. Ms. University of York". Presented at *MIT Workshop on Phases*. January 2003.
 Chomsky, N., 1993, "A minimalist program for linguistic theory". In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by K. Hale and S. J. Keyser, 1-52. Cambridge, Mass.: MIT Press.
 —, 1995a, *The Minimalist Program*. Cambridge, Mass.: MIT Press.
 —, 1995b, "Bare Phrase Structure". In G. Webelhuth (ed.), *The Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell, pp. 383-439.
 —, 2000, "Minimalist Inquiries: The framework". In *Step by step: Essays on minimalism in honor of Howard Lasnik*, R. Martin, D. Michaels, and J. Uriagereka (eds.), 89-155. Cambridge, Mass.: MIT Press.
 —, 2001, "Derivation by phase". In *Ken Hale: A life in language*, ed. M. Kenstowicz. Cambridge, Mass.: MIT Press.
 —, and M. Halle, 1968, *The sound pattern of English*. New York: Harper and Row.
 Cinque, G., 1993, "A null theory of phrase and compound stress", *LI* 24: 239-97.

- Epstein, S. D., E. M. Groat, R. Kawashima, and H. Kitahara, 1998, *A Derivational Approach to Syntactic Relations*. Oxford: Oxford U.P.
- Felser, C., 2004, "Wh-copying, phases, and successive cyclicity", *Lingua* 114: 543-74
- von Fintel, K., 2000, *Singleton indefinites* (re. Schwarzschild 2000). Ms. MIT.
- Fodor, J. D. and I. Sag, 1982, "Referential and Quantificational Indefinites", *Linguistics and Philosophy* 5: 355-98.
- Franks, S., 1995, *Parameters of Slavic Morphosyntax*. New York: OUP.
- Kahnemuyipour, A., 2003, "Towards a Phase-Based Theory of Sentential Stress". Ms. U. of Toronto. Presented at the *MIT Workshop on EPP and Phases*, January 2003.
- Kratzer, A., 1997, *Scope or Pseudoscope? Are there Wide-Scope Indefinites?* Ms. UMass Amherst. Obtained from <http://www.semanticsarchive.net/> Aug. 2004.
- Larson, R. & P. Ludlow, 1993, "Interpreted Logical Forms", *Synthese* 95: 305-55.
- Legate, J. A., 2001, *Some interface Properties of the Phase*. Ms. MIT/Harvard.
- , 2003, *Identifying Phases*. Ms. Harvard. To appear in: M. McGinnis and N. Richards (eds.): *Proceedings of the MIT Workshop on Phases*, January 2003. *MIT Working Papers in Linguistics* 47.
- Marantz, A., 1997, "No escape from syntax. Don't try morphological analysis in the privacy of your own lexicon", *Penn Working Papers in Linguistics* 4.2: 201-25.
- Marušič, L., 2001, "The placement and scope (f-projections) of sentence accent in Slovenian". Ms. Stony Brook U.
- , 2003, "CP under Control". Ms. Stony Brook U. Presented at FDSL 5.
- , In prep., "A prosodic account of clitic placement". Ms. Stony Brook University.
- & R. Zaucer, 2004a, "On the Intensional FEEL-LIKE Construction in Slovenian: A case of a phonologically null verb". Ms. Stony Brook U. & U. of Ottawa.
- & —, 2004b, "The intensional FEEL-LIKE construction in Slovenian: A case of a phonologically null verb". Ms. Stony Brook U. and U. of Ottawa. [http://www.ic.sunysb.edu/Stu/fmarusic/pub/FEELLIKE_Draft_June_2004.pdf]
- Marvin, T., 2002, *Topics in the stress and syntax of words*. PhD thesis, MIT.
- May, R., 1985, *Logical Form: Its structure and derivation*. Cambridge, Mass: MIT Press.
- Merchant, J., 2001, *The syntax of silence*. Oxford: Oxford U.P.
- Nissenbaum, J., 2000, *Investigations of Covert Phrase Movement*. PhD Thesis. MIT, Cambridge Mass.
- Rivero, M.-L. & M. Milojević Sheppard, 2003, "Indefinite Reflexive Clitics in Slavic: Polish and Slovenian", *NLLT* 21.1: 89-155.
- Saito, M., 1989, "Scrambling as semantically vacuous A'-movement". In *Alternative conceptions of phrase structure*, ed. by M. Baltin and A. Kroch, 182-200. Chicago: University of Chicago Press.
- Sauerland, U. & P. Elbourne, 2002, "Total Reconstruction, PF Movement, and Derivational Order", *LI* 33.2: 283-319.
- Schwarzschild, R., 2002, "Singleton Indefinites", *JS* 19.3.
- Trockenbrodt, H., 1999, "On the Relation between Syntactic Phrases and Phonological Phrases", *LI* 30.2: 219-255.
- Uriagereka, J., 1999, "Multiple Spell-Out". In S. D. Epstein and N. Hornstein (eds.) *Working Minimalism*. Cambridge, Mass. MIT Press, pp. 251-82.
- Wagner, M., 2003, "Asymmetries in Prosodic Domain Formation". Ms. MIT, to appear in: M. McGinnis and N. Richards (eds.): *Proceedings of the MIT Workshop on Phases*, January 2003. *MIT Working Papers in Linguistics* 47.
- Wurmbrand, S. & J. D. Bobaljik, to appear, "Domains of Agreement", *NLLT*.

A TYPOLOGICAL APPROACH TO THE ORDERING OF ADVERBIALS: WEIGHT, ARGUMENTHOOD AND EPP¹

Roberto Mayoral Hernández

University of Southern California and Universidad Autónoma de Madrid

1. Introduction

The analysis of adverbial expressions has been an area frequently forgotten in the linguistics literature. Besides the wide range of grammatical classes that can fulfill this category, such as AdvP, PP, NP, etc., many of these expressions appear in different positions in the sentence. However, there have been some relevant attempts to elaborate a coherent description of this grammatical class, providing a unifying account. In this respect, Cinque (1999) attributed the different kinds of adverbials to the existence of multiple aspectual projections, based on their position with respect to each other. The underlying position of the different types of adverbials was claimed to be universal. Nevertheless, this method was unable to explain the alternation in positions that frequency adverbials show. In their grammar, Fernández Lagunilla and Anula Rebollo (1995) also derived the position of adverbials from the existence of different syntactic projections in different positions in the tree. Both analyses would require the incorporation of a complicated movement theory to successfully describe the multiplicity of positions in which adverbials can appear and both approaches fail in this respect.

Apart from the two purely syntactic systems described above, there have been more descriptive analyses of this subject, like the extensive description presented in Rodríguez Ramalle (2000). In this paper, I will adopt a variationist approach to describe the ordering of frequency adverbials in Spanish. This approach is based on the belief that different factors interact with each other to bring about the final collocation of constituents in a given sentence. Specifically I will consider three factors, building on Mayoral Hernández (2004): (i) argumenthood, (ii) type of verb and (iii) the position of a co-occurring agreeing subject. The results will be

¹ This research was partially supported by a 'La Caixa' Fellowship. I would like to thank Jack Hawkins, Carmen Silva-Corvalán and Mario Saltarelli. I would also like to thank the contributions and comments made by Violeta Demonte, Rebeka Campos, Abe Kazemzadeh, and Tomomasa Sasa. However, any possible mistake is my own responsibility.

related to the Extended Projection Principle (EPP) and will be assigned typological relevance.

The results obtained in Mayoral Hernández (2004) indicated that the argumental condition of the constituents that co-occur with a frequency adverbial in postverbal position did not have an influence on the collocation of postverbal elements, which goes against the existing literature (Hawkins 1994, 1999, 2000..., Wasow and Arnold 2003). However, these contradictory results might be due to the fact that the definition of argument adopted in Mayoral Hernández (2004) was too wide, including many PPs and other XPs. Here I will adopt a more restricted definition of argumenthood to find out the relevance of this factor.

Another factor that might trigger the adverbial alternation is the nature of the verb. Mayoral Hernández (2004) showed that the position of overt agreeing subjects in Spanish determines to a large extent the choice between preverbal and postverbal adverbials. In this paper I will analyze if the subject effects on position are derived from the type of verb rather than from the subject position itself. While accusative, copulative or unergative verbs will tend to have preverbal subjects, unaccusatives will show a higher percentage of postverbal subjects. In this way, the factor that triggers the alternation would be the nature of the verb, while the subject position would only be a secondary effect.

The data used for this research have been obtained from the online corpus CREA (Corpus de Referencia del Español Actual). The use of the statistical tool SPSS and the Pearson's Chi Square test will ensure the statistical significance of this analysis.

2. Alternation of frequency adverbials

The adverbial expressions that denote frequency in Spanish can be realized by different syntactic and morphological categories. Thus they can appear as DPs or NPs (1), AdvPs (2), PPs (3), etc.

- | | |
|---|---|
| (1) <i>Todos los martes</i> Juan come patatas.
Every Tuesday Juan eats potatoes. | (2) Juan come patatas <i>frecuentemente</i>
Juan frequently eats potatoes. |
| (3) Juan come patatas <i>en muchas ocasiones</i> .
Juan eats potatoes on many occasions. | |

Moreover, frequency adverbials can appear in different positions in the sentence without altering necessarily the meaning of the sentence, although they cannot appear in the middle of a different XP, for example in between a determiner and the modified noun, as sentence (4) shows.

- (4) *Juan habla a sus *todos los martes* padres.
*Juan talks to his every Tuesday parents.

There are four possible positions that will be analyzed in this paper. Examples (5) to (8) have been extracted from the corpus CREA:

- (5) Before a co-occurring XP in preverbal position
- '**Frecuentemente** los miembros de las comunidades reciben cursos de protección ambiental.'
 - Frequently the members of the communities receive courses on environmental protection.
- (6) After any co-occurring preverbal XP:
- 'Los agentes del SIN **frecuentemente** realizan redadas en empresas...'
 - The SIN agents frequently carry out raids in companies...
- (7) Immediately adjacent to the right of the verb:
- 'la actividad del citado empresario trasciende **frecuentemente** el mero aspecto comercial'
 - The activity of the aforementioned businessman frequently transcends the merely commercial aspect.
- (8) Following any co-occurring postverbal XP, at sentence final position:
- 'La situación ha sido muy tensa **frecuentemente**'
 - The situation has frequently been very tense.

In the previous examples the position of the adverbial can be changed without altering the meaning. However, Kovacci (1999) noted that there are some asymmetries between postverbal and preverbal frequency adverbials that can cause a change in meaning. For example, she explained that postverbal adverbials are circumstantial, so they behave like adjuncts. Also she mentioned that a sentence with postverbal adverbial, such as (9a), would imply the text without it (9b) and could be paraphrased using *como* "how" or *cuando* "when", like in (9c). The following examples show these properties:

- (9) a. Mis amigos comen patatas frecuentemente
My friends eat potatoes frequently
- b. Mis amigos comen patatas
My friends eat potatoes.
- c. Es frecuentemente cuando/como mis amigos comen patatas.
It's frequently when/how my friends eat potatoes.

Following Kovacci (1999), preverbal adverbials modify the whole sentence, while the sentences including them do not imply the text without them and cannot be paraphrased using *como* "how" or *cuando* "when". However, in (10) we can see that a sentence containing a preverbal adverb (10a) does entail the same sentence without the adverb (10b) and can be paraphrased using *como* "how" or *cuando* "when" (10c).

- (10) a. Frecuentemente mis amigos comen patatas
My friends eat potatoes frequently"
- b. Mis amigos comen patatas
My friends eat potatoes"
- c. Es frecuentemente cuando/como mis amigos comen patatas
It's frequently when/how my friends eat potatoes.

The previous examples show that there is not a necessary change in meaning derived from the alternation. It is also necessary to bear in mind that even in the cases

where there might be a slight change in meaning, due to quantifier scope for example, abundant research has shown that ambiguity avoidance is not relevant when accounting for ordering alternations (Hawkins 2000, Wasow and Arnold 2003). It might be possible to get different meanings from certain sentences whose only difference is the position of the adverbials, but the examples in (10) show that for most native speakers of Spanish there is no difference in meaning derived from the adverbial alternation.

Now let's imagine that some native speakers might be able to get a difference in meaning between (9a) and (10a), opposing the most frequent judgments. This would only show that there is variation when it comes to the interpretation of a given sentence, which would make the current analysis the most appropriate, since the variationist approach is designed to explain variation. Thus, an account based on the existence of meaning differences associated to different positions would be difficult to hold.

Finally, I must acknowledge that some sentences, as Kovacci (1999) noted, do not leave place for variation and can only be interpreted in a single way because of different syntactic phenomena. However, this research is not concerned with this kind of single meaning sentences derived from syntactic constraints.

3. Hypotheses

In order to give an explanation to the variable ordering of frequency adverbials in Spanish, I have tested the validity of three hypotheses: *argumenthood*, *type of verb* and *subject position*. Following Mayoral Hernández (2004) I have also added the *weight* hypothesis to make the current analysis more complete, although no further comments on this issue have been made, accepting previous results.

3.1. Argumenthood

In the linguistic literature dealing with constituent ordering (Hawkins 2000, Wasow and Arnold 2003,...) lexical dependencies appear as a factor that is able to determine the collocation of postverbal constituents. There is a strong preference for place arguments immediately adjacent to the verb, as long as there are not weight effects involved. Thus sentence (11a) would be preferred rather than (11b).

- (11) a. John waited for his mother in the rain.
 b. John waited in the rain for his mother.

However, Mayoral Hernández (2004) showed that the collocation of frequency adverbials with respect to other co-occurring XPs in postverbal position was not influenced by argumenthood, which contradicts previous research. Nevertheless, this fact might be due to the definition of argumenthood adopted in his analysis. Following Hawkins (2000), he considered arguments not only direct objects (DOs) but also all the PPs whose interpretation depended on the meaning of the verb, or vice versa, using the tests in (12) and (13):

- (12) Verb entailment test
 "If [X V PP PP] entails [X V], then assign V_i . If not, assign $V_{d...}$." (op. cit: 242)

In the previous definition, V_i means ‘independent verb’, which is a verb whose interpretation does not depend on the appearance of any other element, while V_d refers to a dependent verb that needs a PP to be interpreted. PPs that are necessary for the interpretation of the verb are considered arguments if one follows this definition.

(13) Pro-verb entailment test

If [X V PP] entails [X Pro-V PP] or [something Pro-V PP] for any pro-verb sentence listed below, then assign P_i . If not. Assign P_d .

Pro-verb sentences: X *did something* PP; X *was* PP; *something happened* PP; *something was the case* PP; *something was done (by X)* PP. (op. cit: 242-243).

The term *dependent preposition* (P_d) refers to a preposition whose interpretation depends on the meaning of the verb, while *independent preposition* (P_i) is used to indicate that the PP headed by that preposition is independent from the event expressed by the verb. Dependent prepositions were considered arguments in Hawkins (2000) and Mayoral Hernández (2004). Following the previous definition, since *John played on the playground* entails *John did something on the playground*, then *on the playground* would be P_i and would not be considered an argument.

The consequence of having adopted this wide definition of argumenthood is that a high percentage of postverbal PPs were considered arguments, which might be the reason why Mayoral Hernández (2004) did not find a statistically significant difference between arguments and adjuncts.

In order to avoid the previous issue only the elements in (14) a-c have been considered arguments. The following variants were analyzed:

(14) Both in preverbal and postverbal position:

- a. Subject: argument that agrees with the verb.
- b. Direct and indirect objects.
- c. Predicates: attributive (i) and predicative (ii) complements.
 - i. *Juan es el doctor* ‘John is the doctor’ or *Juan es agradable* ‘John is nice’
 - ii. *Juan viene cansado* ‘John comes tired’
- d. Other XPs: PPs that are not IOs, and CPs that are not Subject, DO or IO.
- e. No XP

Thus, the argument hypothesis predicts that the restrictive interpretation of argumenthood adopted here as shown in (14) will influence the ordering of frequency adverbials, unlike Hawkins’ (2000) definition.

3.2. Subject presence

Mayoral Hernández (2004) showed that the presence of an overt subject influences the choice between preverbal and postverbal adverbials. It seems that when there is an overt preverbal subject, frequency adverbials will tend to appear in postverbal position. However, when the subject is postverbal or it is omitted, adverbials will normally appear in preverbal position. Because of the data provided in his analysis, Mayoral Hernández (2004) states that the position of subjects determine the choice between preverbal and postverbal adverbials. Thus the subject hypothesis

predicts that agreeing subjects will tend to appear in complementary distribution with frequency adverbials.

However, the literature dealing with unaccusativity in Spanish has related the appearance of certain postverbal subjects to the unaccusative nature of the verb. Thus, the presence of postverbal plural subjects in Spanish has been claimed to be an exclusive feature of unaccusative verbs. Also, the theme properties associated with unaccusative subjects and their underlying complement of V position could make us think that unaccusative verbs, but not unergatives and transitives, should show a higher percentage of postverbal occurrences in languages that allow for this position, like Spanish. Because of this, one might be led to think that the position of overt subjects is not the relevant factor that could determine the collocation of adverbials (*contra* Mayoral Hernández 2004), but a derived one. The type of verb would, therefore, be the relevant factor that determines the choice between preverbal and postverbal positions, while the position of overt subjects could be easily derived from the argument structure of verbs.

Summarizing, if the kind of verb was in fact the factor that determines subject position, Mayoral Hernández's (2004) claim might be inadequate. However, if the kind of verb is not the determining factor the current analysis would support an EPP effect (Fernández Soriano 1999), where the adverbials behave as subjects.

3.3. Type of verb

As I mentioned in the previous section, unaccusative verbs in Spanish have the peculiarity of allowing for the occurrence of postverbal plural subjects with no determiner. The linguistics literature dealing with this matter associates this feature to the underlying complement of V position of unaccusative subjects. Because of this underlying position, it would also be plausible to imagine that postverbal subjects should frequently appear with unaccusative verbs, but not with transitive or unergative verbs. Mayoral Hernández (2004) showed that overt subjects and frequency adverbials tend to appear in complementary distribution. Therefore, the type of verb could be the factor that determines the choice between preverbal and postverbal positions, and not the position of the subject.

In order to analyze the relevance of the type of verb factor, the following classes have been adopted:

1. Transitive verbs, which are those that take an overt direct or indirect object.
2. Following Mendikoetxea (1999), unaccusative verbs are those that can be included in the following groups:
 - 2.a. Verbs of change of state or location, such as *abrir(se)* "open", *hundir(se)* "sink", *caer* "fall", *florecer* "bloom",...
 - 2.b. Verbs of appearance or existence, such as *aparecer* "appear", *llegar* "arrive", *existir* "exist", *venir* "come", *suced* "happen",...

When a verb could not clearly be inscribed in the unaccusative class French and Dutch were used for feedback, since both languages show overt unaccusative morphology (use of auxiliaries *être* and *zijn* "to be")

with unaccusatives and *avoir* and *hebben* ‘to have’ with transitive and unergative verbs in French and Dutch respectively).

3. Reflexive verbs. Spanish reflexive pronoun ‘se’ precedes verbs with reflexive meaning.
4. Intransitive verbs are those with no overt object and agentive subject.
5. Copular verbs are those that take an attributive or predicative complement.
6. Impersonal verbs are characterized by their inability to appear with an overt subject.
7. Verbs with passive morphology.
8. Pronominal passive. The use of the pronominal clitic *se* can make a verb with active form acquire a passive meaning. For example, a sentence like (15a) could have an unaccusative interpretation with no overt agent, like in (15b), or a passive interpretation with a covert agent/cause (15c). Sentences like (16a) can only have a passive interpretation, since verbs such as *construir* ‘to build’ always imply the presence of a volitional agent.

- | | |
|---|---|
| (15) a. El barco se hundió.
The ship sank. | b. El barco se hundió solo.
The ship sank by itself. |
| c. El barco se hundió para simular un ataque pirata.
The ship was sunk to imitate a pirate attack. | |
| (16) a. Se construyó un puente.
A bridge was built. | b. *El puente se construyó solo.
The bridge built by itself. |

Therefore, the type of verb hypothesis predicts that unaccusative verbs will be characterized by a higher appearance of postverbal subjects, which will imply the appearance of preverbal adverbials. Unergatives (intransitives), transitives, and copulas will have a higher occurrence of preverbal subjects, which will imply postverbal adverbials. Pronominal passives will have a higher appearance of postverbal subjects, in the same way as unaccusatives... However, this paper will only be concerned about the position of adverbials in sentences with unaccusative, transitive, copulas and unergative verbs. If the previous predictions hold, we might be able to derive the position of adverbials without the need to consider the position of overt subjects.

3.4. Weight

Following Hawkins (1994, 1999, 2000, 2001), weight will be determined by the number of words that a certain constituent has. A higher number of words increases the weight of the constituent. The concept of weight is linked to universal processing constraints, since a higher number of words would increase the number of syntactic nodes that have to be processed before a certain constituent can be interpreted. This analysis is based on the belief that syntactic constituents can be interpreted when the head has been mentioned. The weight hypothesis predicts that (17a) will be easier to process than (17b) because there is a smaller number of words (or nodes) that need to be processed to interpret the sentence.

(17) a. John **waited for** Peter **in** the dark but moonlit night.

1 2 3 4

b. John **waited in** the dark but moonlit night **for** Peter.

1 2 3 4 5 6 7 8

Therefore, the weight of the different constituents will interact with each other and the heavier ones will tend to appear at the end of the sentence. When there are differences in weight between two different constituents, the weight hypothesis predicts that the heavier element will tend to appear at the end of the sentence in a higher percentage of occurrences.

In this paper, we will adopt Mayoral Hernández's (2004) results about weight, and no further comment will be added. However, this factor has been included in the present research to provide a more generalizing and complete analysis.

4. Tests

The aim of this paper is to provide a statistically supported analysis of frequency adverbial alternation in Spanish. A total number of 1,033 Spanish sentences, obtained from the online Corpus de Referencia del Español Actual (CREA), were analyzed for the purpose of this research. Three adverbs were selected for this study: *frecuentemente* "frequently", *en muchas ocasiones* "on many occasions" and *en más de una ocasión* "on more than one occasion" because they represent instances of adverbials with different weights, ranging from one to five words. These sentences were annotated using the coding in (18).

- (18) 1. Dependent variable: Position of adverbs:
- Before XP in Prev. b
 - Adjacent to the left of the verb 1
 - After XP in Postv. a
 - Adjacent to the right of the verb d
2. Weight of co-occurring postverbal XP
- XP with 1 or 2 words 1
 - XP with 3 or 4 words 3
 - XP with 5 or 6 words 5
 - XP with 7 words or more 7
 - No co-occurring XP z
3. Weight of co-occurring preverbal XP
- XP with 1 or 2 words 2
 - XP with 3 or 4 words 4
 - XP with 5 or 6 words 6
 - XP with 7 words or more m
 - No co-occurring XP c
4. Argumenthood of postverbal XP
- XP argument r
 - XP non-argument n
 - No XP i

5. Argumenthood of preverbal XP	
— XP argument	s
— XP non-argument	x
— No XP	P
6. Adverbs	
— <i>Frecuentemente</i>	f
— <i>En muchas ocasiones</i>	e
— <i>En más de una ocasión</i>	+
— <i>Diariamente</i>	t
7. Position of Agreeing subject	
— Preverbal subject	v
— Postverbal subject	u
— Omitted subject	o
— <i>Wh-</i> subject	w
8. Argumenthood of Preverbal XP(D/IO)	
— Subject	S
— DO or IO	O
— Predicative	A
— Other XP	X
— No XP	Z
9. Argumenthood of Postverbal XP(D/IO)	
— Subject	E
— DO or IO	D
— Predicative	V
— Other XP	P
— No XP	C
10. Type of verb	
— Transitive	T
— Unaccusative	U
— Intransitive	I
— Reflexive	F
— Copula	K
— Impersonal	—
— Passive	M
— Pronominal passive	R

The argumenthood codes in (18.4) and (18.5) represent the wide definition of argumenthood adopted by Mayoral Hernández (2004), which was based on Hawkins (2000), as indicated in section 3.1. of this paper. However, (18.8) and (18.9) represent the narrower definition of argumenthood that has been tested here.

Cross-tabulations were applied to compare the factors, and the Pearson's Chi-Square test was used to elucidate the relationship between them and, therefore, their statistical significance.

5. Results

5.1. Weight effects

Through the use of the statistical program SPSS, Mayoral Hernández (2004) showed that both the weight of the adverbials and the weight of any co-occurring postverbal XP could determine the collocation of postverbal constituents when differences in weight were in play. He also showed that when there are both an adverbial and an XP in preverbal position, heavier XPs tend to appear attached to the left of the verb in a higher percentage of occurrences. These results supported the previous theories on constituent ordering (Hawkins 1994, 1999, 2000...) and Wasow and Arnold (2003) and provided new information on preverbal ordering: in VO languages the heavier elements tend to be placed in the rightmost position available.

Here we will adopt these results and incorporate them to the current analysis to reach a better understanding of the factors that trigger frequency adverbial alternation in Spanish. As an example, table 1, extracted from Mayoral Hernández (in press), shows how the weight of the adverbial expressions determines their ordering in postverbal positions, with a Pearson's Chi-Square value of $P < 0.05$.

TABLE 1. Ordering of postverbal adverbials, depending on weight²

Position of adverbials	Type of adverb			Total
	En más de una ocasión	En muchas ocasiones	Frecuentemente	
Adjacent to the right of the verb	176 73.3%	50 84.7%	63 94.0%	289 79.0%
After XP in postverbal position	64 26.7%	9 15.3%	4 6.0%	77 21.0%
Total	240 100.0%	59 100.0%	67 100.0%	366 100.0%
Chi-Square tests				
	Value	df	Asymp. Sig. (2-sided)	
Pearson Chi-Square	14.922 ^a	2	.001	
Likelihood Ratio	17.527	2	.000	

When it comes to the position of the adverbials in the sentence, table 2 shows that the default positions are those immediately adjacent to the verb, with a total of 80.8% occurrences.

² The numerical data in every cell represent the total number of tokens and the percentage within the column.

TABLE 2. Influence of weight on adverbials

Position of adverbials	Type of adverb			Total
	En más de una ocasión	En muchas ocasiones	Frecuentemente	
Before XP in preverbal position	26 6.4%	47 15.5%	24 7.4%	97 9.4%
Adjacent to the left of the verb	155 38.4%	127 41.8%	135 41.5%	417 40.4%
Adjacent to the right of the verb	161 39.9%	98 32.2%	159 48.9%	418 40.5%
After XP in postverbal position	62 15.3%	32 10.5%	7 2.2%	101 9.8%
Total	404 100.0%	304 100.0%	325 100.0%	1,033 100.0%

However, although weight could explain postverbal ordering and it proved a determining factor in the collocation of preverbal XPs, its influence on preverbal adverbials was not so clear, which suggested the incorporation to the analysis of other factors. Table 3 shows how weight does not have the desired effect on adverbial ordering, since the two rows are not the mirror image of each other, and there is not a direct increase or decrease of the percentages related to weight differences.

TABLE 3. Influence of weight on preverbal adverbials

Position of adverbials	Type of adverb			Total
	En más de una ocasión	En muchas ocasiones	Frecuentemente	
Before XP in preverbal position	26 72.2%	47 78.3%	24 46.2%	97 65.5%
Adjacent to the left of the verb	10 27.8%	13 21.7%	28 53.8%	51 34.5%
Total	36 100.0%	60 100.0%	52 100.0%	148 100.0%

The incorporation to the analysis of different factors should improve its general validity. In the next section the argumenthood factor will be analyzed, adopting a narrower definition than the one used in Mayoral Hernández (2004).

5.2. Argumenthood

The effects of argumenthood or lexical dependencies on constituent ordering have been clearly shown in the linguistics literature. Hawkins (2000), for example, showed that postverbal arguments tend to appear immediately adjacent to the right of the verb in VO languages like English or Spanish, while they tend to precede the verb in OV languages like Japanese. However, Mayoral Hernández (2004) found no difference between arguments and adjuncts when analyzing the collocation of frequency adverbials and XPs in postverbal position, although he noticed that argumenthood does seem to influence the collocation of preverbal XPs.

These contradictory results might be due to the wide definition of argument adopted in his paper, as we noted before. A more restrictive definition of argumenthood, as indicated in section 3.1, was predicted to yield different results.

In spite of the narrower definition of argumenthood adopted in the present research, the influence of argumenthood on postverbal positions was not statistically significant, with a P value of 0.429 ($P > 0.05$), as shown in table 4.

TABLE 4. Influence of argumenthood on the position of postverbal adverbials

Position of adverbials	Argumenthood of postverbal XP		Total
	Co-occurring postverbal XP is argumental	Co-occurring postverbal XP is non-argumental	
Adjacent to the right of the verb	106 75.2%	135 78.9%	241 77.2%
After XP in postverbal position	35 24.8%	36 21.1%	71 22.8%
Total	141 100.0%	171 100.0%	312 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.625 ^b	1	.429
Likelihood Ratio	.623	1	.430

The results of this research support Mayoral Hernández's (2004) observations. Even when only DOs and IOs were considered arguments the P value is still far greater than 0.05, which implies that the fact that any co-occurring postverbal XP is an argument or an adjunct does not influence the position of postverbal frequency adverbials. However, there is a kind of XP that proved to behave differently from other arguments and adjuncts: predicative complements. Predicative complements are the only type of XP whose appearance influences adverbial ordering, as shown in table 5.

TABLE 5. Influence of argumenthood on the position of postverbal adverbials:
Predicative Complements

Position of adverbials	Argumenthood of postverbal XP		Total
	Predicative complement	Other XPs	
After XP in postverbal position	25 41.0%	76 22.2%	101 25.0%
Adjacent to the right of the verb	36 59.0%	267 77.8%	303 75.0%
Total	61 100.0%	343 100.0%	404 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9.790 ^b	1	.002
Likelihood Ratio	8.977	1	.003

Although these tables show that argumenthood does not have a statistically significant influence on the collocation of postverbal constituents, it is worth noting that Mayoral Hernández (2004) showed that the argument condition of preverbal XPs is statistically significant. Thus, subjects tend to appear immediately adjacent to the left of the verb when only preverbal positions are taken into account³. Table 6 shows that the argument or adjunct condition of preverbal XPs is a statistically significant factor, with a P value of .000. Our data only confirms Mayoral Hernández's (2004) analysis.

At this point we have seen that weight has an important influence on the collocation of postverbal XPs, while the argument/non-argument distinction is highly relevant when determining preverbal ordering. However, we still need a factor that can explain when an adverbial will appear in preverbal or postverbal position. This is why Mayoral Hernández (2004) introduced the *subject position* factor, while the present research will try to explain it through the incorporation of the *type of verb* factor.

³ Out of 1033 sentences analyzed in this research in which an argument and an adverbial co-occurred in preverbal position (a total of 147 sentences), only two contained a non-subject argument. The rest of the sentences, i.e. 145, contained a subject and an adverbial. Since it is very uncommon to have left dislocations in Spanish unless they are separated by commas, it is not possible to provide an analysis of preverbal non-subject arguments in this research.

Table 6. Influence of argumenthood on the position of preverbal adverbials

Position of adverbials	Argumenthood of preverbal XP		Total
	Co-occurring preverbal XP is argumental	Co-occurring preverbal XP is non-argumental	
Before XP in preverbal position	93 69.4%	2 15.4%	95 64.6%
Adjacent to the left of the verb	41 30.6%	11 84.6%	52 35.4%
Total	134 100.0%	13 100.0%	146 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	15.126 ^b	1	.000
Likelihood Ratio	14.814	1	.000

5.3. Subject position

Mayoral Hernández (2004 and *in press*) noted that frequency adverbials tend to appear in complementary distribution with the co-occurring agreeing subject. Therefore, when the subject appears in postverbal position, frequency adverbials tend to appear in preverbal position, and *vice versa*. These facts are illustrated in table 7, where a Pearson's Chi-Square of $P=.000$ is provided.

TABLE 7. Influence of overt subjects on adverbial ordering

Position of adverbials	Subject Position		Total
	Postverbal	Preverbal	
Postverbal position	28 33.3%	239 64.8%	267 58.9%
Preverbal position	56 66.7%	130 35.2%	186 41.1%
Total	84 100.0%	369 100.0%	453 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.941 ^b	1	.000
Likelihood Ratio	27.636	1	.000

However, as seen in section 3, these effects might be due to the class of verb that appears in the sentence. Postverbal subjects can be the result of the appearance of unaccusative verbs, while preverbal subjects should occur when the main verb is transitive, unergative or a copula.

5.4. Type of verb

As predicted in section 3.3 the type of verb does determine the position of the subject and, therefore, adverbial ordering. The analysis shows that intransitives, copulas and transitives are not significantly different with respect to subject position, and in fact Pearson's Chi-Square could not find a statistically significant difference between them, as shown in table 8, with a P value of .678. The subject tends to appear in preverbal position with these types of verbs and there are no significant differences between them in their percentages of appearance in postverbal or preverbal position.

TABLE 8. Position of subject with transitive, intransitive and copulative verbs

Position of subject	Type of verb			Total
	Intransitive	Copulative	Transitive	
Postverbal	6 12.2%	12 15.4%	23 11.5%	41 12.5%
Preverbal	43 87.8%	66 84.6%	177 88.5%	286 87.5%
Total	49 100.0%	78 100.0%	200 100.0%	327 100.0%
Chi-Square tests				
	Value	df	Asymp. Sig. (2-sided)	
Pearson Chi-Square	.777 ^a	2	.678	
Likelihood Ratio	.748	2	.688	

However, the results in table 9 show that, when compared to intransitive, copulative and transitive verbs, unaccusatives tend to have postverbal subjects in a significantly higher percentage. With a Pearson's Chi-Square of $P = .000$. These statistical data show that the unaccusative nature of the verbs have an overt influence on constituent ordering.

But at this point, it is necessary to ask if there is still need of a subject hypothesis or if we can do without it. If subject position is indeed derived from the class of verb that appears in the sentence, then excluding this factor would result in a more economical theory.

TABLE 9. Position of subject with transitive, intransitive, copulative and unaccusative verbs

Position of subject	Type of verb		Total
	Intransitive, transitive and copulative	Unaccusative	
Postverbal	41 12.5%	25 33.8%	66 16.5%
Preverbal	286 87.5%	49 66.2%	335 83.5%
Total	327 100.0%	74 100.0%	401 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	19.809 ^b	1	.000
Likelihood Ratio	17.102	1	.000

5.5. Subject position revisited

In order to find out if we still need the subject position factor, as Mayoral Hernández (2004 and in press) claim, we need to come up with a context in which unaccusativity cannot interfere. If we only take into consideration sentences in which the verb is transitive, intransitive or copulative, after we have seen that they behave very similarly with respect to subject position, we should be able to find out if subject position is still relevant when determining the ordering of frequency adverbials in Spanish.

In table 10, only sentences in which the verb is transitive, intransitive or copulative have been selected. However, we can still see that the influence of subject position is very relevant, since it tends to appear in complementary distribution with frequency adverbials, as Mayoral Hernández suggested before. Pearson's Chi-Square test $P=.003$ supports Mayoral Hernández (2004) theory.

When only sentences with unaccusative verbs are selected, there is still a statistically significant influence of the subject position on adverbial ordering, as shown in table 11, where $P<0.05$. These data oblige us to accept that the presence of an overt subject determines the choice between preverbal and postverbal positions, which makes this factor indispensable.

As suggested in Mayoral Hernández (in press) the complementary distribution in which agreeing subjects and adverbials appear could be linked to EPP effects, in which adverbials can fulfill the EPP when subjects are not occupying preverbal positions or they are omitted. This EPP effect can be linked to the work by Fernández Soriano (1989, 1999a and 1999b), in which datives and preverbal XPs with locative meaning (locative subjects) are claimed to fulfill the EPP. Bear in mind that the re-

sults provided in this paper are not based on theory internal explanations, but on actual data analyzed through statistical software (SPSS). We would like to suggest that a statistical analysis of language in use, specially written texts, can be a tool to test linguistic hypotheses, apart from theory internal explanations.

TABLE 10. Position of adverbials depending on subject position with transitive, intransitive and copulative verbs

Position of adverbial	Position of Subject		Total
	Postverbal	Preverbal	
Postverbal	16 39.0%	180 62.9%	169 59.9%
Preverbal	25 61.0%	106 37.1%	131 40.1%
Total	41 100.0%	286 100.0%	327 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.539 ^b	1	.003
Likelihood Ratio	8.351	1	.004

TABLE 11. Position of adverbials depending on subject position with unaccusative verbs

Position of adverbial	Position of Subject		Total
	Postverbal	Preverbal	
Postverbal	10 40.0%	35 71.4%	45 60.8%
Preverbal	15 60.0%	14 28.6%	29 39.2%
Total	25 100.0%	49 100.0%	74 100.0%
Chi-Square tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.861 ^b	1	.009
Likelihood Ratio	6.818	1	.009

6. Conclusion

A variationist approach that incorporates different factors as triggers of the alternation in position of frequency adverbials has been proven adequate. As discussed in the last section, the validity of a linguistic theory can be tested through the use of statistical analyses or through the use of theory internal explanations. We would like to suggest that the use of written text can reflect the linguistic competence and Saussure's *langue* easier than spoken language (see Newmeyer 2003 for a discussion on this subject).

In this paper we have shown that a restrictive definition of *argumenthood*, in which only direct and indirect objects were considered as internal arguments, cannot provide a better generalization than Hawkins' (2000) definition of dependency. These results support Mayoral Hernández's (2004 and *in press*) analysis, where he adopted Hawkins' *wide* definition of argumenthood. We also proved that the only XP that has the argument properties suggested by Hawkins for English, i.e. the tendency to appear immediately adjacent to the verb, is the predicative complement.

The results obtained about the importance of argumenthood as a factor that determines the collocation of adverbials and other XPs show that it cannot account for postverbal ordering, but it explains preverbal positions. In this respect, subjects tend to appear immediately adjacent to the verb. In postverbal position, frequency adverbials behave like arguments, in the sense that they tend to appear immediately adjacent to the verb.

The *type of verb* hypothesis has been shown to influence the position of the subject, which fulfills the predictions of the unaccusative hypothesis. Unaccusative verbs tend to have postverbal subjects in a higher percentage than transitive, intransitive and copulative verbs. The present study provided for the first time a statistical analysis of one of the purported features of unaccusative verbs: postverbal subjects.

Even if the type of verb could determine, to a certain extent, the position of agreeing subjects, we have also proven that the *subject position* factor is indispensable, since it determines the choice between preverbal and postverbal adverbs, as Mayoral Hernández (2004 and *in press*) showed. Because of this, we have come to the conclusion that this factor cannot be substituted or entirely derived from the kind of verb.

The fact that adverbials tend to appear in complementary distribution with subjects, could suggest that *EPP* can be fulfilled by both agreeing subjects and frequency adverbials, as seen before.

Finally, adopting the results obtained by Mayoral Hernández (2004 and *in press*), we have seen that *weight* determines the collocation of postverbal constituents (Mayoral Hernández 2004 and *in press*). Weight has been claimed to be a typological universal in the literature, and therefore this analysis has typological validity.

Summarizing, the main contributions of this paper have been to provide a detailed description of subject position with different verb classes. As predicted, postverbal subjects tend to appear with unaccusative verbs. We have also shown that a more restricted definition of argumenthood does not provide better results than a

wider definition. We have finally shown that a statistical analysis of linguistic phenomena obtained from a corpus can be an excellent tool when testing linguistic theories, like the unaccusative hypothesis.

References

- Bosque, I. and Demonte, V. (eds.), 2000, *Gramática descriptiva de la lengua española*, Madrid, Espasa.
- Cinque, G., 1999, *Adverbs and functional heads: A crosslinguistic perspective*, New York and Oxford, Oxford U.P.
- Corpus de Referencia del Español Actual* (CREA), in www.rae.es
- Fernández Lagunilla, M. & Anula Rebollo, A., 1995, *Sintaxis y cognición. Introducción al conocimiento, el procesamiento y los déficits sintácticos*. Madrid. Editorial Síntesis.
- Fernández Soriano, O., 1989, *Rección y ligamiento en español: Aspectos del parámetro del sujeto nulo*. Doctoral dissertation, Universidad Autónoma de Madrid.
- , 1999a, “Two types of impersonal sentences in Spanish: Locative and dative subjects”, *Syntax* 2: 2, 101-40.
- , 1999b, “Datives in construction with unaccusative *se*”, *CatWPL* 7.
- Hawkins, J., 1994, *A performance theory of order and constituency*, CUP, Cambridge.
- , 1999, “Processing complexity and filler-gap dependencies across grammars”, *Lg* 75.2. 224-85.
- , 2000, “The Relative order of Prepositional Phrases in English: Going beyond Manner-Place-Time”, *Language Variation and Change* 11, 231-66.
- , 2001, “Why are categories adjacent?”, *JL* 37.1, 1-34.
- , in press, “Efficiency and complexity in grammars: Three general Principles”. In: M. Polinsky & J. Moore (eds.), *Explanation in Linguistics*, CSLI Publications, Stanford.
- Kovacci, O., 1999, “El adverbio”. In Bosque and Demonte (eds.).
- Mayoral Hernández R., 2004, “Importance of weight and argumenthood on the ordering of adverbial expressions”, WCCFL 23 Proceedings. Somerville, MA: Cascadilla Press.
- , in press, “On the position of frequency adverbials in Spanish”, *USC Working Papers in Linguistics*.
- Newmeyer, F. J., 2003, “Grammar is grammar and usage is usage”, *Lg* 79.4.
- Rodríguez Ramalle, T., 2000, *Algunos aspectos de la sintaxis semántica de los adverbios y de ciertas expresiones adverbiales*. Doctoral dissertation, Universidad Autónoma de Madrid.
- Rojo, G., 1999, “Sobre los complementos adverbiales”, in VVAA, *Profesor Francisco Marsá /Jornadas de Filología*, Universidad de Barcelona, 153-71.
- Schutze, C. T. & Gibson, E., 1999, “Argumenthood and English prepositional phrase attachment”, *Journal of Memory and Language* 40, 409-431.
- Silva-Corvalán, C., 1982, “Subject expression and placement in Mexican-American Spanish”, in Amastae, J., and L. Elias-Olivares (eds.). *Spanish in the United States: Sociolinguistic Aspects*, N. Y., Cambridge U. P., 93-120.
- Wasow, T. & Arnold, J., 2003, “Post-verbal constituent ordering in English”, in Rohdenburg, G. & Britta, M. (eds.), *Determinants of Grammatical Variation in English*, Mouton de Gruyter, Berlin.

NOMINAL PARTICIPLES, A CASE OF CATEGORIAL ALTERNANCE: EVENTIVE NOMINALIZATIONS IN *-DA**

Aysa Mondoñedo C.

Instituto de la Lengua Española (CSIC),
Instituto Universitario Ortega y Gasset/Universidad Autónoma de Madrid

Abstract

The purpose of this paper is to present a first approximation to the description of the formation and behavior of a type of deverbal noun in Spanish, formed as of the suffixation of a morpheme -da (feminine '-ed') to a verbal base, which is very productive in Spanish. As it will be shown, the suffixation of this morpheme gives rise to eventive and resultative nominalizations as well as object nouns. Here, I focus on the first case and I assume that this kind of nominalization is a subset of deverbal nouns in -DO ('-ED'), given that both types share morphological, semantic and syntactic properties.

1. Introduction

In the 80's, within the framework of Generative Grammar, it was argued that grammatical categories could be axiomatically defined by means of a restricted number of binary features N and V. This was the case, for instance, of nouns, which were defined as elements with the features [+N, -V], or verbs, whose features were [+N, -V]. In this sense, the participle, which was not one of the four basic grammatical notions (N, V, A and P), was the object of many studies aimed at understanding the nature of grammatical categories and their grammatical behaviour. Examples of these studies are, among others, the work of Jaeggli (1986), where he argued the existence of a passive morpheme with the feature [+N], which absorbed the thematic role to be assigned to the external argument of the passive English sentence, and the work of Lefebvre and Muysken (1988), who, departing from the analyses of nominalizations in Quechua, assumed that participles could be defined in terms of features [+N, +V].

* I wish to express my special gratitude to Soledad Varela for the attentive reading and for her opportune comments to the draft. I also wish to thank most sincerely José Camacho and Liliana Sánchez, Jorge Pérez Silva and Isabel Pérez for contributing with valuable suggestions and for their good-natured disposition.

However, the past participle, generated from a verbal base to which a suffix is added, namely *-do* ('-ed') or its allomorphs *-to*, *-so* and *-cho*, and the corresponding feminine and plural forms (*-DO*) ('-ED'), may appear in several syntactic contexts and presents, in Spanish, different forms and different associated properties. Among these contexts we find the following in 1:

- (1) (i) *haber* (aux. 'to have') + past participle
 - a. El partido de fútbol ha *causado* alboroto en la ciudad
'The football match has caused disturbance in the city'
- (ii) *ser* (aux. 'to be') + past participle
 - b. Los jinetes fueron *vencidos* por los toros salvajes
'The riders were beaten by the wild bulls'
- (iii) *estar* (aux. 'to be') + past participle
 - c. El agua está *contaminada*
'The water is polluted'
- (iv) Verbal participle in absolute clauses
 - d. *Escritas* las declaraciones, partió al aeropuerto
'Having written the statements, he left for the airport'
- (v) Verbal participle in agreed clauses
 - e. *Asustados* como estaban, se rindieron de inmediato
'Frightened as they were, they surrendered immediately'
- (vi) Adjective participle¹
 - f. Ese hombre es muy *decidido*
'That man is very resolute'
- (vii) Nouns
 - g. El *revelado* cuesta 10 euros
'Developing costs 10 euros'

The purpose of this paper is to present a first approximation to the description of the formation and behavior of a type of deverbal noun in Spanish formed as of the suffixation of a morpheme *-da* (feminine '-ed') to a verbal base,² which is very

¹ Varela (2001) and (2002) establishes a difference between verbal and adjectival participles departing from a dual aspectual projection of the *-do* suffix (a progressive and a stative one) and states that genuine *-do* adjectives are active and related to non-agentive verbs.

² For this paper, I will assume that deverbal nouns in *-da* (and also in *-DO*) derive from verbs. I believe that, for this specific case, the discussion about the base of the derivation of these nouns does not lead to a solution that takes into account how the past participle, that can appear in any of the constructions mentioned above in 1, can exhibit the different properties it shows in each case. For example, for the case of nouns in *-DO* (*-da* included), it could be possible to assume that this morpheme suffixes to a past participle, as has been sustained by Beniers (1977) and Bordelois (1993), for instance, to a kind of root, or to a verb (as I assume). If it is assumed that the past participle that constitutes a noun is the base of the derivation, then it would be necessary to suffix (some way) some kind of element that makes possible that this past participle, in a given case, behaves like a noun. If it is assumed that the base of the noun is a root, the suffix *-da* (or, in the general paradigm, the suffix *-DO*) should be, to some extent, responsible for the nominal behavior of the elements we study, that is, nouns in *-da* (and *-DO*), or any other particle that produces this effect. And if we assume that the verb is the base of these constructions, we also have to explain why, in

productive in Spanish. As it will be shown, the suffixation of this morpheme gives rise to eventive and resultative nominalizations as well as object nouns. Here, I will focus on the first case and I will assume that this kind of nominalization is a subset of deverbal nouns in -DO (1g), given that both types share morphological, semantic and syntactic properties.³ The following examples seem to be ascribable to the general paradigm of deverbal nouns in -DO: Nouns referring to persons:

- (2) a. El *herido* muestra señales de hipotensión
 'The [V 'wound' -do] (wounded) shows signs of hypo-tension'
 b. Tuvimos que acercarnos a ver a la *muertita*
 'We had to approach to see the [V 'die' -t (dim. suf.) -a] (little dead woman)'
 c. Llegaron unos *abogados* trayendo la demanda contra la compañía
 'Some [V 'advocate' -do -pl] (lawyers) arrived, with the demand against the company'

Object nouns:

- (3) a. A mí no me gusta el *cocido*
 'I don't like [V 'stew' -do] (stew)'
 b. El niño dejó toda la *comida*
 'The child left the whole [V 'eat' -da] (meal)'
 c. A todos engaña el *parecido* de Carlos con su hermano
 'Carlos' [V 'resemble' -do] (resemblance) to his father misleads everybody'
 d. Luis me regaló unas *entradas* para el cine
 'Luis gave me some [V 'enter' -da -pl] (tickets) for the cinema'
 e. Subraye el *predicado* de la oración
 'Underline the sentence's [V 'predicate' -do] (predicate)'
 f. Yo preparo la *picada* y vos hacés la ensalada
 'I prepare the [V 'niddle' -da] (snacks) and you make the salad'

this case, the case of nouns in -DO, the suffix or some other element allow these formations to act like nouns.

[...]	...]	-DO / -da]	?]
Root	V	p.p.	N

But any of these possibilities in itself allows us to explain how a formal past participle can act like a noun here (1g). Moreover, how (1a) it can also act like a compound verb, or with an auxiliary *ser* ('to be'), that is, in a passive sentence (1b), with an auxiliary *estar* ('to be') (1c), in absolute or agreed clauses (1d, e), or like an adjective (1f). In other words, it seems always necessary to consider the presence of an extra element to describe the different structures where the past participle participates or, else, another process that can explain this multiple behavior of this element.

³ Although the cases of deverbal nouns in -DO have got the form of past participles, I assume them to be nouns because of its syntactic and semantic behavior. I will leave aside the semantic arguments that lead me to consider them nouns, but from the syntactic point of view, these elements present the typical features of the nominal flexion (gender and number), appear inside DPs (specified by determinants and agreeing with them in gender and number also), and admit relative clauses. Furthermore, they distribute in the sentence along with the phrase they project in three typically syntactic argument positions: the specifier of a VP (VP subject), the complement of a verb, and/or the complement of a preposition.

Resultative action nouns:

- (4) a. El *lavado* del coche dura cuando no llueve
'The [V 'wash' -do] (washing) of the car lasts as long as it doesn't rain'
 b. La reciente directora no podía hacerse cargo de todas las *llamadas*
'The new director couldn't look after all the calls [V 'call' -da -pl] (calls)'
 c. La *llegada* del vuelo está anunciada para medianoche
'The [V 'arrive' -da] (arrival) of the flight is announced by midnight'
 d. A esa hora, los *silbidos*⁴ de la calle despiertan a los vecinos
'At that time, [V 'whistle' -do -pl] (whistles) on the street awake the neighbors'

Eventive action nouns in -DO:

- (5) a. Ana está a cargo del *cuidado* de los niños durante las mañanas
'Ana is in charge of [V 'take care' -do] (caring for) the children during the mornings'
 b. La *corrida* del domingo se prolongó hasta las nueve
'The [V 'run' -da] (bullfight) on Sunday lasted till nine'
 c. Los cultivos se perdieron con la *crecida* del río
'Crops were lost due to the river's [V 'grow' -da] (flood)'
 d. En su *venida* de setiembre, el Papa santificó a San Clemente
'In his [V 'come' -da] (visit) in September, the Pope sanctified St. Clement'

Eventive action nouns in -da:

- (6) a. La *leída* del Quijote de Juana fue emocionante
'Juana's [V 'read' -da] (reading) of Don Quixote was moving'
 b. Julia le dio una *calada* al cigarrillo
'Julia had a [V 'pull' -da] (pull) at the cigarette'
 c. La *desaparecida* de Juan nos preocupó a todos al final
'Juan's [V 'disappear' -da] (disappearance) got all of us worried in the end'
 d. Ayer me metí una *perdida* tal que casi no llego
'Yesterday I had such a [V 'loose' -da], I hardly managed to arrive'
 e. La *nadada* lo ha dejado cansado
'The [V 'swim' -da] (swimming session) has left him worn out'
 f. El rector se pegó una *bostezada* en medio de su propio discurso
'The rector had a [V 'yawn' -da] (yawn) in the middle of his own speech'

Sentences presented from (2) to (6) are intended to be examples of the complexity of the set of nouns in -DO. On the one hand, according to the nature of the referent they denote, we find nouns referring to people, nouns referring to objects and nouns referring to actions. More so, within this last group, and depending on the

⁴ *silbido* 'whistling', from *silbar* 'to whistle', as well as *chillido* 'scream', from *chillar* 'to scream', among others, seem to belong to a homogeneous semantic group that denotes noises and that change the thematic vowel -a- (from the first verbal conjugation in -ar in Spanish) to vowel -i- when there is a verbal base. Inside this group, we could also find the noun *alarido* 'outcry'.

type of predicate they constitute, we may distinguish between resultative nouns and eventive ones. On the other hand, focusing on the lexical and syntactic properties of the verbal base from which these nouns derive, it is remarkable that they come from either accusative or unaccusative verbs as well as from unergative ones, as figure (A) below shows:

REFERENT → BASE VERB ↓	Ns referring to people	Ns referring to objects	ACTION NOUNS		
			Resultative	Eventive -DO	Eventive - <i>da</i>
ACCUSATIVE	el/la herido, - <i>da</i> 'the wounded'	el <i>cocido</i> 'the stew'	el <i>lavado</i> 'the washing'	el <i>cuidado</i> (de los niños) 'the care' (of the children)	la <i>leída</i> 'the reading'
		una <i>comida</i> 'a meal'	una <i>llamada</i> 'a call'	una <i>corrida</i> (de toros) 'a bullfight'	una <i>calada</i> 'a pull' (at the cigarette)
UNACCUSATIVE	El/la <i>muerto</i> , - <i>ta</i> 'the dead'	el <i>parecido</i> (de Carlos con su padre) 'resemblance' (to his father)	la <i>llegada</i> 'the arrival'	la <i>crecida</i> (del río) (the river's) 'flood'	la desaparecida (de Juan) the 'disappearance'
		?	una <i>entrada</i> (billete) 'a ticket'	?	Una <i>venida</i> (del Papa) 'a visit (of the Pope)'
UNERGATIVE	el/la <i>abogado</i> , - <i>da</i> 'the lawyer'	el <i>predicado</i> (el elemento que predica) 'the predicate (the element that predicates)'	el <i>silbido</i> 'the whistle'	?	la <i>nadada</i> 'the swimming session'
		una <i>picada</i> (Arg. 'tapa') 'a snack'		?	una <i>bostezada</i> 'a yawn'

FIGURE (A)

2. Eventive nouns in -*da*

The extreme right column of figure A shows the cases of eventive action nouns in -*da* (feminine '-ed'), which I consider to be a subset of deverbal nouns in -DO ('-ED'). Nouns in -*da* differentiates from eventive nominalizations in -DO in that the first group can appear in constructions with light verbs and also has a more eventive reading, as we will see further on. This is the case of *leída* in sentences like (6a, c, and e) rewritten below:

- (6) a. La *leída* del Quijote de Juana fue emocionante
'Juana's [V 'read' -*da*] (reading) of Don Quixote was moving'
c. La *desaparecida* de Juan nos preocupó a todos al final
'Juan's [V disappear' -*da*] (disappearance) got all of us worried in the end'
e. La *nadada* lo ha dejado cansado
'The [V 'swim' -*da*] (swimming session) has left him worn out'

These nominalizations are very frequent in Spanish and, even though they are not socially marked, they are frequently found in informal speeches. However, in Peninsular Spanish, occurrences of these participles are less frequent than in the dialects of Latin America Spanish, and, though we find forms such as *lavada* 'washing', *colada* 'washing', or *pasada* 'the act of passing the iron, a rub, etc.' —sometimes frequently co occurring with the diminutive suffix *-it-* (as in *lavadita*, *pasadita*) also is present in Latin American Spanish—, most of these occurrences are apparently registered in the borders of the lexicon, in nouns such as *una chupada* (from accusative *chupar* 'to suck') or *una corrida* (from unaccusative 'to cum', 'to have an orgasm'), etc.

Nevertheless, morphology in these eventive nouns is regular, i.e. they always have resort to the suffix *-da*, as in *imprimida* 'printed', *limpiada* 'cleaned', or *venida* 'sold', even when the system presents the irregular, the truncated or the Latin form (*impresa*, *limpio*, or *venta* correspondingly) for the rest of uses in the examples in (1) and presents the feminine morpheme *-a*.⁵

Furthermore, these nominalizations are derived from both accusative and unaccusative verbs in all the dialects of Spanish. Nonetheless, in the varieties of Latin American Spanish, these nouns may also be derived from unergative verbs,⁶ i.e. they may be found in the whole set of verbs,⁷ as it is shown below:

- (6) a. La *leída* del Quijote de Juana fue emocionante
'Juana's [V 'read' -*da*] (reading) of Don Quixote was moving'
(accusative verbal base)

⁵ Even though, for the case of *venta*, it must be considered that the Latin feminine past participle was not transferred to Spanish as such.

⁶ For example, in Argentina: *la corrida*, from unergative *correr* 'to run'; in Chile, *una dormida*, from *dormir* 'to sleep'; in Peru: *una nadada*, from *nadar* 'to swim', etc. The fact that this phenomenon is so frequent in Latin American Spanish and almost non existing in the peninsular dialect, requires an explanation that goes beyond the limits of this paper. However, a parallel solution could arise from the analysis of another phenomenon also absent in Peninsular Spanish exposed by Bartra y Suñer (1997). In cases like *El avión voló alto* 'The plain flew high', the authors proposed that verbs like *volar* 'fly', which seem to be unergative in the S-Structure, present a direct object in the D-Structure, meaning they are ergative verbs. Thus, what has been considered a truncated adverb that modifies the verb would actually be an adjective.

In this sense, one could think that a similar situation could be occurring with unergative eventive nouns in *-da*: the apparent unergative verbs that constitute these nouns in *-da* could be transitive or, at least, ergative ones, that is, with an internal argument or a quantifier. In this way, these would be bounded and, so, we would not be dealing with unergative verbal bases, but with ergative ones.

⁷ It will remain unexplored here why these nominalizations are not possible with other type of verbs, such as causative verbs like *hacer* 'to do', or *causar* 'to cause', etc.

- c. La *desaparecida* de Juan nos preocupó a todos al final
'Juan's [V disappear' -*da*] (disappearance) got all of us worried in the end'
(unaccusative verbal base)
- e. La *nadada* lo ha dejado cansado
'The [V 'swim' -*da*] (swimming session) has left him worn out'
(unergative verbal base)

2. Syntactic formation of nouns in -*da*

Forms in -*da* share an important syntactic feature allowing their being grouped together under a well-defined type: syntactic contexts where they regularly occur, both in every Latin American and Peninsular dialects, are periphrasis with a light verb of the type of *dar*, *meter*, *echar*, *pegar*, *hacer* (English 'to give', 'to have', 'to put', 'to make'), and the corresponding forms with -*se* (3p pronoun). In fact, the possibility they have to appear in these contexts differentiates them from the set of eventive nouns in -*DO*. These are the cases in (6b, d, and f):

- (6) b. Julia le dio una *calada* al cigarrillo⁸
'Julia had a [V 'pull' -*da*] (pull) at the cigarette'
- d. Ayer me metí una *perdida* tal que casi no llego
'Yesterday I had such a [V 'loose' -*da*], I hardly managed to arrive'
- f. El rector se pegó una *bostezada* en medio de su propio discurso
'The rector had a [V 'yawn' -*da*] (yawn) in the middle of his own speech'

which can be paraphrased as:

- (6) b'. Julia fumó del cigarrillo (una vez)
'Julia smoked the cigarette (once)'
- d'. Ayer me perdí de tal manera que casi no llego
'Yesterday I got so terribly lost, I hardly managed to arrive'
- f'. El rector bostezó en medio de su propio discurso
'The rector yawned in the middle of his own speech'

These constructions may take the same arguments admitted by the verb from which the noun in -*da* is derived, since they are eventive nouns.⁹ In other words, in (6f), where *bostezada* 'a yawn', derives from the unergative verbal root of *bostezar*, the derived noun can take the agent argument (*el rector* 'the rector'). And, in (6d) where *perdida* derives from *perderse* 'getting lost', the derived nominal will maintain the internal argument (pro, 1p, sg) and will also show a monadic argument structure.

⁸ *dar una calada* 'to have a pull' seems to be so hardly lexicalized that it cannot be paraphrased as *Julia caló el cigarrillo* 'Julia pulled the cigarette'.

⁹ I assume, along with Grimshaw (1990), that eventive structures are the ones that can deploy an argument structure.

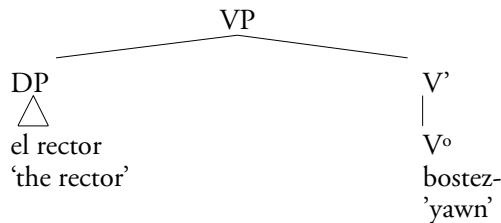


FIGURE 7

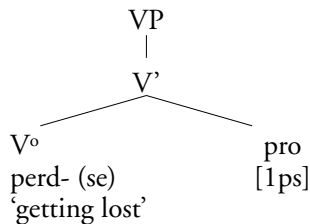


FIGURE 8

However, deverbal nouns in *-da* may also appear in syntactic contexts other than the one described above, i.e. without the presence of a light verb, in a nominal sentence which will take any of the argument positions of the syntactic derivation. This is the case, for instance, of (6a):

- (6) a. La *leída* del Quijote de Juana fue emocionante
 'Juana's reading of Don Quixote was moving'

which may be paraphrased as 'by reading *D. Quixote*, Juana got the audience moved'. Unlike (7), as we may observe in this case, the nominalizations in *deal*, may also deploy a diadic argument structure, where, *Juana* has the agent role and *Don Quixote* has the theme role:¹⁰

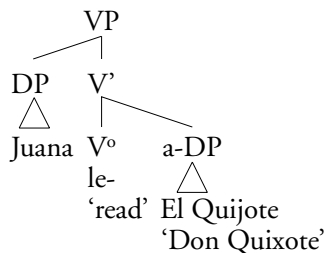


FIGURE 9

¹⁰ In these respect, as it may be noticed in (6f), both theme and agent are introduced by the preposition *de* 'of', at least in the Spanish from Lima, Peru. However, it will be left for further research whether the preposition 'de' introduces the agent in every Latin American dialect or whether this is only characteristic of Andean Spanish, which reduces the possibilities of forming passive sentences and which, furthermore, does not admit the use of *por* 'by', for the case considered here.

Regarding the formation of nouns in *-da*, one possible explanation may arise from considering the possibility that these nouns are formed in the syntax from a verbal root (with an evident [eventive] feature) which, along the derivation, internally merges with the perfective suffix *-da* and check its eventive feature with a nominal category and, then, gives rise to partially deverbalized form with features [+N, +V] (*leída* 'the reading', *calada* 'pull', etc.). This form, at its turn, by merging with the functional features of the determiner (D), becomes determined and part of a DP in the derivation. The syntactic representation of this derivation could be the following:

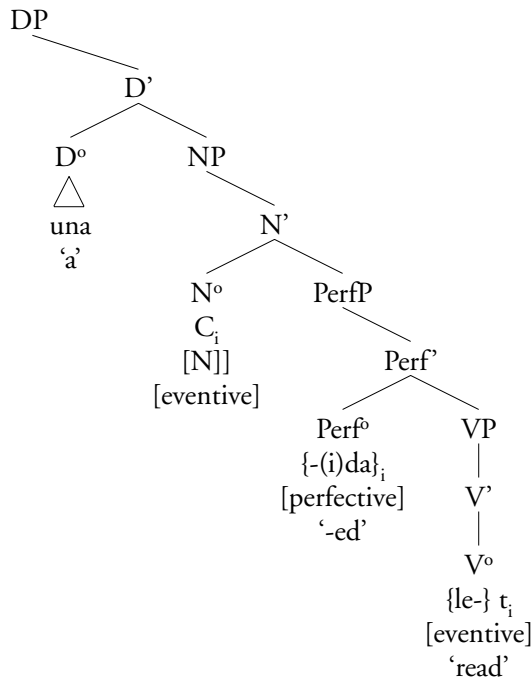


FIGURE 10

The double categorial nature I am assuming for this formation in *-da* may then constitute a sentence with a light verb, which will yield to the noun its functional and inflectional features allowing agreement with the subject, in case there is one. Moreover, it preserves the eventiveness in the meaning of the noun and makes the following construction possible:

- (11) La *bostezada* del rector, en medio de su discurso, duró 10 segundos
 'The rector's yawn, in the middle of his speech, lasted 10 seconds'

According to these criteria, the following representation is a first approximation to the derivation of:

(12) Juan dio una *leída* al examen
 'Juan had a reading at the exam'

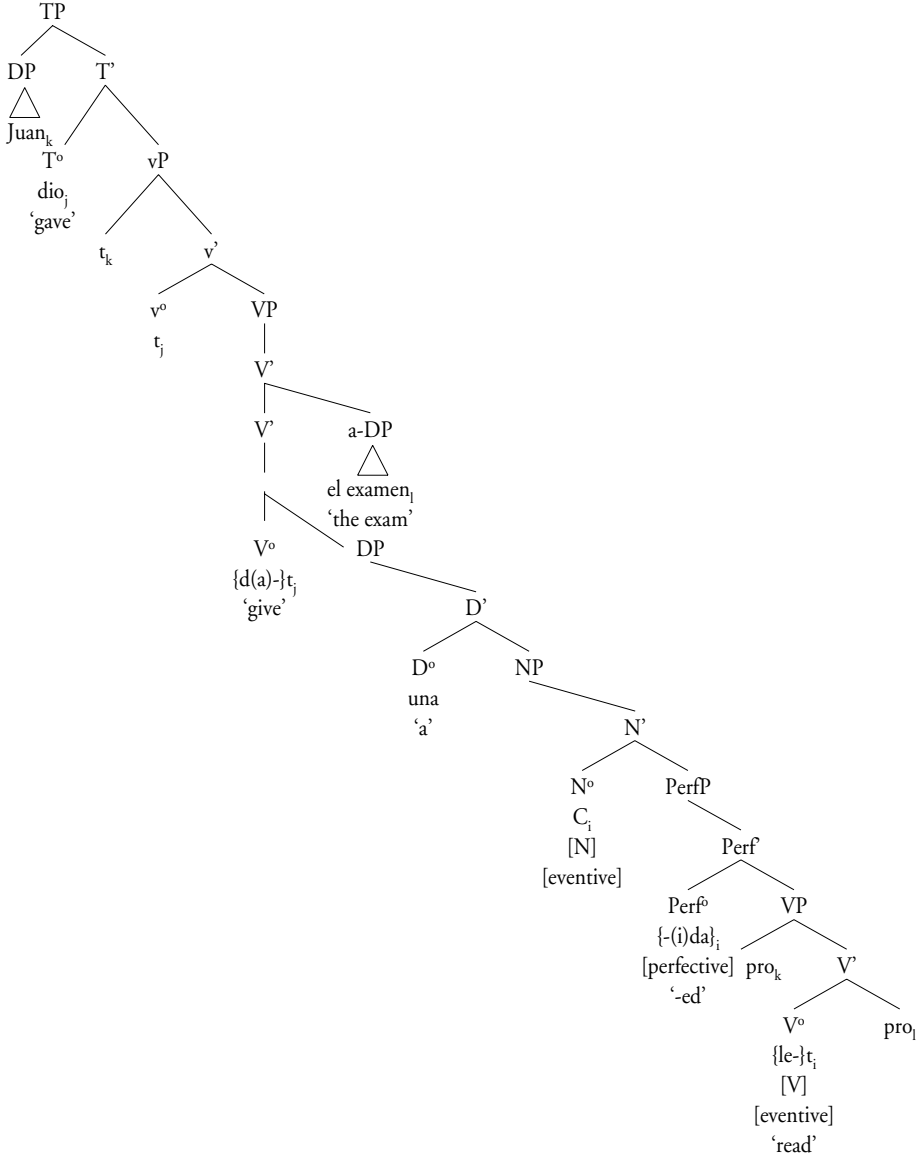


FIGURE 13

In this derivation, the DP is inserted in a TP where *Juan* and *el examen* are generated as arguments of the light verb and are co indexed with two empty categories (*pro*), in the specifier and the complement of the last VP.

The assumption that nominalizations in *-da* are the result of a set of morphosyntactic operations rises from the belief that, in this way, it is possible to capture the fact that past participles can participate in any of the constructions shown in (1), on the one hand, and that speakers seem to be making use of a rather recursive and creative process to generate this type of nominalization, which is possible in many varieties and from almost any verbal base.

3. Semantics of nouns in *-da* and *-DO*

Finally I would like to make some reflections on the semantic relationship between the suffix *-da* and the rest of suffixes referred herein as *-DO*.

The contrast between the great number of eventive nouns in *-da* and the general paradigm of nominalizations in *-DO*, which exhibits a great amount of object nouns, give rise to a reflection on the existence of discrete limits between action nouns—both eventive and resultative—, object nouns, person nouns, etc. In other words, the difficulty in classifying these nouns within the borders of this semantic framework, on the one hand, and the different judgments on whether a noun in *-DO* is an eventive action noun, a resultative or even an object one,¹¹ on the other hand, are both factors which once again call for an argument about the real existence of such limits. Therefore, in my opinion, we should rather approach the classification of these nouns in terms of a semantic *continuum*.

In this sense, it is possible to observe that the boundaries fade in the case of some verbs which, by the suffixation of *-da*, results into eventive action nouns, which, at their turn, may be considered resultative nouns, or even object nouns. Here it turns out to be difficult to find a well-defined semantic parameter allowing the differentiation of one type of noun from the others. This is the case, for example, of:

- (14) *llegada*
 V 'arrive' -*da*
 'arrival'

This noun may refer to:

- (15) (i) an event, as in:
 a. La *llegada* del nuevo profesor (ayer) se produjo de forma inesperada
 'The new professor's arrival (yesterday) was unexpected'
 (ii) a result, as in
 b. La *llegada* del nuevo profesor de ayer fue sorprendente
 'The arrival of the new professor yesterday was surprising'

¹¹ For instance, speakers of Peninsular Spanish of Madrid, classified *lavado* 'washing', as an eventive noun, whilst speakers of American Spanish in Lima, Perú, considered the same noun to be a resultative one. Judgments were provided by speakers trained in metalinguistic thinking, and their remarks also varied depending on the dialect they spoke.

but it can also refer to:

- (iii) an object, or rather, a place, such as the finish line in a course, for instance.
 c. Todos sus amigos esperaron a Juan en la *llegada*
 'All his friends waited for Juan at the *arrival spot*'

Therefore, although nouns like *llegada* may have an ambiguous behavior, the set of nouns in -da that we have been looking at do not seem to have the same possibilities.¹² That is, nouns like *leída*, *chupada*, *desaparecida*, *perdida*, *nadada* or *bostezada* cannot be used referring to object or results. Within a continuum of eventiveness to resultativeness of nouns in -DO, these seem extremely eventive.

The remarks above may also apply to the general paradigm of nominalizations in -DO. That is the case, for instance, of:

- (16) *el encendido*
 [V 'ignite' -do]
 the ignition'

In a sentence like:

- (17) El *encendido* falla todas las mañanas
 'The ignition fails every morning'

all the following interpretations are possible:

- (i) Every morning, the process of switching on the car fails.
 (eventive noun)
 (ii) Every morning, the result of switching on the car fails (i.e. the car does not start on).
 (resultative noun)
 (iii) Every morning, the set of items which make possible the ignition of the car (sparking plugs, battery, carburattor, fuel pump —which are known as a whole as 'the ignition') fails.
 (object noun)

However, if we observe the data introduced in figure A for nouns in -DO, it does not seem possible to attribute the categorial properties of nouns in -da (i.e. [+N, +V]) also to object and person nouns in -DO. Clearly, the set of person and object nouns in -DO, and maybe also the set of resultative nouns, which exhibit a more lexicalized meaning, show nominal properties and excludes verbal ones.

In this way, it seems plausible to assume that the first step in the lexicalization of an eventive noun would be turning it into a resultative one. From this point of view, an eventive noun would stop having features [+N, +V] to have features [+N, -V], becoming then a resultative noun.

Finally, the contrast between the limited set of deverbal eventive nouns in -da in Peninsular Spanish and the great amount of such a formation in Latin American

¹² In the same way, eventive nouns from the contiguous column (where *llegada* is included) cannot occur in the syntactic context previously mentioned.

Spanish leads to consider that, within a linguistic typology regarding these nominalizations, American dialects could be somewhere half the way between a language with [+nominal] nouns (with quite lexicalized meanings and abundance of object, person or resultative nouns) such as Peninsular Spanish,¹³ and a language with abundance of [+verbal] nouns, i.e. a language with abundance of nominalizations, as is the case of Quechua, for instance.¹⁴

References

- Bartra, A. and A. Suñer, 1997, "Inert Agreement Projections and the Syntax of Bare Adjectives", *Probus* 9, 1-31.
- Beniers, E., 1977, "La derivación de sustantivos a partir de participios", *NRFM* 26, 316-31.
- Bordelois, I., 1993, "Afijación y estructura temática: *-da* en español". In *La formación de palabras*, ed. by S. Varela, 162-179. Madrid: Taurus.
- Bosque, I., 1990, *Categorías gramaticales*. Madrid: Síntesis.
- , Demonte, V. (eds.), 1999, *Gramática descriptiva de la lengua española*. Madrid: Espasa.
- Chomsky, N. [1970], 1979, "Observaciones sobre la nominalización". *Sintáctica y semántica en la gramática generativa*. Madrid: Siglo XXI.
- , 1995, *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- , and M. Halle, 1968, *The sound Pattern of English*. New York: Harper and Row.
- Grimshaw, J., 1979, *Argument structure*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, R., 1990, *Semantic structures*. Cambridge, Massachusetts: MIT Press.
- Jaeggli, O., 1986, "Passive", *LI* 17.
- Lefebvre, C. and P. Muysken, 1988, *Mixed categories. Nominalizations in Quechua*. London: Kluwer Academic Press.
- Mendikoetxea, A., 1999, "Construcciones inacusativas y pasivas". In Bosque and Demonte (eds.), 2.1575-1629.
- Picallo, C., 1999, "La estructura del SN: las nominalizaciones y otros sustantivos con complementos argumentos". In Bosque and Demonte, (eds.) 1.363-393.
- Sánchez, L., 1999, "Why does Southern Quechua agree nominally?", *MIT working papers in Linguistics*. Cambridge, Massachusetts: MIT Press.
- Varela, S., 2001, "Lexical morphology revisited: Form Meaning correspondences in Psych Adjectival Participles". *Topics in Morphology: Selected papers from the third Mediterranean Morphology Meeting*. ed. by G. Booij et alii, 51-74. Barcelona: U. Pompeu Fabra.
- , 2002, *Gramática y formación de palabras. Aspectos de morfología derivada del español*, ed. by J. García Medall, Lugo.

¹³ For example, *corrida de toros* 'bullfight', *calada/chupada al cigarro* 'pull at the cigarette', *pedida de mano* 'engagement', *calada*, in the sense of getting wet when it rains, etc.

¹⁴ According to Lefebvre and Muysken (1988), nominalized verbs in Quechua can be the head of a noun phrase or a clause and have obligatory subject. See also Sánchez (1999) for an analysis of [+verbal] Ns in Quechua.

- Williams, E., 1981, "Argument structure and morphology", *The Linguistic Review* 1.81-114.
- , 1989, "The anaphoric nature of θ -roles", *LI* 20, 425-56.
- , 1995, *Theta theory. Government and binding theory and the minimalist Program*, ed. by G. Webelhuth. Cambridge: Blackwell.
- Zagona, K., 2002, *The Syntax of Spanish*, ed. by J. Herschensohn, E. Mallen and K. Zagona. Cambridge: Cambridge U.P.

ANTECEDENT-GAP RELATIONS AND LOCALITY IN VERBAL ELLIPSIS

Elixabete Murguía

University of Deusto

This paper investigates the different locality restrictions that apply to some verbal ellipsis constructions in English; namely, Verb Phrase Ellipsis (VPE), Pseudogapping and Gapping. It is proposed that locality restrictions can be given a natural answer from the processing domain. Locality is analyzed as the result of the interaction of different factors: (i) tense presence/absence (Fodor 1985), (ii) low initial attachment of coordinates, and (iii) Spell-Out operations which render syntactic structure unavailable (Uriagereka 1999).

1. Introduction

It has been observed in the literature (i.e. Chao 1987 and Neijt 1980 on the competence side; and Berwick and Weinberg 1985, together with Fodor 1985 on the processing side) that not all elliptical constructions are subject to the same locality restrictions. The relation between the antecedent and the gap in Verb Phrase Ellipsis (henceforth VPE) and Pseudogapping can be either local (see examples 1a and 2a) or non-local (see examples 1b and 2b), while in the case of gapping, locality has to be respected (see example 3a versus 3b). If locality is not respected, then the sentence turns out to be ungrammatical:

- (1) a. Mary accepted the job offer, and Peter did too.
b. Mary accepted the job offer, and I believe Peter did too.
- (2) a. Tom talked to his wife, and Beth to her husband.
b. Tom talked to his wife, and I heard Beth did to her husband.
- (3) a. Susan prepared lunch, and John dinner.
b. *Susan prepared lunch, and I think John dinner.

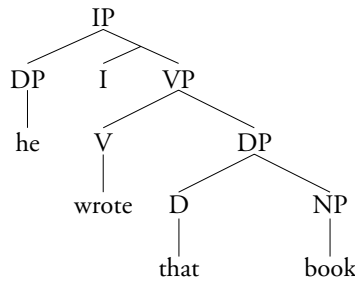
The analysis of locality effects advanced in this work is based on the minimalist framework (Chomsky 1993, 1995); in particular, on the economy principle that governs minimalism. Locality in coordinate elliptical structures is determined by Spell-Out operations, in the sense of Uriagereka (1999) —we will see that an antecedent remains in the local context of the gap if the former has not been spelled out. We assume Weinberg's (1999) human sentence processing algorithm (defined below in 6) and extend it to coordination and ellipsis. Before getting into the analy-

sis of the ellipsis facts, a brief comment on some of the theoretical assumptions taken should be included first. This is done in the next two subsections.

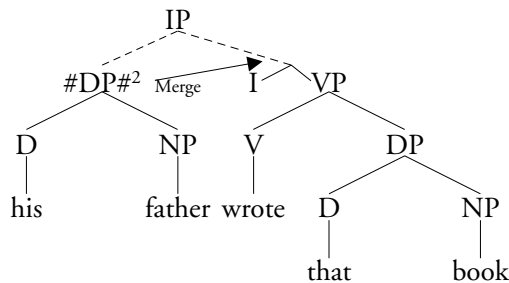
1.1. Multiple Spell-Out Theory (MSO) (Uriagereka 1999)

Uriagereka's (1999) Multiple Spell-Out (MSO) theory is an attempt to reduce Kayne's Linear Correspondance Axiom (LCA) to a more minimalist basis.¹ According to Uriagereka, once the D- and S-structure level have been abandoned, there is no reason to restrict Spell-Out to one unique application. He presents a dynamically split model, in which multiple application of Spell-Out applies, accessing PF and LF in separate derivational cascades. Following Epstein (1999), he proposes that command is a reflex of merge, and also that this command relation codes precedence relations: command maps to precedence in simple Command Units (CUs), because it is the simplest state of affairs. Consider example (4) and (5):

(4) He wrote that book.



(5) His father wrote that book.



Sentence (4) has been assembled through the monotonic application of the operation Merge—the word *that* is merged with *book*, the resulting object is merged with *wrote* and the object of this last operation is in turn merged with *he*—thus, constituting a so-called Command Unit. In the case of (5), however, we have a sen-

¹ Linear Correspondance Axiom (LCA):

Base step: If @ commands &, then @ precedes &.

Induction step: If \$ precedes & and \$ dominates @, then @ precedes &.

² The symbols # # mean that a category has been spelled-out.

tence which has been assembled through the non-monotonic application of merge to two separately assembled objects: (i) object A: *his* is merged with *father* constituting a Command Unit, and (ii) object B: *that* is merged with *book*, and then with *wrote*, constituting another Command Unit. Then, object A (the DP) —once it is spelled out— is merged with object B creating a single CU —which Uriagereka calls the mother CU.

Spell-Out applies to every CU in a derivation, linearizing the elements that compose them—or in other words, establishing the command-precedence relation among terminals. So, in example (4) the CU which is assembled monotonically through subsequent merge operations is spelled out and precedence relations established on the basis of command relations. Thus, the base step of the LCA is accounted for.

In the case of (5), however, the situation is somehow more complicated. The CU constituted by the DP *his father* is spelled out and the precedence relation between these two elements is established. After Spell-Out, what remains is not a phrase marker any longer. The resulting element (the DP) is frozen, it is a lexical compound, so the syntax cannot operate with it any longer —its syntactic structure cannot be altered. However, it can associate further up: the DP can be merged as a unit with the mother CU *wrote that book*. The command-precedence relation between the DP and the elements in this other CU is established in the following way: (i) the node DP commands the elements that constitute the mother CU, since the DP has been merged to those, and (ii) the elements that the label DP dominates should act as the label DP does within its mother CU —this is a consequence of the fact that they have been spelled out separately from that CU the DP has been attached to (they have been spelled out in a different derivational cascade); their place in the structure is frozen, so they cannot interact with the rest of the elements in the mother CU. Thus, the induction step of the LCA is also deduced.

1.2. A Minimalist Theory of Human Sentence Processing (Weinberg 1999)

Weinberg (1999) assumes the minimalist program (Chomsky 1993, 1995), and applies minimalist operations—Merge, Move and Spell-Out (as defined by Uriagereka 1999 above)—together with minimalist principles —economy principles— to parsing. She defines a minimalist algorithm for human sentence processing which not only accounts for some attachment preferences observed in the literature, but also offers a theory of reanalysis. Her algorithm definition is included below:

- (6) A derivation proceeds left to right. At each point in the derivation, merge using the fewest operations needed to check a feature on the category about to be attached. If merger is not possible, try to insert a trace bound to some element within the current command path. If neither merger nor movement is licensed, spell out³ the command path. Repeat until all terminals are incorporated into the derivation.

³ Spell-Out operations are carried out in a phrase-by-phrase manner. In other words, when a command unit is spelled-out, the whole command unit is not linearized at once, but rather each of the phrasal nodes are spelled-out one by one.

Weinberg (1999) assumes a MSO theory for performance in order to account for the mapping between precedence and dominance relations without the need for the LCA in parsing as well.⁴ The base step of the LCA (in footnote 4) is deduced from the fact that it is the simplest mapping relation between precedence and dominance (a one-to-one mapping relation). The induction step is not necessary if MSO applies. Spell-Out applies whenever two categories cannot be merged together (see algorithm definition above): if neither Merge nor Move can apply, then the category being built is spelled out; linearized —or in other words, turned into an unstructured string. For this spelled out string the only important precedence relations are those already established. Precedence does not need to be established between the elements in this string and the rest of the items in the structure.

The algorithm defined above accounts for certain parsing preferences, e.g. Argument-over-Adjunct attachment (Pritchett 1992 and Gibson 1991) and Minimal Attachment Principle (Frazier and Rayner 1982) —see Weinberg (1999) for discussion. It also offers a theory of reanalysis: we are going to discuss this last point in some detail, since it is crucial for the analysis of verbal ellipsis and locality that follows. In order to do so, we look at two examples: in the first one reanalysis is possible, but in the second one reanalysis is blocked by the prior application of Spell-Out.

Reanalysis to a different reading remains possible within a domain where Spell-Out has not applied. Consider sentence (7) below. A verb like *believe* subcategorizes both for a DP and an IP (see examples in 8):

- (7) The man believed his sister to be a genius.
 (8) a. He believed [_{DP} his sister].
 b. He believed [_{IP} [_{DP} his sister] to be clever].

At the point where the determiner *his* is encountered in sentence (7), the parser has two possibilities for attachment: (i) attach the DP as the object of the verb *believe* (as in 8a), or (ii) attach it as the subject of the embedded IP (as in 8b). The parser goes for the first option, since not only is it the most economical one (fewest nodes), but also it allows feature checking (case and theta-role) for the DP. Attachment as the subject of the embedded clause does not allow any feature checking at this point, since the head of the IP has not been processed yet.

When the embedded verb *to be* is processed, the parser needs to reanalyze the syntactic structure it assigned to sentence (7) so as to accommodate the new input items (it needs to reanalyze the attachment of the DP from object into subject of the embedded clause). Thus, however, it does not present any problem, because

⁴ The LCA as proposed by Kayne (1994) derives linear precedence from dominance relations. Weinberg (1999) inverts the claim so as to make it relevant for parsing purposes:

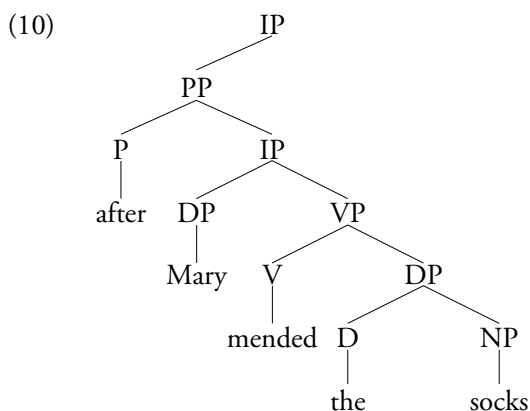
- (i) Linear Correspondance Axiom (LCA)
 Base step: If α precedes β then α dominates β
 Induction step: If γ precedes β , and γ dominates α , then α precedes β .

both the verb and the DP are available: Spell-Out has not applied to these categories yet.

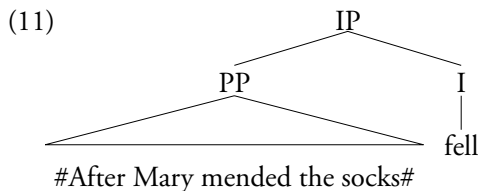
On the contrary, if Spell-Out applies, then extraction or insertion of syntactic material is not possible: once a syntactic structure is spelled out it is frozen, and it cannot be affected by operations such as merge or move. Consider sentence (9) below:

(9) *After Mary mended the socks fell off the table.

When we start processing this sentence a PP is built headed by the preposition *after*. Once we arrive to the DP *the socks*, there are two possible attachments: (i) attachment of the DP as the object of the verb *mend*, or (ii) attachment as the subject of the matrix clause. The first option is the one that the parser chooses (as in 10 below), since it allows features of the DP to be checked by the verb *mend*. Attachment as the subject of the matrix clause does not allow the checking of any feature at this point, since the head of the matrix clause *fell* is not part of the structure yet.



At this point of the parse, the next item to be attached is the verb *fell*. However, it cannot be merged with anything in the preceding clause. Thus, Spell-Out applies: the adverbial clause is linearized (c-command is established for the elements of this string) and the verb *fell* is attached to the structure. The resulting structure is that shown below:



At this point, reanalysis to the non-preferred reading (where the DP is the subject of the matrix clause) is not possible, because the domain where the DP is at-

tached has been spelled out, and consequently, it is not possible to retrieve it in order to attach it as the matrix subject.

2. Verbal Ellipsis and Locality Restrictions

In this section, a minimalist processing account of verbal ellipsis and locality restrictions is introduced. The starting point is Weinberg's (1999) human sentence processing algorithm (defined in 6 above), which is extended to coordination and ellipsis here.

An analysis for parsing different verbal ellipsis constructions is advanced, based on an algorithm defined on the minimalist operations (Merge, Move and Spell-Out), which takes into account economy considerations and which makes use of local information.

Locality is explained as a result of the interaction of different factors: (i) Tense presence/absence, (ii) low initial attachment of coordinates, and (iii) Spell-Out operations which render syntactic structure unavailable. These last two together determine when left-context, i.e. the antecedent in these ellipsis contexts, is available.

The problem for parsing is (i) to detect the gap, and (ii) to resolve/interpret it. As a preview of what is coming, it should be mentioned that there is a contrast between VPE and Pseudogapping constructions on the one hand, and Gapping constructions on the other. For the first two, the gap is detected by the presence of an auxiliary—the auxiliary signals the gap and allows us to predict a VP (the antecedent only needs to be accessed to interpret the predicted VP)—while in the case of gapping the antecedent needs to be consulted to assign structure to the gap and for interpretation purposes.⁵

2.1. Tense Presence/Absence and Locality

As it has already been noticed, there exists a crucial difference between VPE and Pseudogapping elliptical constructions on the one hand, and gapping on the other. In the case of VPE and Pseudogapping, there is an auxiliary overtly realized in the elision site (see 12 and 13). On the contrary, in gapping sentences there is no auxiliary present (see 14):

- (12) Mary is very hungry, and I am too.
- (13) Peter gave his corrections to Susan, and John did to Bill.
- (14) These students ate bagels, and the visitors pizza.

This auxiliary difference is crucial for detecting and resolving the gap. In the case of VPE and Pseudogapping, since there is an auxiliary, an IP can be built and a VP

⁵ In this paper, I just deal with the work that is done on-line by the parser. For a further discussion of how the interpretation process is carried out in ellipsis contexts I refer the reader to Murguía (2004), where both on-line and off-line processes are discussed.

predicted (functional categories like “I” select lexical categories like “V”): the auxiliary is recognized on the basis of the input string (bottom-up), an IP is built, and a top-down prediction of a VP can be made. All this is done by using local context, i.e. the information provided by the auxiliary. There is no need to access the antecedent to detect the gap and assign structure to it. However, in the case of gapping, there is no overt auxiliary or verb from which to build an IP, and the antecedent needs to be accessed in order to detect the possibility of a gap. The antecedent is needed to postulate a node for the gap. This Tense effect was already noticed by Fodor (1985), and discussed by Berwick and Weinberg (1985).

In VPE and Pseudogapping sentences, the VP that is predicted is assigned a pointer to the antecedent VP and it shares the structure with the latter.⁶ The antecedent structure is accessed only for interpretation. It is not accessed on-line to build the structure of the gap. For gapping, the antecedent is accessed on-line to assign structure to the gap.

The difference proposed here is supported by some findings which have been reported in the psycholinguistics literature. Frazier and Clifton (2001) report what they call “missing complexity effects” in VPE sentences. In a self-paced reading experiment, they did not find any difference in the reading times of those sentences below, even though the structure of the antecedent in (15) is more complex than in (16):

(15) Sarah left her boyfriend last May. Tina did too.

(16) Sarah got the courage to leave her boyfriend last May. Tina did too.

This contrasts with complexity effects found by Carlson (2002) for gapping sentences. This difference between VPE and gapping sentences supports the distinction that we have proposed above for VPE and gapping. In the case of VPE, it seems that the antecedent is not accessed on-line for gap detection; otherwise, if the structure of the antecedent is computed for the gap, then there should be differences in the reading times of the sentences in (15) and (16).

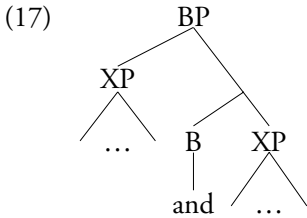
2.2. Low Initial Attachment of Coordinates

Weinberg (1999) evaluates ambiguity of attachment with respect to economy: the most economical structure is preferred, i.e. that one that involves fewest nodes or operations. We translate this economy preference into initial low attachment for coordinates.

As we will see in this section, there is ambiguity of attachment in the case of coordinates too. Coordinators are initially attached low, and this decision is revised into high attachment if later incoming material forces reanalysis—we will see how reanalysis is carried out in detail when I discuss some examples in the next section.

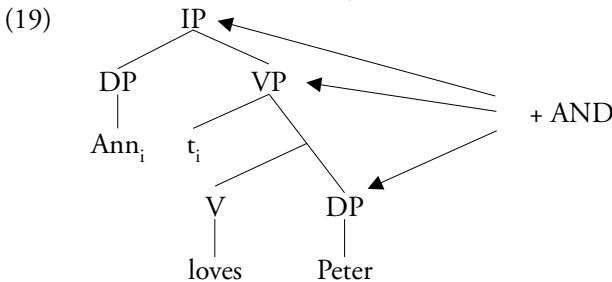
⁶ For a detail discussion of this pointer assignment and of the sharing relation that exists between the antecedent and the gap see Murguía (2004).

We assume that coordinators head Boolean Phrases (BPs) as proposed by Munn (1987) —coordinate sentences have the structure in (17) below. As proposed in Murguía (2000), coordinates are spelled out in different CUs in order to preserve the precedence-command relationship among terminal elements:



Now, let us consider why low attachment is more economical than high attachment. For a sentence like (18), at the point where the coordinator is encountered the structure computed so far is that one in (19), where the structure of what is going to be the first conjunct is already built. At this point, the next input item to be attached is the coordinator *and*. How is the coordinator attached? There are three possible attachment sites, marked with arrows—the three possible attachment sites are the (i) IP, (ii) VP, and (iii) DP nodes:

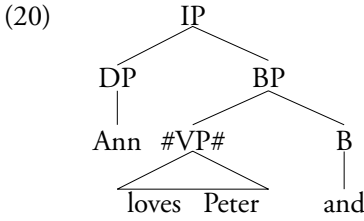
(18) Ann loves Peter, and Mary does too.



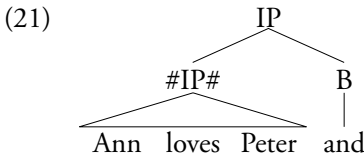
How does the parser choose among these possibilities? Recall that a minimalist grammar is assumed here, and the most important principle in minimalism is the principle of economy —derivations must be as economical as possible (fewest number of steps/operations and fewest number of nodes). This economy principle is what is going to guide the parser in choosing among the three alternatives. Let us consider these in turn.⁷

⁷ Using the strategy of actually considering all the possible alternatives will be a problem for efficiency in parsing—if all the different attachment sites were considered on-line, then the algorithm will be n^2 proportional to attachment sites, which does not respect efficiency. Therefore, economy is enforced in a serial way. Low attachment is the initial choice the parser takes, because it is more economical. The goal of the discussion that follows is to illustrate why it is the case that low attachment of coordinates is more economical. This then justifies the assumption that it becomes the automatic first option without the need for global comparison.

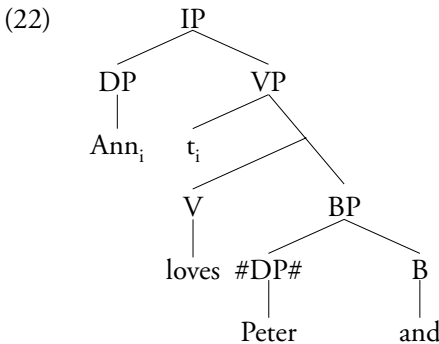
Start by considering alternative (ii): attachment to the VP. In order to attach the coordinator to the VP node and preserve the command-precedence relationship, the intervening material must be spelled out so that this position becomes available. Since Spell-Out occurs in a phrase-by-phrase manner, the DP will be spelled out first, followed by the VP. After these two Spell-Out operations, the coordinator can be merged to the structure. This is what the structure will look like after all these operations:



Alternative (i), attachment to the IP node, will include the same steps as attachment to the VP plus one more Spell-Out operation —Spell-Out of the IP phrase. The structure after attaching the coordinator will be that in (21):



Turning now to the third possibility: attachment to the DP. If the coordinator is attached to the DP then only one Spell-Out operation is necessary: spelling out of the DP as in (22) below. Thus, this third possibility is the most economical one (it involves fewest steps/operations), and the one that the parser chooses—decisions are taken locally, this algorithm is not a global one, and at this point in the derivation attachment to the DP is the best option in terms of economy:



Thus, according to economy, it looks like low attachment should be preferred for coordinates too. This tendency for attaching low has already been observed in the

parsing literature in other contexts different from coordination: Minimal Attachment (Frazier and Rayner 1982), which has been explained in terms of a minimalist parsing algorithm that favors feature checking, and respects economy (Weinberg 1999).

2.3. Locality Effects and Verbal Ellipsis

As we have mentioned, it is a well-known fact in the literature that elliptical constructions are subject to different locality restrictions (e.g. Chao 1987, Fodor 1985, and Berwick & Weinberg 1985). VPE and Pseudogapping are not constrained by any locality restrictions, while gapping is. See examples below, which exemplify this contrast:

- (23) a. Ann loves Peter, and Mary does too.
 b. Ann loves Peter, and Susan thinks Mary does too.
- (24) a. I gave money to Susan, and Peter did to Beth.
 b. I gave money to Susan, and you heard that Peter did to Beth.
- (25) a. John saw Carmen, and Tom Othello.
 b. *John saw Carmen, and Bill thinks Tom Othello.

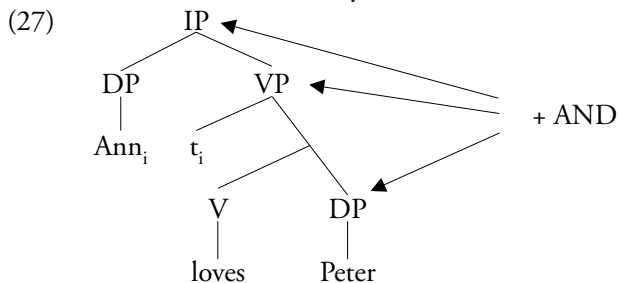
Both for sentences (23) and (24) (examples of VPE and pseudogapping, respectively) the antecedent verb phrase, and the elided constituent can be separated by intervening material—the gap can be embedded (as in the “b” examples), and still result in a grammatical sentence. However, in the case of gapping (sentence 25), if the antecedent and gap are not local (if the elided constituent is embedded as in 25) then the sentence is ungrammatical.

In this section, we propose an analysis for the presence/absence of locality effects in ellipsis which is based on (i) the presence/absence of the auxiliary, and (ii) the availability of left context (i.e. of the antecedent), which in turn is a result of low initial attachment of coordinates and of Spell-Out operations that render syntactic structure unavailable.

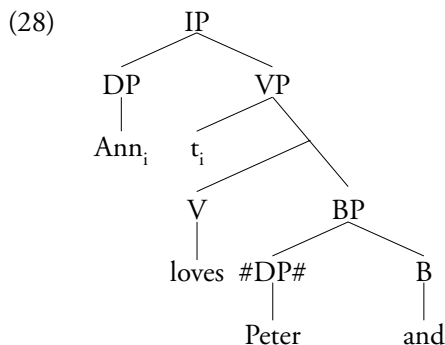
2.3.1. VPE

Let us start with the VPE example (26). Through subsequent Merge and Move operations, the first conjunct structure is built. The next input item to be attached is the coordinator *and* and the three possible attachment sites are those in (27):

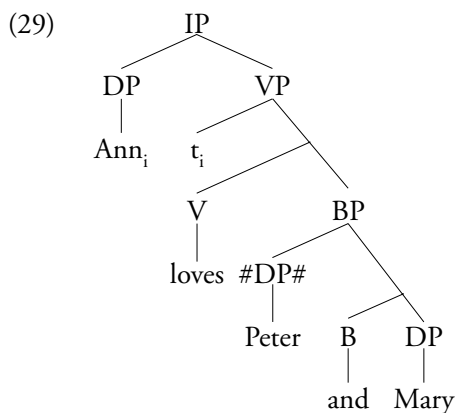
- (26) Ann loves Peter, and Mary does too.



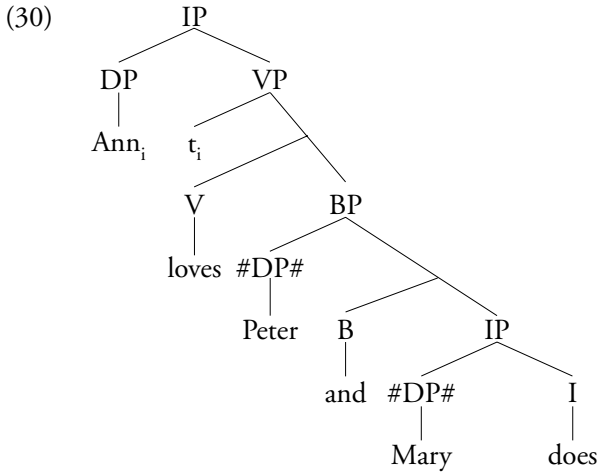
As we saw in the previous section, the parser chooses low attachment (i.e. attachment to the DP), since this is the most economical option; the one that involves fewest steps.



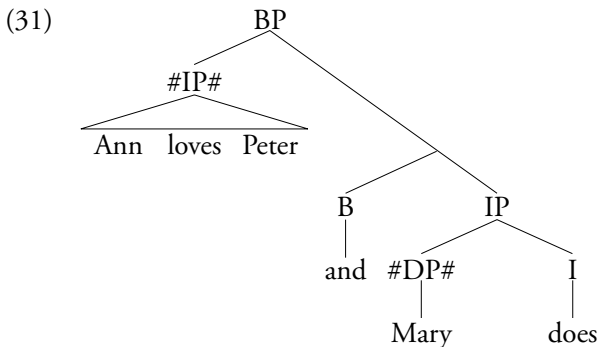
Once the coordinator has been merged to the structure, the next input item to be attached is the DP *Mary*, which will be attached as follows:



There is a condition on coordination that must be respected: the coordinator must conjoin two identical categories. (Coordination of Likes: Williams 1981). If two different categories are coordinated, then this constraint is violated and the sentence is ungrammatical. In (29) above, two DPs are coordinated, so the condition on coordination of likes is respected. The next input item to be attached is the auxiliary *does*. As before, the most economical option is to attach low for the same reason: it involves fewest steps.

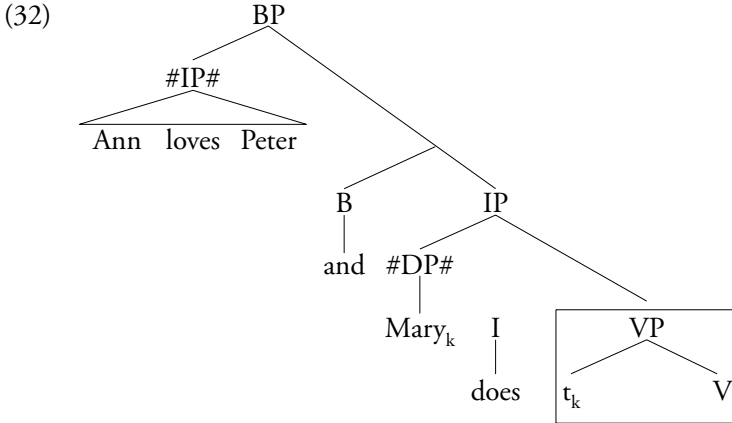


However, one more node (IP) had to be postulated (the DP *Mary* that was coordinated with the DP *Peter* has been transformed into an IP). Now *and* is coordinating a DP *Peter* and an IP *Mary does*, which violates the condition on coordination. At this point reanalysis is necessary. Low attachment is reanalyzed as high attachment, in other words, coordination of objects is reanalyzed as coordination of IPs. To respect the condition on coordination there is only one possibility now, and that is attachment to the IP, as in (31) below:



To attach high as in (31) above, the whole antecedent IP needs to be spelled out so as to preserve the precedence-command relationship among terminals. Spelling out the antecedent makes its internal syntactic structure unavailable. However, at this point where an IP has been built bottom-up for the second conjunct, we can predict a VP; since functional categories select lexical categories (a top-down prediction can be made). We do not need to look back to the antecedent to do this, so whether its syntactic structure is available or not is irrelevant for the parser to successfully detect the gap and assign a category to it. We can also relate the subject in

the specifier of IP to its base position —the specifier of VP— and build all the structure in (32) below:

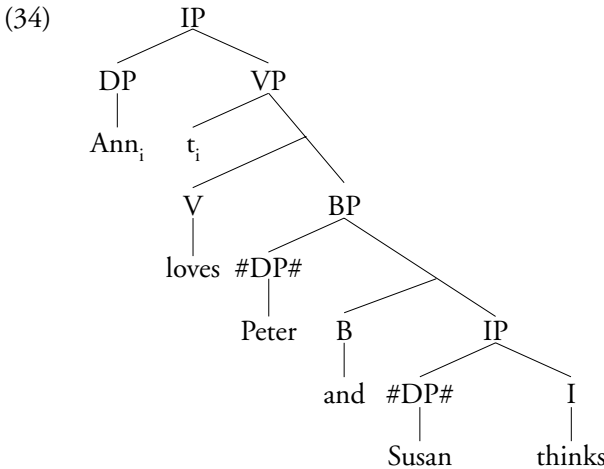


The complete internal structure of this VP, however, as well as the lexical content of the V-category and its complement are not fully specified. The rest of the VP structure and the lexical content are recovered from the antecedent, by following the pointer that the elided VP is assigned. So the antecedent is retrieved for interpretation purposes at LF, but not on-line when the gap is encountered.

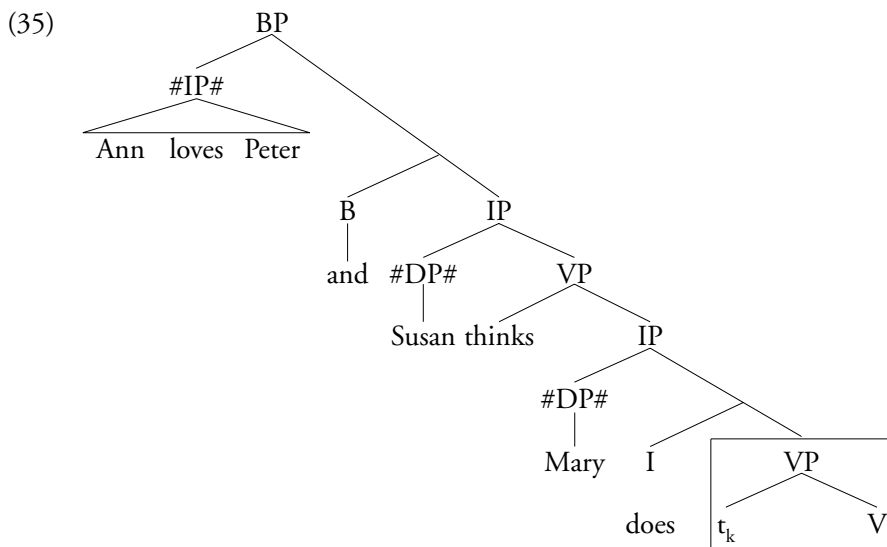
Consider now the second example of VPE mentioned above —where the antecedent and the elided VP are separated by intervening material— repeated here for the reader’s convenience:

(33) Ann loves Peter, and Susan thinks Mary does too.

This sentence will be parsed in the same way as the previous example —the coordinator will be attached low, since this is the most economical option. Reanalysis from low into high attachment here, however, will be triggered by the attachment of the intervening clause *Susan thinks* in (34) below:



At this point, the condition on coordination is not satisfied (a DP and an IP are coordinated) and reanalysis is necessary. The coordinator is attached to the higher IP (as in (35)). In order to attach high, the antecedent clause needs to be spelled out (recall that after spelling out a category, its internal syntactic structure is no longer available); consequently, the antecedent VP will not be accessible —the antecedent cannot be retrieved to assign structure to the gap. Nevertheless, this is not a problem for VPE examples since the elided VP can always be predicted from the IP (built top-down based on the auxiliary *does*), without resorting to the antecedent's help:



2.3.2. Pseudogapping

Consider now the pseudogapping (or subdeletion) examples mentioned above, which are repeated below:

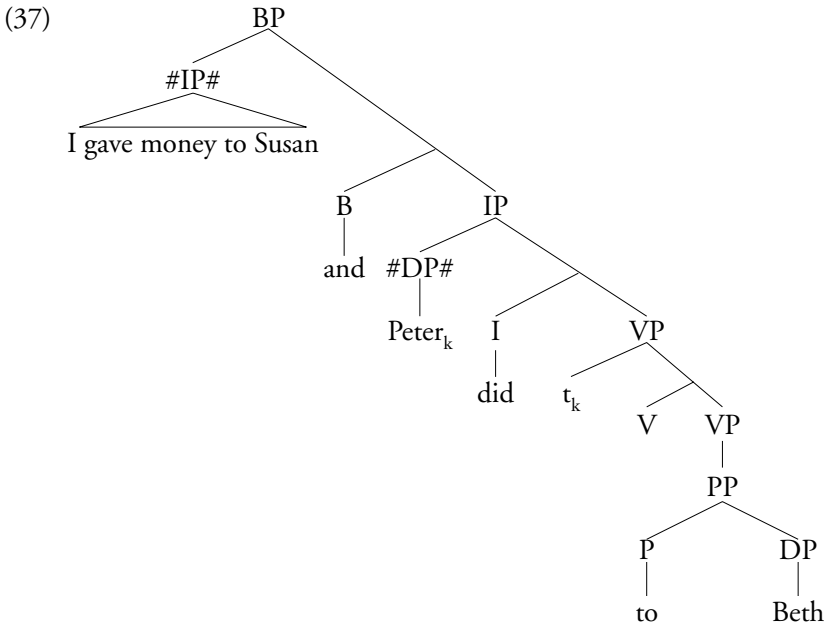
- (36) a. I gave money to Susan, and Peter did to Beth.
 b. I gave money to Susan, and you heard that Peter did to Beth.

Pseudogapping sentences, like VPE, are grammatical whether the antecedent and the elided clause are local or not: in (36) for example there is intervening material between both clauses, but the sentence is still grammatical. An auxiliary is always present, as in the case of VPE, so an IP is built and a VP can be predicted without the need to access the antecedent.

One difference between VPE and pseudogapping examples is that one of the verb arguments/adjuncts in the elided conjunct is overtly realized only in the latter. In example (36) above, the indirect object *to Beth* has not been elided. How is this overtly realized argument attached, when the verb phrase is elided? Since a VP is

predicted top-down (as in the VPE cases) for all pseudogapping cases the argument is attached as part of that predicted VP.

Consider the parse for example (36). The argument will run in the same way as for the VPE examples above—attachment of the coordinator starts low and this is reanalyzed to IP attachment when the auxiliary *did* (in 36) and the clause *you heard* (in) are attached. Once the second conjunct is reanalyzed as an IP, a VP is predicted (a top-down prediction). Finally, the non-elided argument is merged to the structure:



The sentence in (36) is parsed in the same way, with the difference that reanalysis in this case is triggered by the intervening material (as in 33 above). But since an auxiliary is present in the elided conjunct an IP is built, and a VP predicted—to which the overtly realized argument is attached. Because of this possibility to predict a VP, the non-availability of the antecedent (it has been spelled out, so it is not available) does not pose a problem neither for the resolution of the gap nor for the attachment of the argument *to Beth*—both can be done without the need to resort to the antecedent.

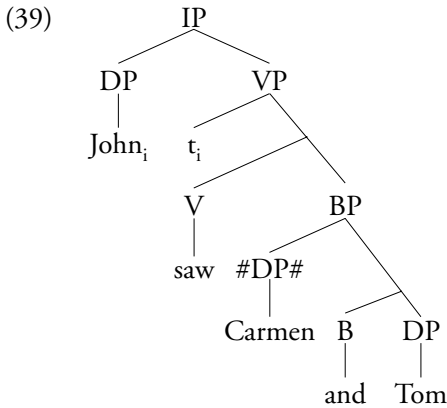
2.3.3. *Gapping*

The gapping examples differ from VPE and pseudogapping by showing locality effects. When the antecedent and the elided clause are separated by intervening material (as in 38 below), then the sentence is ungrammatical:

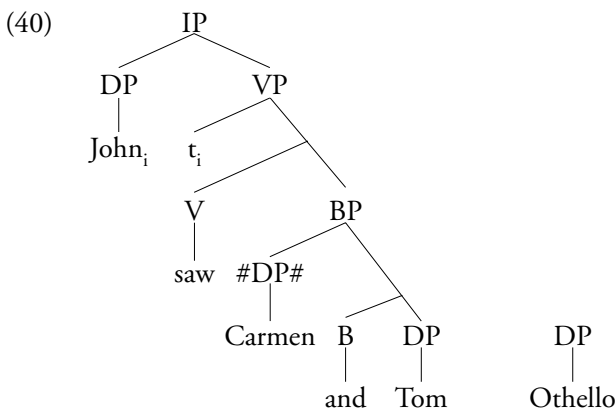
- (38) a. John saw Carmen, and Tom Othello.
- b. *John saw Carmen, and Bill thinks Tom Othello.

With the minimalist parsing algorithm that we have assumed, particularly with a theory of MSO where material is rendered inaccessible for further computation after being spelled out, we can account for why these locality effects are observed in gapping.

Consider sentence (38). The first conjunct is parsed and the coordinator *and* once more is attached low for economy reasons. The next item attached is the DP *Tom*, which is attached low, as a coordinated object. The structure at this point looks like:

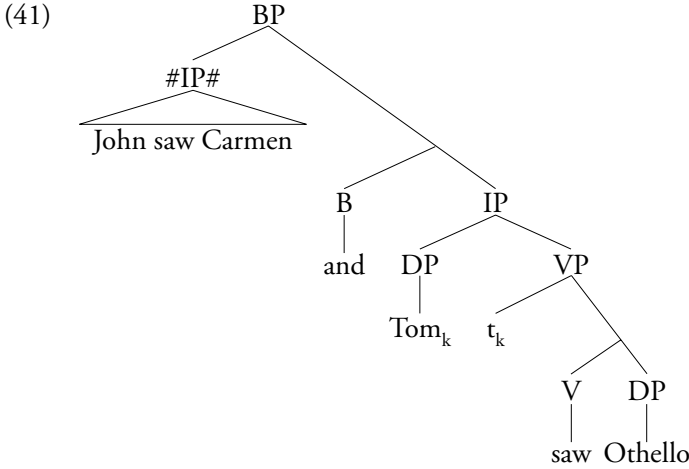


Now, the next word to be attached is the DP *Othello*, but there is no way in which it can be attached to the structure. In this case, we do not have an auxiliary or verb in the current clause, as in the VPE or Pseudogapping examples, that will help us predict a VP. The parser needs to go back to the antecedent clause and use the information about the predicate in that antecedent clause to relate the two arguments *Tom* and *Othello*:

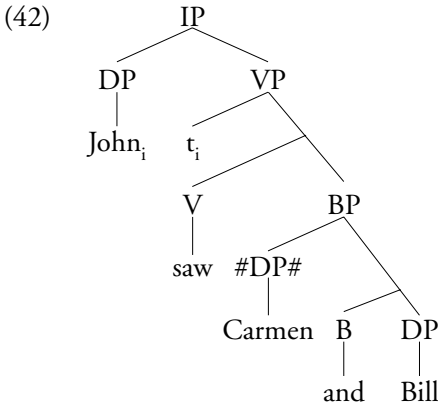


The parser reanalyzes the structure by looking for an antecedent in the c-command path. The antecedent is still available because of the initial mistake of attach-

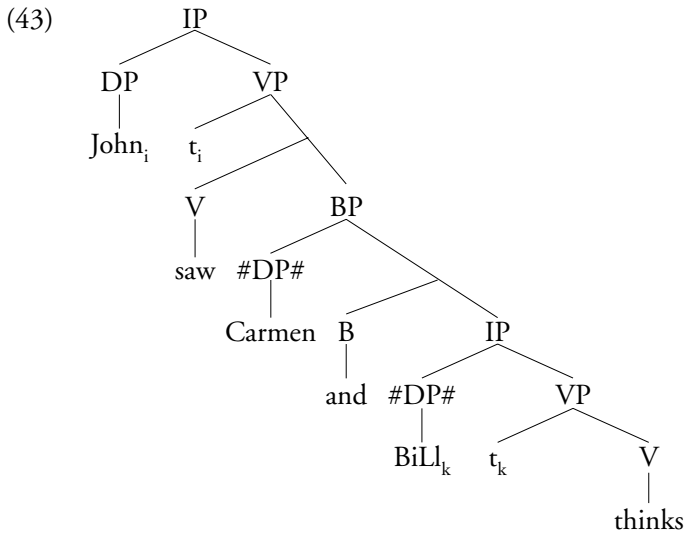
ing low triggered by economy. The coordination attachment is reanalyzed and the resulting structure is that in (41) below:



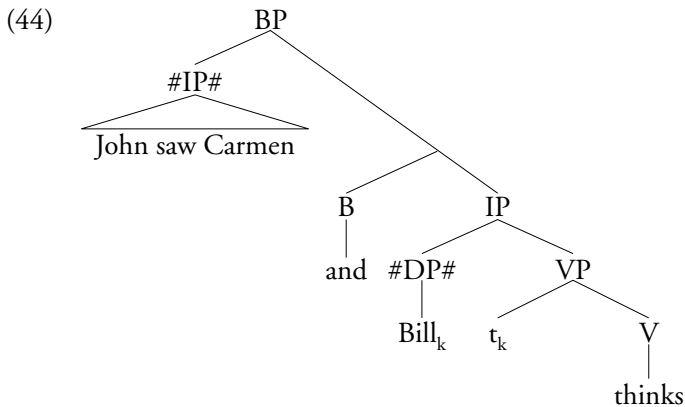
However, in the case of (38) the verb gap cannot be reconstructed, because by the time the parser gets to the gap the antecedent has already been spelled out. Let us see this in some more detail. The coordinator and the DP *Bill* are attached low (for the same reasons we have claimed for the previous examples), as in (42):



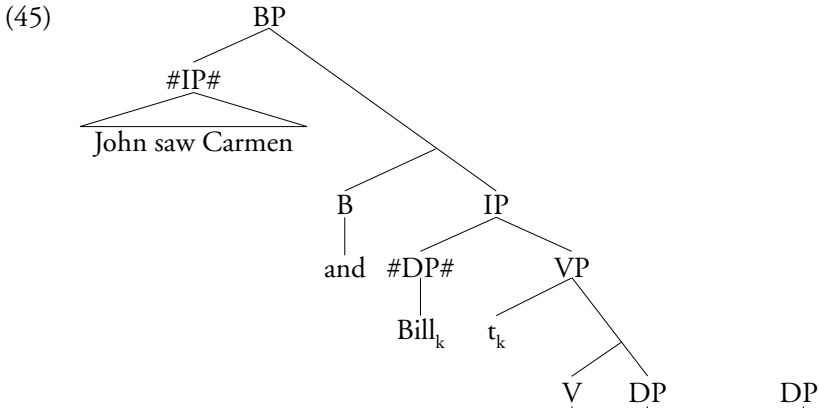
Reanalysis for sentence (38) is triggered when the verb *think* is merged to the already existing structure, since the resulting structure violates the condition on identity of categories for coordination:



This structure is reanalyzed as follows. We have an IP so the only possibility is to attach as a coordinated IP. To do so, the first conjunct (i.e. the antecedent) is spelled out in order to preserve the command-precedence relations:



The parse will follow by attaching the DP *Tom* as in (45) below, but then the next word *Othello* cannot be attached to the existing structure in any way. An auxiliary is not present in the elided clause, as in the VPE or pseudogapping cases, so a VP cannot be predicted and the gap cannot be interpreted. Because the antecedent has already been spelled out, it cannot be accessed to license the gap, resulting in an unacceptable sentence:



Thus, we have seen how the locality effects that gapping cases display can be explained in terms of MSO, and of the minimalist algorithm that we are assuming. In gapping, the antecedent and the elided clause must be local; otherwise, the gap cannot be licensed. The gap depends on the antecedent to be reconstructed.

3. Conclusions

We have offered an account for parsing elliptical constructions which makes use of the minimalist operations: Merge, Move and Spell-Out; which takes into consideration economy issues, and which makes use of local information.

We have accounted for the presence/absence of locality restrictions in ellipsis as a result of the interaction of the following factors: overt tense presence/absence, and of the availability of left-context (i.e. the antecedent), which in turn is a consequence of (i) low initial attachment of coordinates, and (ii) Spell-Out operations which render syntactic structure unavailable. We have seen that in the case of gapping (an ellipsis construction where the relation between the antecedent and the gap must be local) the antecedent needs to be accessed to assign structure to the gap; therefore locality restrictions between the antecedent and the gap apply. However, in the case of VPE and Pseudogapping, a VP may be predicted top-down without resorting to the antecedent, which is only accessed for interpretation purposes. Therefore, these two are not subject to locality restrictions.

Thus, it has been showed that locality restrictions in ellipsis, which have not been properly accounted for from the competence side, do find a natural and satisfactory answer in the processing domain. Locality in ellipsis is reformulated here in terms of c-command: an antecedent is available for gap resolution if it remains in the same c-command path of the gap, or in other words, if they belong to the same CU.

Acknowledgements

I would like to acknowledge the contributions of different people to this paper. First, I have to thank Amy Weinberg who has directed this research. Thanks to Juan Uriagereka and Colin Philips who have also discussed the contents of this paper. I am also indebted to the audiences of the Euresco Conference on Theoretical and Computational Linguistics, the 16th Annual CUNY Conference on Sentence Processing, the CUNY Syntax-Psycholinguistics Supper Club Meeting, the CNL Lab Lunch Meeting, and the Euskal Mintegia. Thanks to Peter Lavery for proof reading and comments. This research was supported by a grant from the Basque Government Department of Education, Universities and Research.

Selected references

- Berwick, B., and A. Weinberg, 1985, "Deterministic Parsing and Linguistic Explanation", *Language and Cognitive Processes* 2, 109-34.
- Carlson, K., 2001, *Parallelism and Prosody in the Processing of Ellipsis Sentences*. Doctoral Dissertation. U. of Massachusetts, Amherst.
- Chao, W., 1987, *On Ellipsis*. Doctoral Dissertation, U. of Massachusetts, Amherst.
- Chomsky, N., 1993, "A Minimalist program for Linguistic Theory". In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, eds. K. Hale and S. J. Keyser. Cambridge, Mass.: MIT Press.
- , 1995, "Categories and transformations". In *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Epstein, S., 1999, "Un-Principled Syntax". In Epstein and Horstein (eds.).
- & N. Hornstein (eds.), 1999, *Working Minimalism*, Cambridge, Mass: MIT Press.
- Fodor, J. D., 1985, "Deterministic Parsing and Subjacency", *Language and Cognitive Processes* 1, 3-42.
- Frazier, L. and Rayner, K., 1982, "Making and Correcting Errors during Sentence Comprehension. Eye Movements in the Analysis of Structurally Ambiguous Sentences", *Cognitive Psychology* 14, 178-210.
- Kayne, R., 1994, *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Munn, A., 1987, "Coordinate Structures and X-bar Theory". *McGill Working Papers in Linguistics* 4-1, 121-40.
- Murguia, E., 2000, "VP Ellipsis as Null Lexicalization: Strict and Sloppy Readings Emerge at Different Stages in the Derivation". In *Maryland Working Papers in Linguistics (UMWPL)* 9, 172-202. M. Guimaraes, L. Meroni, C. Rodriguez, and I. San Martin (eds.).
- , 2004, *Syntactic Identity and Locality Restrictions on Verbal Ellipsis*. Doctoral Dissertation. University of Maryland at College Park, USA.
- Neijt, A., 1980, *Gapping: A Contribution to Sentence Grammar*. Doctoral Dissertation. University of Utrecht, Dordrecht, Holland.
- Pritchett, B., 1992, *Grammatical Competence and Parsing Reference*. Chicago: U.P.
- Uriagereka, J., 1999, "Multiple Spell-Out". In Epstein and Horstein (eds.).
- Weinberg, A., 1999, "A Minimalist Theory of Human Sentence Processing." In Epstein and Horstein (eds.).

IDENTIFYING AND PROCESSING TOPICALIZATION IN DANISH

Sofie Raviv

Department of Foreign Literatures and Linguistics, Ben-Gurion University of the Negev

Abstract*

In Danish there is a surprising asymmetry between pronouns and full DPs: pronouns can always be fronted, whereas full DPs are subjugated to more restrictions. This puzzle has not been addressed in the literature, and I argue that the split-CP analysis as presented by Rizzi (1997) cannot distinguish between pronouns and full DPs when they are marked with a topic feature. The analysis I propose is based on the claim that in Danish the purpose of fronting a constituent is to mark it as the main topic, rather than the need or necessity to check a certain feature. It is furthermore based on the observation that the presence of an overt case marked pronoun, an adverb or a modall/auxiliary enables topicalization. In order to account for the data I propose a (PF) processing constraint on dislocation, which ensures that the derivation has an unambiguous interpretation.

1. Introduction

In Danish pronouns can always be fronted, whereas full DPs are subjugated to more restrictions. The full DP-pronoun asymmetry poses problems for any theory dealing with topicalization. The fact that pronouns can always front, whereas topicalization of full DPs is restricted, seems to be an obstacle for a uniform account of full DPs and pronouns. The goal of this paper is to provide an account of topicalization in Danish. The main hypothesis is that there is a strong correlation between information structure, movement, and word order.

I propose an alternative account of movement, in which topicalization is governed by phonological constraints rather than syntactic principles. This analysis is based on the following observations: first, most topicalized sentences are grammat-

* A version of this paper was presented at the Bilbao/Deusto Student Conference in Linguistics BIDE04. I am thankful for the feedback I received from the participants of this conference. I am also very grateful to Nomi Erteschik-Shir for the meticulous comments and criticism and to Lisa Rochman for helpful suggestions and remarks.

- (5) Q: Tell me about Peter
 a. HAM kender jeg ikke (Danish)
 Him know I not
 b. ?Oto ani lo makir (Hebrew)
 Him I not know
 c. *Him I don't know (English)
 ≈ *I don't know him.*

The answer in Hebrew, however, is possible in a contrastive context (6):

- (6) Q: Tesaper al Peter veSara
 Tell me about Peter and Sara
 A: Oto ani lo makir, aval he meod nexmada
 Him I not know, but she very nice
 ≈ *I don't know him, but she is very nice.*

I have here illustrated that fronting is disallowed in English and only allowed in contrastive contexts in Hebrew. In Danish, on the other hand, fronting of all topics is possible. In the next section, I examine the reasons for fronting in Danish.

2.1. Topic alignment in Danish

Thrane (2003) observes that in Danish 'the topic/comment opposition [is] the favored organization of information' (Thrane 2003: 330).² I claim that this alignment is a result of the preference to place the main topic sentence initially in order to link it to previous discourse. Fronting in Danish can therefore be viewed as a way of marking the main topic of a clause. It follows from this assumption that the sentence initial DP by default will be interpreted as the main topic.

First, I argue that although fronting is often related to contrast, contrast does not, in and of itself, yield fronting.³ As can be observed in (7) stress on contrastive topics is obligatory (Ørsnes 2002) whereas fronting is truly optional.⁴

- (7) Q: Kender du Peter og Sara?
 Do you know Peter and Sara?
 a. SARA kender jeg, men Peter har jeg aldrig mødt
 Sara know I, but Peter have I never met
 b. Jeg kender SARA, men jeg har aldrig mødt Peter
 I know Sara, but I have never met Peter

Furthermore, fronting is not possible when the embedded subject pronoun co-refers to the matrix subject (8). In (8a) the matrix subject is a possible antecedent

² A similar preference can be observed in Spanish.

³ According to Erteschik-Shir (1997: 12) contrast on topics occurs if the context provides a list of possible topics, and a single topic is picked out from that list.

⁴ Stress on topicalized elements is marked with a double underline (STRESS), and contrastive stress is marked with a single underline (STRESS).

ject. Furthermore, I claim that fronting only occurs when the sentence contains two topics and the subject is not the main topic. I have shown that even if the object is the sole topic of the clause it does not move since movement is a matter of last resort.

2.2. The pronoun - full DP asymmetry

In Danish there is a surprising asymmetry between pronouns and full DPs: pronouns can always be fronted.

- (11) a. Peter så mig/dig/ham/hende/os/jer/dem ved festen
 Peter saw me/you/him/her/us/you/them at the party
 b. Mig/dig/ham/hende/os/jer/dem så Peter ved festen
 Me/you/him/her/us/you/them saw Peter at the party

Full DPs appear to be more restricted, and, according to my informants, fronting of a full object DP is not possible if the subject is also a full DP:

- (12) *PETER så Sara
 Peter saw Sara
 ≈ Sara saw Peter.

Danish topicalization provides some quite intriguing facts that have not been observed before: topicalization of full DPs is possible if the sentence contains an adverb/negation (marking the left boundary of the VP), a modal or an auxiliary.

- (13) a. PETER så Sara ikke b. PETER kender Sara selvfølgelig
 Peter saw Sara not Peter knows Sara of course
 ≈ Sara didn't see Peter. ≈ Of course Sara knows Peter.
 c. PETER burde Sara kende fra fjernsynet
 Peter ought to Sara know from TV-the
 ≈ Sara ought to know Peter from TV.
 d. PETER har Sara set over hækken
 Peter has Sara seen over fence-the
 ≈ Sara has seen Peter over the fence.

In sentences without these elements fronting is not possible and the topicalized element will by default be interpreted as the subject even though it is stressed, thus (14b) is infelicitous with the intended topicalized interpretation.

- (14) a. Sara kender Peter fra fjernsynet b. *PETER kender Sara fra fjernsynet
 Sara knows Peter from TV-the Peter knows Sara from TV-the
 ≈ Sara knows Peter from TV.

Bobaljik (to appear) points out that the OVS reading could be achieved in sentences such as (12) and (14) without any added elements, the right context and intonation provided. Nevertheless, as Ørsnes (2002: fn. 8) notes, such constructions 'involving nominals with no case marking and common semantic features, [are] strikingly rare' (Ørsnes 2002: 344), since nominals, unlike pronouns, cannot be disambiguated based on their case properties. According to my informants, however,

an initial DP can only be interpreted as the subject, although pragmatics might help disambiguate in certain sentences. This can be observed in sentence (15), in which *Peter* must be interpreted as the subject, since *bogen* ('the book') clearly cannot function as the subject of the verb *læste* ('read') due to the verb's selectional restrictions.

- (15) *Bogen læste Peter*
 The-book read Peter
 ≈ Peter read the book

In this paper, I focus on the examples where pragmatics and the verb's selectional restrictions cannot disambiguate, and explore what other elements might enable full DPs' fronting.

The full DP-pronoun asymmetry poses problems for any theory dealing with topicalization. The fact that pronouns can always front, whereas fronting of full DPs is restricted, seems to be an obstacle for a uniform account of full DPs and pronouns. In what follows, I explore to what extent syntax can account for the topicalization asymmetry in Danish.

3. Syntax and Topicalization

Discourse functions such as topic and focus have been argued to influence word order (e.g. Rizzi 1997, Szendrői 2004, Thrane 2003). These syntactic theories employ feature checking in their account of topicalization. I argue that fronting cannot be constrained by syntax.

Rizzi (1997) argues that a derivation contains two functional layers instead of one CP layer: the Force phrase (ForceP) and the Finite phrase (FinP). In addition to these two functional layers a derivation can also contain a Focus phrase (FocP) and several Topic phrases (TopP). Yet, these phrases, unlike ForceP and FinP, are optionally projected if needed, i.e., if the derivation contains movement of the topic or focus constituent. FinP relates to the tense and mood of the clause, and its head Fin° is assumed to contain the feature [+ finite], which can be checked either covertly by moving the tense feature of the verb, or overtly as in V2 languages where the verb moves from T° to Fin° (Rizzi 1997: 328, fn. 5). Furthermore, if either FocP or TopP is projected the verb in V2 languages is assumed to move to the head of these phrases, in order to conform to the V2 principle.

Rizzi's Split-CP hypothesis differs from the proposal that the topic constituent moves to spec-CP, where the C° contains the strong top-feature (e.g. Christensen 2003, Schwartz and Vikner 1996). However, both approaches employ the top-feature in more or less the same manner. In the next section I explore the properties of the top-feature.

3.1. A syntactic top-feature

If all topicalized DPs are marked with a top-feature (Rizzi 1997, Chomsky 1995, Christensen 2003), the question is what the properties of this feature are.

Rizzi accounts for the optionality of topicalization by stating that the TopP and

FocP are only projected if needed (Rizzi 1997: 288). Thus movement of a topic constituent is a matter of last resort and ‘it is triggered by the necessity of properly interpreting certain expressions’ (Rizzi 1997: 287) rather than by the need to check a certain feature. Hence Rizzi’s analysis differs from the general assumption of the Minimalist Program (Chomsky 1995) (henceforth MP) in which movement is motivated by the need to check morphological features. I argue that neither the need to check a certain feature nor the ‘necessity to interpret’ can account for the topicalization facts in Danish.

Rizzi’s interpretation of the top-feature by and large follows the suggestion made by Chomsky (1995) and it is argued to be equal to the wh-feature. Moreover, ‘[t]opicalization and focus could be treated the same way [as wh-operator on C^o]. If the operator feature of C is strong, the movement is overt’ (Chomsky 1995: 199). Assuming that the top-feature can be treated on a par with the wh-feature, it is then understood to be uninterpretable on the C^o or whatever head carries the top-feature and thus trigger movement. Yet, it is not clear on what ground the top-feature is assigned. Besides, if all topics are assigned a top-feature, a problem arises in sentences with more than one topic, in which at the most one is fronted.

- (16) Jeg_[TOP] så ham_[TOP] igår
I_[TOP] saw him_[TOP] yesterday

If both topics move, the derivation would render the following word order assuming that the verb moves to the second position.

- (17) *Jeg så ham
I saw him

If on the other hand only one of them projects, the question is which one of them will move. Only if the object moves, will the desired topicalized word order be obtained.

Yet another issue remains unsolved: if a sentence contains more than one top-feature marked DP: which DP should move? Another problem is posed by sentences containing more than one wh-word:

- (18) Who_i saw t_i what?

In sentences containing two wh-marked words as well as in sentences containing two top-marked DPs one element is left with an interpretable feature that cannot be checked. Notwithstanding this problem, the main predicament of treating the top-feature on a par with the wh-feature is that the two features differ with respect to the Minimal Link Condition⁵ (Chomsky 1995: 296). If a sentence contains more than one wh-word the first one will move in order not to violate the Minimal Link Condition (hereinafter MLC).

- (19) *What_i saw who t_i

⁵ The Minimal link Condition: “ α can raise to target K, only if there is no legitimate operation Move β targeting K, where β is closer to K” (Chomsky 1995: 296).

If wh-features and top-features are equivalent in other matters, the prediction is that if the sentence contains more than one top-feature, the closest one, i.e., the subject will move in order not to violate the MLC. However, this prediction does not hold with topicalization, since either one of the topic marked elements can move.

- (20) a. Jeg_[TOP] så ham_[TOP]
I saw him
- b. HAM_[TOP] så jeg_[TOP]
Him saw I

According to the MLC the sentence in (20b), should be ruled out, since movement of the object will violate this condition given that the subject is closer. The discrepancy between the top-feature and the wh-feature needs further investigation, and I conclude that the two features should not be treated as being equivalent.

Even if merging the topicalized DP sentence initially instead of moving circumvents the MLC,⁶ the question still remains which one of the top-marked DPs should be merged sentence initially. Moreover, as it stands one would have to assume different top-features for full DPs and pronouns, since pronouns can always front whereas full DPs are restricted to constructions containing an adverb, a modal, or an auxiliary (13).

I conclude that the split-CP hypothesis as it stands cannot adequately explain the topicalization facts in Danish, and I claim that it is unlikely that any feature based syntactic analysis can account for these data.

3.2. A syntactic constraint on topicalization

One of the major problems for syntactic theories is how to allow for optionality. First of all, morphological case cannot be the feature that constraints topicalization, since full DPs, which are not overtly marked for case in Danish, can also topicalize. Moreover, 3rd person pronouns ('den/det' *it*) do not exhibit a case distinction between nominative and accusative case, yet they do topicalize (21).

- (21) A: Hvad med bogen
What about the book?
- a. Sara læste den b. Den læste Sara
Sara read it It read Sara

I argue that at least in Danish the top-feature is not limited to contrastive topics, contrary to Rizzi's (1997) definition of fronted DPs. In (22) the object pronoun is contrastive and if it remains in situ, as in (22a), the word order is awkward. Fronting is preferred (22b).

- (22) a. ?Jonas kender HAM (men han kender ikke hans nabo)
Jonas knows HIM (but he knows not his neighbor)
- b. HAM kender Jonas (men han kender ikke hans nabo)
HIM knows Jonas (but he knows not his neighbor)
≈ *Jonas knows him but he doesn't know his neighbor.*

⁶ See Platzack (2004) for an analysis of initial merging and the MLC.

The split-CP hypothesis and the assumption that contrastive topics have to move can explain the preference for (22b), as the contrastive pronoun *ham* ('him') is topic marked, and thus predicted to front. Nonetheless, recall that full DPs cannot topicalize even though they are contrastive, thus (23b) is infelicitous with the topicalized interpretation although according to Rizzi the full DP *Peter* should be able to move.

- (23) a. Jonas kender PETER (men han kender ikke hans nabo)
 Jonas knows PETER (but he knows not his neighbor)
 b. *PETER kender Jonas (men han kender ikke hans nabo)
 PETER knows Jonas (but he knows not his neighbor)
 ≈ *Jonas knows Peter but he doesn't know his neighbor.*

Rizzi would wrongly predict (23b) to be possible since the topic object has to move in order to check the strong top-feature on Top^o. On the other hand, topicalization of full DPs is possible if the sentence contains a modal.

- (24) a. Jonas burde kende PETER (men ikke hans nabo)
 Jonas should know PETER (but not his neighbor)
 b. PETER burde kende Jonas (men ikke hans nabo)
 PETER should know Jonas (but not his neighbor)
 ≈ *Jonas should know Peter but not his neighbor.*

As it stands, Rizzi's approach cannot account for these facts. Another problem with the top-feature assignment is the fact that non-contrastive topics can front as well. In (25), the object is the topic in both answers, however once again movement is optional.

- (25) Q: Hvad med Jonas?
 What about Jonas?
 a. Sara så ham over hækken b. HAM så Sara over hækken
 Sara saw him over the-fence Him saw Sara over the-fence

According to Rizzi the two answers in (25) will have different derivations, and despite the fact that one must assume that in both cases the object is the topic, it can only have been assigned this feature in (25a). The only way out would be to assign a strong top-feature in the former case and a weak one in the latter, yet such assignment is arbitrary. I therefore conclude that contrast does not constrain topicalization in Danish, and other constraints are needed in order to account for the Danish topicalization data.

4. The syntax and phonology of topicalization

Having established that syntactic feature checking approaches fail with respect to the data presented here, I turn to an alternative account of movement, in which topicalization is governed by phonological constraints rather than syntactic principles. The argument that movement is related to the phonology is not new (e.g. Holmberg 1999, 2000, Zubizarreta 1998). However, my proposal differs essentially from that of e.g. Holmberg (2000).

Holmberg introduces phonology into narrow syntax in connection with Stylistic Fronting. I adopt his view that '[s]yntactic categories enter syntactic derivation in the form of words, that is triples of formal, semantic and phonological features' (Holmberg 2000: 16). In the analysis presented here the top-feature is incorporated into the grammar and assigned optionally as a part of merge after being selected freely from the lexicon. Given that top-features are available in the lexicon, the derivation does not violate the Condition of Inclusiveness (Chomsky 1995: 288).

The main problem with Holmberg's approach, as I see it, is the fact that phonology is incorporated into narrow syntactic processes; a suggestion that entails syntactic processes still constrain phonological features. Thus nothing is gained with respect to the problems outlined in the previous section concerning the top-feature and the restrictions on fronting that a syntactic analysis presents. As an alternative I propose that movement is not constrained by syntax but is rather believed to take place in p(honological)-syntax (Erteschik-Shir 2003). Consequently, movement is subject to phonological restrictions rather than syntactic principles, that is, movement is either to the left or to the right edges of sentences (Erteschik-Shir, to appear). I propose that all optional reordering processes take place in the phonology and I argue that topicalization is displacement in the phonology on a par with object shift (Erteschik-Shir 2001, 2003, to appear, Holmberg 1999).

Although processing is often assumed to be extra-grammatical I argue here that processing is an integral part of the grammar. In Danish, no syntactic principles constrain topicalization; therefore all topicalized sentences are grammatically well-formed sentences. In spite of this, the intended interpretation is not always attainable in topicalized sentences, and I argue that the only way to account for this is by incorporating processing into the grammar. Moreover, topicalization is optional, a fact that cannot be accounted for in traditional syntactic theories. I argue that processing is phonological, since it is subject to restrictions on pronunciation, stress, and intonation. In what follows, I argue that the processing constraint I introduce accounts for the Danish data.

Recall that the presence of a pronoun (26), negation (27a), adverb (27b), modal (27c) or auxiliary (27d) enables topicalization in Danish:

- | | | | |
|---------|--|----|---|
| (26) a. | <u>HAM</u> så Sara
Him saw Sara | b. | <u>SARA</u> så han
Sara saw he |
| (27) a. | <u>PETER</u> så Sara ikke
Peter saw Sara not
≈ <i>Sara didn't see Peter.</i> | b. | <u>PETER</u> kender Sara selvfølgelig
Peter knows Sara of course
≈ <i>Of course Sara knows Peter.</i> |
| c. | <u>PETER</u> burde Sara kende fra fjernsynet
Peter ought to Sara know from TV-the
≈ <i>Sara ought to know Peter from TV.</i> | | |
| d. | <u>PETER</u> har Sara set over hækken
Peter has Sara seen over fence-the
≈ <i>Sara has seen Peter over the fence.</i> | | |

I propose that what overt case and the added elements in (27) have in common is that they force the OVS interpretation on fronted objects. Furthermore, topicalization in Danish is allowed only when the resulting OVS structure is unambiguous. These two facts together argue for an account of topicalization that involves disambiguating factors. In order to formalize this idea, I propose a processing constraint (28) adapted from Erteschik-Shir (2003).⁷ This constraint is part of UG, however disambiguating factors are language specific and can even vary across dialects with differing morpho-phonological features.

(28) ID(entification)

In a string X...Y, ID the X as the subject and the Y as the object if neither is identified otherwise.

(Erteschik-Shir 2003)

This constraint applies to the final output of phonological strings of words and not to syntactic structures. The string of words (order irrelevant) allows for intervening material [a X b V c Y d], where a,b,c,d are non-arguments. In what follows I outline several “identifiers” related to topicalization in Danish.

First of all morphological case functions as an identifier: In Danish, all subject pronouns in canonical subject position are nominative. All other pronouns, (disregarding genitive case) are accusative. Therefore, all fronted objects are accusative (29b).⁸

- (29) a. Peter så mig/dig/ham/hende/os/jer/dem ved festen
 Peter saw me/you/him/her/us/you/them at the party
 b. Mig/dig/ham/hende/os/jer/dem så Peter ved festen
 Me/you/him/her/us/you/them saw Peter at the party

Surprisingly, extracted subjects are also accusative (30b).

- (30) a. Peter tror jeg/du/han/hun/vi/I/de vinder løbet
 Peter thinks I/you/he/she/we/you/they (will) win the race
 b. Mig/dig/ham/hende/os/jer/dem tror Peter vinder løbet
 Me/you/him/her/us/you/them thinks Peter (will) win the race

Accusative case on the sentence initial pronoun identifies it as a fronted object (Ørsnes 2002). If the sentence initial DP were nominative it would be the subject.

⁷ The idea that such a processing constraint is imposed on the grammar has been around since.

⁸ Norwegian differs from Danish in that morphological case on pronouns cannot function as an identifier. I ascribe this to the fact that the Norwegian pronoun paradigm differs from the Danish paradigm in that pronouns do not differ in their nominative and accusative forms. Thus the following sentence (i) is potentially ambiguous between a fronted and non-fronted interpretation:

- (i) Han så Peter
 He saw Peter
 ≈ He saw Peter/ Peter saw him

I suggest that intonation might play a disambiguating role in these sentences, but further studies are needed in order to solve this problem.

Nonetheless, there is another interesting fact, which is not captured in Ørnsnes' analysis: subject as well as object pronouns can function as identifiers:

- (31) a. HAM så Sara
Him saw Sara
- b. SARA så han
Sara saw he

In (31a) the accusative pronoun identifies the sentence initial pronoun as object. In (31b) the nominative pronoun in the canonical object position identifies it as subject, and consequently the sentence initial full DP is the fronted object.

The second form of identifiers mentioned above is the left boundary identifier: negation and adverbs (left adjoined to VP)⁹ identify the argument immediately preceding them as the subject independently of topicalization.

- (32) a. Sara så ikke Peter
Sara saw not Peter
- b. PETER så Sara ikke
Peter saw Sara not
≈ Sara didn't see Peter.

In both sentences in (32) Sara is identified as the subject, since it is the first argument preceding negation. The only exception to this generalization is sentences with object shift, in which the first argument preceding the negation is the object.

- (33) Jeg så ham ikke
I saw him not
≈ I didn't see him.

The ID constraint still imposes an OVS interpretation in these cases given that pronouns are always marked for case (Erteschik-Shir, to appear), thus the morphological case on the pronoun functions as an identifier and these sentences are therefore unambiguous.

The final form of identifiers that I will discuss here are modal and auxiliary identifiers. Generally, the occurrence of a modal or an auxiliary enables OVS:

- (34) a. SARA har Peter set
Sara has Peter seen
≈ Peter has seen Sara.
- b. SARA burde Peter have set
Sara ought Peter have seen
≈ Peter ought to have seen Sara.

The general intuition is that the auxiliary and the modal identify the subject because of the agreement marking on the verb/modal/auxiliary. Yet, in Danish there is no overt agreement, and non-overt agreement does not identify the subject: the sentence in (35) is unambiguous despite the fact that the subject and object share the same agreement features. If agreement functioned as an identifier one would expect the sentence in (35) to be ambiguous.

- (35) ANDERS har Peter set
Anders has Peter seen
≈ Peter has seen Anders.

⁹ Holmberg (2000: 16) suggests that the negative element in Mainland Scandinavian is classified as an adverb, thus it is no surprise that left adjoined adverbs and negation function the same way with respect to identification.

Moreover, if agreement could identify the subject, one would predict that full DPs could topicalize with no restrictions, which is not the case: the sentence in (36) has the unambiguous SVO interpretation.

- (36) ANDERS så Peter
Anders saw Peter

As an alternative I propose that the main verb identifies the first argument adjacent to it as the subject in all sentences: in (37) the sentences in the subject *Peter* is the first argument adjacent to the main verb. As a result, (37a-b) have the unambiguous interpretation of SVO, whereas (37c) has the unambiguous OVS interpretation.

- (37) a. Peter så Anders b. Peter har set Anders
Peter saw Anders Peter has seen Anders
c. ANDERS har Peter set
Anders has Peter seen
≈ *Peter has seen Anders.*

The only exception to this generalization is topicalized sentences containing an adverb, in which the first argument preceding the main verb is the object (38). However, I claim the ID constraint still imposes an unambiguous OVS interpretation in these instances since the adverb identifies the first argument preceding it as the subject. Thus, these sentences remain unambiguous.

- (38) SARA så Peter ikke
Sara saw Peter not
≈ *Peter didn't see Sara.*

I have argued for a processing constraint (ID), which constrains topicalization to unambiguous strings. I have furthermore illustrated the different identifiers in cases in which selectional restrictions and pragmatics do not disambiguate. In the following section, I will discuss the p-syntax of topicalization and illustrate how the ID constraint proposed here interacts with the grammar.

4.1. The p-syntax of topicalization

In this section I illustrate the mechanisms of the syntax and the phonology of p-syntax based on the analysis of object shift (Erteschik-Shir, to appear). I argue that the analysis adopted in this paper accounts for all the different topicalization facts presented here. My main point is that in principle topicalization is always possible, if the object or the embedded subject is the main topic. I argue that the reason topicalization is not possible in all constructions is because the ID constraint imposes a SVO interpretation in the absence of markings to the contrary.

I propose the constraints in.¹⁰ These constraints are absolute and cannot be violated.

¹⁰ These constraints are adapted from Erteschik-Shir (to appear), nonetheless the versions presented here are my own.

- (39) V2: finite verb is in second position
 TOP(IC) ALIGN: Place main topic sentence initially
 ID: Subject and object must be identified

Sentence (40) illustrates a simple sentence in which no movement takes place, yet all the constraints are satisfied; the verb occurs in second position and the topic is placed sentence initially. The ID constraint is also satisfied since both the subject and the object are marked for case, yielding an unambiguous SVO interpretation.

- (40) Jeg så ham
 I saw him
- a. [jeg_{TOP} så ham_{TOP}]_{FOC}
 b. V2: [jeg_{TOP} så ham_{TOP}]_{FOC}
 c. TOP ALIGN: [jeg_{TOP} så ham_{TOP}]_{FOC}
 d. ID: [jeg_{TOP/SUBJ} [så ham_{TOP/OBJ}]]_{FOC}

Unlike in Optimality Theory the constraints are not ordered, as can be observed in the derivation in (41).

- (41) Jeg så ham
 I saw him
- a. [jeg_{TOP} så ham_{TOP}]_{FOC}
 b. TOP ALIGN: [jeg_{TOP} så ham_{TOP}]_{FOC}
 c. V2: [jeg_{TOP} så ham_{TOP}]_{FOC}
 d. ID: [jeg_{TOP/SUBJ} [så ham_{TOP/OBJ}]]_{FOC}

However, another ordering of the derivation could in certain instances yield uninterpretable results. This is best illustrated in sentences in which movement is obligatory, as is the case for fronted objects (42).

- (42) HAM så jeg
 Him saw I
- a. [jeg_{TOP} så ham_{TOP}]_{FOC}
 b. TOP ALIGN: ham_{TOP} [jeg_{TOP} så]_{FOC}
 c. V2: dna
 d. ID: *ham_{TOP/OBJ} [jeg_{TOP/SUBJ} så]_{FOC}

In (42) the V2 constraint is violated as the verb does not occur in second position because it is prevented from moving as it does not target an edge position, which is necessary in a phonological account (Erteschik-Shir, to appear). The ID constraint still identifies the subject and the object accordingly because the pronouns are marked for case.

It follows from the analysis that in a sentence without any identifiers, topicalization is not possible. So far I have only presented sentences with fronted pronouns. Since these are always marked for case in Danish, they are identified as either the subject if nominative or as the object if accusative. In Danish full DPs, which are not marked for case, cannot topicalize as shown in (43). In these cases the ID constraint imposes a SVO interpretation.

- (43) *PETER så Sara
Peter saw Sara
≈ *Sara saw Peter.*

a.	[Sara _{TOP} så Peter _{TOP}] _{FOC}
b. V2:	[så Sara _{TOP} Peter _{TOP}] _{FOC}
c. TOP ALIGN:	Peter _{TOP} [så Sara _{TOP}] _{FOC}
d. ID:	*Peter _{TOP/SUBJ} [så Sara _{TOP/OBJ}] _{FOC}

The sentence in and of itself is a well-formed sentence. However the intended OVS interpretation is not available.

(44) shows a non-topicalized sentence in which a negative adverb identifies the first argument preceding it as the subject:

- (44) Peter så ikke Sara
Peter saw not Sara
≈ *Peter didn't see Sara.*

a.	[ikke Peter _{TOP} så Sara _{TOP}] _{FOC}
b. V2:	[så ikke Peter _{TOP} Sara _{TOP}] _{FOC}
c. TOP ALIGN:	[Peter _{TOP} så ikke Sara _{TOP}] _{FOC}
d. ID:	[Peter _{TOP/SUBJ} så ikke Sara _{TOP/OBJ}] _{FOC}

In (44b) the verb moves to second position in order to satisfy the V2 constraint. The topic, in this case the subject, moves in order to align the topic with the focus structure and link it to the previous discourse. The negative element identifies *Peter* as the subject, since it is the first argument preceding it. It follows that the derivation has an unambiguous interpretation. However, this derivation poses a slight problem for the analysis presented here, since the subject in topicalized sentences would have to move to a non-edge position, a move that is not possible in the phonology.

- (45) SARA så Peter ikke
Sara saw Peter not
≈ *Peter didn't see Sara.*

a.	[ikke Peter _{TOP} så Sara _{TOP}] _{FOC}
b. V2:	[så ikke Peter _{TOP} Sara _{TOP}] _{FOC}
c. TOP ALIGN:	<u>SARA</u> _{TOP} [så ikke Peter _{TOP}] _{FOC}
d. <u>SUBJ MOVE?</u> :	<u>SARA</u> _{TOP} [så Peter _{TOP} ikke] _{FOC}
e. ID:	<u>SARA</u> _{TOP/OBJ} [så Peter _{TOP/SUBJ} ikke] _{FOC}

In what follows, I revise the analysis slightly to accommodate for the above raised problem as well as to explain the fact that in Danish all fronted DPs are stressed whereas subject topics, however, are not stressed. I argue that the stress discrepancy between subject and topicalized DPs is explained by the TOP ALIGN constraint. Consequently, if a DP moves in order not to violate TOP ALIGN, it receives stress. In (46) no movement of the subject takes place, thus the subject is not stressed.

(46) Jeg så Peter
I saw Peter

- a. [jeg_{TOP} så Peter_{TOP}]_{FOC}
 b. V2: [jeg_{TOP} så Peter_{TOP}]_{FOC}
 c. TOP ALIGN: [jeg_{TOP} så Peter_{TOP}]_{FOC}
 d. ID: [jeg_{TOP/SUBJ} så Peter_{TOP/OBJ}]_{FOC}

Conversely, in the derivation in (47) the object moves in order not to violate the TOP ALIGN constraint, and is accordingly assigned stress. (Erteschik-Shir, to appear, Raviv, in prep)

(47) HAM så jeg
Him saw I
≈ *I saw him.*

- a. [jeg_{TOP} så ham_{TOP}]_{FOC}
 b. V2: [så jeg_{TOP} ham_{TOP}]_{FOC}
 c. TOP ALIGN: HAM_{TOP} [så jeg_{TOP}]_{FOC}
 d. ID: HAM_{TOP/OBJ} [så jeg_{TOP/SUBJ}]_{FOC}

Due to the prosodic discrepancy between subjects DPs and fronted DPs, I suggest that fronted DPs move in order not to violate TOP ALIGN, subjects, on the hand do not move.

Yet, in order to account for the fact that the subject appears to have left its base-generated position within the VP, and occurs to the left of the negation I adopt Erteschik-Shir's (to appear) analysis of adverb projection.¹¹ Following Bobaljik (2002), Chomsky (2001), and Áfarli (1997), Erteschik-Shir proposes that adverbials do not move; instead they are merged on a separate level (in a third dimension), and they can linearize either to the left or the right of the verb depending on processing and pronunciation constraints.

(48) a. Peter så han ikke b. HAM så Peter ikke
Peter saw him not Him saw Peter not
≈ *Peter didn't see him.*

The outcome of the analysis is that adverbs in non-topicalized sentences are projected to the left of the object (48)b whereas in topicalized sentences they are projected to the right of the subject (48)a. Thus, I assume that in non-topicalized sentences the subject does not move, which means that it is not stressed. In topicalized sentences, on the other hand, the object/embedded subject moves and it will therefore receive stress.

In this section I have argued for a phonological account of topicalization. I have demonstrated that although topicalization is always possible, the intended interpretation is not always available. I have imposed a processing constraint on the grammar, which I argue can account for the topicalization facts presented here.

¹¹ See Erteschik-Shir (to appear) for arguments against subject-movement.

5. Conclusion

The goal of this paper was to account not only for the optionality of topicalization but also for the apparent asymmetry of pronouns and full DPs in Danish. I have argued that syntactic feature checking approaches face explanatory difficulties when confronted with this data. Instead I have provided an alternative account in which processing is an integral part of the grammar. I have argued for a processing constraint, which constrains topicalization in Danish and provides a uniform account of this asymmetry. If indeed topicalization is restricted by this constraint, my argument will have repercussions for syntactic theory.

References

- Bobaljik, J. D., 2002, "A-chains at the PF interface: copies and 'covert' movement", *NLLT* 20, 197-267.
- , to appear, "Re: CycLin and the role of PF in Object Shift", *Theoretical Linguistics*.
- Casielles-Suarez, E., 1997, *Topic, Focus and Bare Nominals in Spanish*. Department of Spanish and Portuguese, University of Massachusetts, Amherst: Ph.D.
- Chomsky, N., 1995, *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- , 2001, "Derivation by phase". *Ken Hale - A Life in Language*, ed. by M. Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Christensen, K. R., 2003, "On the Synchronic and Diachronic Status of the Negative Adverbial *ikke/not*", *Working Papers in Scandinavian Syntax WPSS* 72.1-53.
- Erteschik-Shir, N., 1997, *The Dynamics of Focus Structure*. Cambridge: Cambridge U.P.
- , 2001, "P-syntactic motivation for movement: imperfect alignment in Object Shift", *Working Papers in Scandinavian Syntax* 68. 49-73.
- , 2003, "The syntax, phonology and interpretation of the information structure primitives Topic and focus". Paper presented at GLOW workshop: *Information structure in generative theory vs. pragmatics*. The University of Lund, Sweden.
- , to appear, "The sound patterns of syntax: the case of object shift", *Theoretical Linguistics*.
- Holmberg, A., 1999, "Remarks on Holmberg's Generalization", *Studia Linguistica* 53.1, 1-39.
- , 2000, "Scandinavian Stylistic Fronting: How can any category become an expletive", *LI* 31, 445-484.
- Jakobson, R., 1936 [1971], "Implications of language universals". *Selected Writings*. The Hague: Mouton. 2: 580-92.
- Ørnsnes, B., 2002, "The Distribution of Personal Pronouns in Danish: A Lexical-Functional Approach", *Copenhagen Studies in Language* 28: 153-61.
- Platzack, Ch., 2004, "Cross-linguistics word order variation and the left periphery: The case of object first in main clauses". *Peripheries: Syntactic Edges and their Effects*, ed. by D. Adger, C. de Cat and G. Tsoulas. Studies in Natural Languages and Linguistic Theory 59, 191-210. Dordrecht: Kluwer Academic Publishers.
- Prince, A. and Smolensky, P., 1993, *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick, N.J.: Technical Report TR-2. Center for Cognitive Science, Rutgers University.

- Raviv, S., in prep., "Identifying and processing topicalization in Danish", *Foreign Literatures and Linguistics*, Ben-Gurion University of the Negev: MA.
- Reinhart, T., 1981, "Pragmatics and linguistics: an analysis of sentence topics", *Philosophica* 27. 53-94.
- Rizzi, L., 1997, "The Fine Structure of the Left Periphery". In *Elements of Grammar*, ed. by L. Haegeman. Dordrecht: Kluwer.
- Szendrői, K., 2004, "Focus and the interaction between syntax and pragmatics [2004/3]", *Lingua* 114, 229-54.
- Schwartz, B. D., and Vikner, S., 1996, "The Verb Always Leaves IP in V2 Clauses". In *Parameters and Functional Heads*, ed. by A. Belletti and L. Rizzi, 11-62. Oxford: Oxford U.P.
- Thrane, T., 2003, "Breakstructures", *Journal of Linguistics* 39, 327-72.
- Zubizarreta, M. L., 1998, *Prosody, Focus and Word Order*. Vol. 33: Linguistic Inquiry Monograph. Cambridge, MA: M.I.T. Press.

PREDICATE NOMINALS IN EVENTIVE COPULAR SENTENCES*

Isabelle Roy

University of Southern California

1. Introduction

This paper presents a comparative study of predicate nominals in eventive copular sentences. I am concerned with languages that grammaticalize the contrast between *eventive* and *non-eventive* copular sentences, and I discuss here three cases: the optionality of the indefinite article in French and German, the two verbs “to be” in Irish, and case marking in Russian.

Setting aside here non-eventive predications, I show that eventive predicate nominals exhibit non trivial semantic and syntactic similarities that argue in favor of a unified account. I argue that the properties of eventive predicate nominals can be derived from the assumption that the small clause is headed by a functional head Pred (cf. Bowers 1993), and that Pred must be realized. Irish and Russian provide evidence that Pred can be spelled-out as a preposition P, whereas French and German show that it can be phonologically realized as N, through head movement.

This paper is organized as follows. In section 2, I will discuss the alternation between the variants with and without the indefinite article in French and German. I will show that the choice of one or the other of the variants is sensitive both to the nature of the predication, and to whether the property denoted by the predicate is perceived as permanent or transient. In section 3, I will discuss how the same contrast is rendered by the alternation between the two verbs ‘to be’ in Irish. I will further show that the eventive predicate nominals in Irish share with their French counterpart similarities that suggest that the difference between the two languages is only superficial. In section 4, I will argue that the contrast between permanent vs. transient predicates must be a syntactic one and cannot be made at the level of the lexicon. This discussion will allow me to clarify the notion of ‘eventive copular sentences’. In section 5 I will present a unified analysis of eventive nominal predication and argue that there are two ways the predicational head of the small clause is realized: by a preposition (either overt or covert), or by head movement of the predicate

* I would like to thank Hagit Borer and Roumyana Pancheva for our discussions, their comments and suggestions. I am also grateful to James McCloskey for discussing the Irish data with me; and to Patricia Cabredo Hofherr, Thomas Leu and Martin Salzmann for their help with the German data. Any error of interpretation is solely mine.

N from its original position. Finally, it will be shown that English patterns, surprisingly enough, with Irish/Russian; such a result will present interesting outcomes in terms of the differences between English and French.

2. Optionality of the indefinite article in French and German

2.1. Predication and identification

Predicate nominals can appear with or without the indefinite article in French¹ and German. The apparent optionality of the article is illustrated in (1) and (2) for the two languages respectively:

- | | |
|-------------------------------|----------------------------------|
| (1) Paul est (un) professeur. | (2) Paul ist (ein) Schauspieler. |
| Paul is (a) professor | Paul is (an) actor |
| "Paul is a professor." | "Paul is an actor." |

The variants with and without the indefinite article have radically different properties. One crucial difference is the "function" (in the sense of Higgins 1979)² in which each sentence can be used (Kupferman 1979 and Pollock 1983 for French): the variant without article appears exclusively in *predicational* sentences; whereas the variant with the indefinite article is used in *identificational* sentences only. The predicational and identificational readings can be distinguished on the basis of the type of question they constitute an appropriate answer to. Consider the sentence *John is a teacher* in English. On the one hand, it can be an answer to *What does John do?*; and thus, it can get a predicational reading, equivalent to *John teaches*. On the other hand, it is also a felicitous answer to the question *Who is John?*; and accordingly is ambiguous with an identificational interpretation. In French however *Paul est un professeur* "Paul is a professor" is not ambiguous. It can only answer the question *Who is Paul?* (see contrast (3) and (4)) and hence is identificational only.³ By opposition, the variant without article is the only option to the question *What does Paul do?*, and therefore the only option in predicational sentences:

- | | |
|-------------------|----------------------------|
| (3) Qui est Paul? | Paul est *(un) professeur. |
| who is Paul | Paul is a professor |
| "Who is Paul?" | "Paul is a professor." |

¹ Similar contrasts exist also in other Romance languages such as Italian, Spanish and Portuguese.

² Since the work of Higgins (1979), four types of copular sentences are commonly distinguished. The two types relevant for our discussion are illustrated in (ia-b):

- | | |
|-----------------------------|---------------------------------------|
| (i) a. <i>Predicational</i> | John is tall. |
| b. <i>Identificational</i> | That man is my teacher. |
| c. <i>Specificational</i> | The problem is his tie. |
| d. <i>Identity</i> | The morning star is the evening star. |

³ The use of the variant with article to identify a person extends to definite predicates also, as in *Paul est le directeur de l'usine* "Paul is the factory director", for instance. I will not discuss these sentences here.

- (4) Que fait Paul dans la vie? Paul est (*un) professeur.
 what does Paul in the life Paul is a professor
 "What does Paul do for a living?" "Paul is a professor."

Identical contrasts can be seen in German as well, as shown by the examples (5-6):

- (5) Wer ist Klaus? Klaus ist *(ein) Schauspieler.
 who is Klaus Klaus is an actor
 "Who is Klaus?" "Klaus is an actor."
 (6) Was macht Klaus? Klaus ist (*ein) Schauspieler.
 what does Klaus Klaus is an actor
 "What does Klaus do?" "Klaus is an actor."

The distribution of predicate nominals in contexts known to be predicational only (see also Rouveret 1998) further shows that only the variant without the article is predicational. First, only the bare variant can be pronominalized by the clitic *le* "it" in French (7):

- (7) Marie voulait être (*une) infirmière, mais elle ne le sera jamais.
 Mary wanted to be a nurse, but she NEG it will-be never
 "Mary wanted to be a nurse, but she'll never be one."

Second, only the bare variant can be the antecedent of a non-restrictive relative clause introduced by *ce que* "which" (8):

- (8) Jean est (*un) médecin, ce que son frère n'est pas.
 John is a doctor, which his brother NEG is not
 "John is a doctor, which his brother is not."

Third, only the bare variant is allowed as predicate of a small clause. Consider, for instance, verbs that select small clause complements as *croire* "to believe" (9) and raising verbs as *s'avérer* "to prove (to be)/ to turn out" (10):

- (9) Je croyais Paul (*un) médecin.
 I believed Paul a doctor
 "I believed that Paul was a doctor."
 (10) Jean s'avèrait (*un) médecin.
 John turned out a doctor
 "John turned out / proved to be a doctor."

2.2. Permanent vs. transitory properties

The interpretational difference between the two variants has been argued to relate to a contrast between transitory *vs.* permanent properties. Kupferman (1991) shows that only the variant without the article can appear in constructions that admit exclusively transitory properties, such as absolutive constructions and existential constructions. Additional support for this dichotomy comes from the fact that only the bare variant is compatible with spatio-temporal modifications (see Roy 2001). First, only bare predicates are compatible with temporal modifications (11-12).

When the property of ‘being a N’ is relative to particular situations, the variant without article is the only option:

- (11) Paul est (*un) médecin le jour, et (*un) chanteur la nuit.
 Paul is a doctor the day, and a singer the night
“Paul is a doctor by day and a singer by night.”
- (12) Paul est (*un) traducteur à ses heures libres.
 Paul is a translator at his hours free
“Paul is a translator on his spare time.”

In a similar way, only the bare variant accepts locative modification (13):⁴

- (13) Paul est (*un) médecin à Paris.
 Paul is a doctor in Paris
“Paul is a doctor in Paris.”

Second, the bare variant does not give rise to “lifetime effects” (cf. Musan 1995) in the past tense. Consider both sentences in (14), only (14b) entails that the individual Paul is now dead. In (14a) no such entailment is made, Paul simply does not practice medicine anymore:

- (14) a. Paul était médecin. (absence of lifetime effects)
 Paul be.PAST doctor
“Paul was a doctor.”
- b. Paul était un médecin. (lifetime effects)
 Paul be.PAST a doctor
“Paul was a doctor.”

Third, only the bare variant can receive an interruptive reading (cf. Fernald 1994). Again, the bare variant is the only option when the state of ‘being an N’ is temporally restricted (15):

- (15) Paul sera (*un) professeur trois fois dans sa vie.
 Paul be.FUT a professor three times in his life
“Paul will be a professor three times in his life.”

Finally, only the bare variant entails the actual practice of the activity denoted by the predicate nominal as shown by the contrast in (16-17). Consider the oddness of example (16): when the actual practice of the activity is negated, the sentence results in a contradiction. Contrastively, the property of being “a N” is independent of the practice of an activity (17):

- (16) # Jean est médecin, mais il ne pratique plus.
 John is doctor but he does not practice anymore
 intended: *“John is a doctor, but he does not practice anymore.”*

⁴ Sentence (13) with the variant with the article has a reading where *à Paris* “in Paris” is interpreted as ‘according to the people in Paris’. However, in this case, the locative PP is not used as a locative binder for the predicate.

- (17) Jean est un médecin, mais il ne pratique plus.
 John is a doctor but he does not practice anymore
 “John is a doctor, but he does not practice anymore”

The contrastive properties of the two variants of predicate nominals in French can be summarized in (18). The same properties hold for German:

- (18)

	variant without article	variant with article
a) Identificational reading	no	yes
b) Lifetime-effects	no	yes
c) Spatio-temporal modification	yes	no
d) Activity reading	yes	no
e) Small clause predicate	yes	no

3. Predicate nominals in copular sentences in Irish

3.1. The two verbs “to be”

The contrast between permanent and temporary properties is rendered by the choice of one of the two verbs “to be” in Irish. Irish has two forms for the verb “to be”, the so-called “substantive auxiliary” *bí* and the copula *is*.⁵ The copula *is* is traditionally said to predicate essential or inherent properties, while *bí* is used to predicate more temporal properties (Stenson 1981). A predicate nominal in construction with *is* expresses “a defining characteristic” of the subject; while in construction with *bí* it rather describes “what someone does, is more dynamic in concept, and suggests a state rather than a property” (Stenson 1981: 94):

- (19) Is shagart è mo dheartháir.
 COP-IS priest AGR my brother
 “My brother is a priest.”
- (20) Tá mo dheartháir ina shagart.
 COP-BI my brother in.AGR priest
 “My brother is a priest.”

⁵ The opposition between *bí* and *is* is analogous to the contrast between *ser* and *estar* in Spanish and Portuguese, as previously noted by Ó Máille (1912) (cited in Stenson 1981 and Ó Siadhail 1989). A further similarity between the two languages is that predicate nominals in Spanish cannot occur with *estar* (*Juan está médico ‘Juan is(estar) doctor’), and must be introduced by a preposition (Juan está de médico ‘Juan is(estar) of doctor’). Compare with (19-20) in Irish. However, Spanish is more complex than Irish in the sense that it does not exhibit simply a binary distinction with the copulas, but rather a three-way distinction, as it allows in addition, for the optionality of the article with the copula *ser* as it is the case in French and German (Juan es (un) médico ‘Juan is(ser) (a) doctor’)). Thus, Spanish has three types of copular sentences with predicate nominals: *ser* \emptyset NP, *ser* article NP, *estar* P NP. For the time being, I will let Spanish aside for future research.

- (24) Bhí Seán ina dhochtúir tráth.
 COP-BI.PAST Sean in.AGR doctor once
 “Sean was a doctor once.” (Doherty 1996)

Finally, the eventive predicate, i.e. PP[*in*-NP], constitutes the only option in small clause predicate. As discussed in the literature (see Chung & McCloskey 1987, for instance) the predicate position of a small clause cannot be filled by an NP in Irish. Irish requires instead that the nominal be introduced by the preposition *in* “in”, exactly as in the *bí*-construction. Consider, for instance, the verb *happen* that selects a small clause complement (25) and a perception verb (26):

- (25) a. *Tharlaigh iad dlíodóiri.
 happened.PAST them lawyers
 intended: “It happened that they were lawyers.”
 b. Tharlaigh na dhlíodóir é.
 happened.PAST in.AGR lawyer him
 “He happened to be a lawyer.” (Chung & McCloskey 1987)
- (26) Chonaic mé Ciarán *(ina) léachtóir.
 see.PAST I Ciaran in.AGR lecturer
 “I saw Ciaran as a lecturer.” (Doherty 1996)

The PP[*in*-NP] predicate that is licensed in eventive constructions is also licensed in small clauses. Similarly to what we have seen in French, only eventive predicate nominals can constitute appropriate small clause predicates in Irish.

To conclude, the PP[*in*-NP] predicate that occurs in eventive copular sentences in Irish shares with the variant without article in French and German not only the same interpretation (temporary, activity-like denoting expression), but also use (predicational, as opposed to identificational) and syntactic environment (as small clause predicate, with temporal modifiers):

- (27)

	variant without article (French; German)	PP[<i>in</i> -NP] (Irish)
a) Identificational reading	no	no
b) Lifetime-effects	no	no
c) Spatio-temporal modification	yes	yes
d) Activity reading	yes	yes
e) Small clause predicate	yes	yes

4. Eventive copular sentences

4.1. A syntactic distinction

The distinction between permanent vs. temporary properties is often formalized as the stage-level vs. individual-level distinction. This distinction is often believed to be a lexical distinction between predicates that express permanent properties (*intell-*

igent, doctor) and predicates that express temporary properties (*sick, absent*) (see, for instance Carlson 1977, Kratzer 1995). However, this view poses a number of problems.⁸ First, it is always possible to coerce a permanent property into a temporary one (28) (Higginbotham 1985):

(28) John went to college dumb and left it intelligent.

Second, in both French and Irish any predicate nominal can enter in either construction, independently of whether it denotes a property perceived as permanent or not. Predicates that express permanent properties can appear with or without the indefinite article (as in (29), for instance, where the properties of being ‘a resistance hero’ and ‘a tyrant’ presumably hold permanently of the individual denoted by the subject, but can still occur without the article):

- (29) a. Il n’a jamais connu son père qui était (un) héros de la Résistance.
 he has never known his father who was (a) hero of the resistance
 “He has never met his father who was a resistance hero.”
 b. L’enfant peut être tyran.
 the child can be tyrant
 “A child can be a tyrant.”

Similarly, a property like being “a hitchhiker”, which does not hold permanently of an individual can be constructed with or without an article:⁹

- (30) Je suis (une) auto-stoppeuse, mais seulement parce que je n’ai pas le choix.
 I am (a) hitchhiker, but only because I have not the choice
 ‘I am a hitchhiker, but only because I don’t have any choice.’

Similarly, in Irish typical permanent properties like ‘to be a man’ can appear in both the *is* and *bí* constructions, showing that the distinction between stage-level and individual-level properties cannot be a lexical one:

- (31) a. Is fear é.
 COP-IS man he
 ‘He is a man.’
 b. Tá sé ina fhear (anois).
 COP-BI he in.AGR man (now)
 ‘He is a man (now).’ (Stenson 1981)

⁸ For further arguments against a lexicalist approach see Higginbotham & Ramchand (1997).

⁹ It is commonly accepted that the variant without article exists only with profession and nationality denoting nominals (which include real professions like *dentiste* “dentist” as well as functions and titles like *ministre* “minister”, *président* “president”, *roi* “king”, and so on). However, this generalization is certainly too strong as we find a considerable number of bare predicate nominals (see (ii)), which cannot be considered as professions as such, but do involve some sort of underlying “activity” in a broader sense:

- (ii) Paul est auto-stoppeur / fumeur / locataire / prisonnier.
 Paul is hitchhiker / smoker / tenant / prisoner
 ‘Paul is a hitchhiker / a smoker / a tenant / a prisoner.’

In German, however, the variant without article seems to be more restricted than in French.

I assume instead that the distinction between stage-level and individual-level predicates is syntactic and relates to the structure of copular sentences involved (following Kupferman 1991, Ramchand 1996, and Higginbotham and Ramchand 1997). In particular, I assume that stage-level predication involves predication over events (32): the property of being an event of a certain kind is predicated of a situation *s*; while individual-level predication involves predication over individuals (33): a particular property is predicated of an individual *x* (following Higginbotham & Ramchand 1997):

- (32) Stage-level:
 $\exists s [\lambda e [\dots] (s)]$
 there exists a situation *s*, such that *s* has the property of being an event of a particular kind
- (33) Individual-level:
 $\exists x [\lambda x [\dots] (x)]$
 there exists an individual *x*, such that *x* has a particular property

Accordingly, on the one hand, stage-level properties, i.e. spatio-temporally dependent properties, correspond to events and are constructed in *eventive* predications. On the other hand, individual-level properties are fundamentally properties predicated of an individual and are constructed in *non-eventive* predications.

4.2. Predicate nominals

It is usually accepted that nominals are individual-level predicates only, while adjectives can be either stage-level or individual-level predicates (cf. Milsark 1974, among others). The traditional diagnostics for the distinction between stage-level and individual-level predicates in English show that nominals are never stage-level: they are excluded from existential constructions (34a), they cannot appear with perception verbs (34b), and they allow for a generic reading of bare plurals (34c):

- (34) a. *There were people doctors b. *John saw Paul a doctor
 c. Dogs are mammals (generic; *existential)

However, the variant without the article in French exhibits the characteristics of stage-level predicates: as we have seen in section 2, it accepts spatio-temporal modifications, it does not give rise to lifetime-effects, and so on. Moreover, it passes the tests for stage-level predicates illustrated above with English. Unlike predicate nominals in English (34a-b),¹⁰ bare predicate nominals in French can appear in existential constructions (35),¹¹ and can appear in the small clause complement of a perception verb (36):

- (35) Il y a des hommes (mauvais) chirurgiens.
 there is indef.pl men (bad) surgeon
 ‘There are men (that are) (bad) surgeons.’

¹⁰ French does not have bare plurals; I leave aside the test (c).

¹¹ The possibility of inserting an adjective shows that *homme chirurgien* ‘man surgeon’ cannot be a compound noun.

- (36) J'ai vu Paul enfant une seule fois; maintenant il est grand.
 I have seen Paul child one only time now he is big
'I've seen Paul as a child only once; he is now grown-up.'

These properties cannot be explained if the distinction between stage-level and individual-level predicates is a lexical one. Assuming that the distinction is syntactic, however, gives us some insight to why predicate nominals, which are traditionally assumed to be individual-level predicates only, can exhibit properties of stage-level predicates in French and Irish.

4.3. Aspect

Furthermore, assuming that bare predicates in French and German are constructed in eventive copular sentences gives us some insight to why they are compatible with aspect, whereas the variant with the article becomes ungrammatical when aspect is realized. As aspect modifies the nature of the event, expressed aspect is only compatible with eventive copular sentences, and hence, with the bare variant.

Two pieces of data support this generalization. First, the article is not allowed when perfectivity is marked, by means of the auxiliary (37a).¹² The imperfective aspect, however, is a default aspect, and can take both variants (37b):

- (37) a. Paul a été (*un) médecin. b. Paul était (un) médecin.
 Paul has been a doctor Paul was a doctor
"Paul has been a doctor." *"Paul was a doctor."*

Second, aspectual copulas like *devenir* "to become", as opposed to the copula *être* "to be", can also only take bare predicates. *To become* entails a change of state that is incompatible with the non-eventive reading associated with the variant with the article:

- (38) Paul deviendra (??un) peintre.
 Paul become.FUT a painter
"Paul will become a painter."

Similarly, the variant with the article cannot appear with perfective verbs such as *to die* either. In both (39) and (40) the state of 'being a N' is bounded, and is only compatible with an eventive predicate:

- (39) Paul mourra (*un) médecin.
 Paul die.FUT a doctor
"Paul will die a doctor."
- (40) Paul sortira de son école (*un) architecte.
 Paul come out. FUT of his school an architect
"Paul will graduate as an architect."

¹² Similar facts have been noted, independently, by Kupferman (1979).

Similar contrasts exist in German. Although the contrast is less strong than in French for my informants, it is, nevertheless, attested:

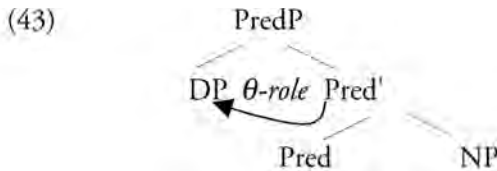
- (41) Paul war (ein) Schauspieler.
 Paul was an actor
"Paul was an actor."
- (42) Paul ist (??ein) Schauspieler gewesen.
 Paul is an actor been
"Paul has been an actor."

To conclude, I assume that the distinction between stage-level and individual-level predicates is a syntactic one and relates to the structure of copular sentences. Stage-level predicates are constructed in eventive copular sentences and involve predications over situations. Predicate nominals can be eventive, and therefore can exhibit the traditional properties of stage-level predicates. Having defined the nature of eventive predicate nominals, I will turn in section 5 to a unified analysis of eventive predicate nominals.

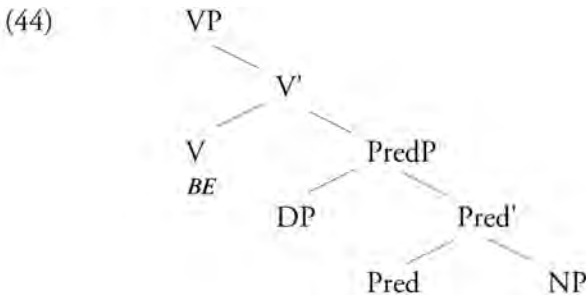
5. A unified account for eventive predicate nominals

5.1. Predicational functional head

I assume that the head of the small clause is a functional projection in a hierarchical structure as exemplified in (43). Following Bowers 1993, for instance, I will take this functional head to be the predicational head Pred. The DP subject is licensed in spec-Pred where it receives an external theta-role from Pred':



I take the verb 'to be' to be a raising verb, as generally assumed, that takes a small clause of the type illustrated above as complement. The structure of eventive copular sentences is as follows:



I further assume that the head Pred must be realized; and I argue in the rest of this section that there are two ways Pred can be realized: either by a preposition (*P-licensing*), or by the predicative head N through head movement (*N-licensing*).

5.2. P-licensing

5.2.1. Overt P

The implementation of the analysis for Irish is quite straightforward: the head Pred is realized by the preposition *in* “in” in Irish.¹³ The predicate nominal *shagart* “priest” in (45a) is theta-marked by the preposition.¹⁴ Accordingly, sentence (45a) has the structure (45b):

- (45) a. Tá Sean ina shagart.
 COP-BI Sean in.AGR priest
 ‘Sean is a priest.’
 b. [VP [V BI [PredP [DP Sean] [Pred in [NP shagart]]]]]

The NP *shagart* “priest” is generated in the complement position of *in* “in”, the head Pred. The sequence *in*-NP is a small clause predicate, and the DP subject is generated in the subject position of the small clause, i.e., in spec-PredP. The fact that PP[*in*-NP] is compatible with small clause predicate positions follows from the fact that PP is indeed a small clause itself, and thus can be selected not only by the copula *bí*, but also by any raising verb or verb that usually selects a small clause.

¹³ Alternatively, Adger & Ramchand (2003) proposed that the preposition *in* *bí*-constructions in Scottish Gaelic (a language closely related to Irish) is needed in order to introduce an event variable that nouns are lacking. However, two arguments based on French support the view that nouns, like adjectives and verbs, are predicates of events also. First, the argument based on the logic of VP modifiers, proposed by Davidson (1967) in favor of the existence of an underlying event for verbs, can be reproduced with bare predicate nominals in French. The fact that the entailment in (iii) is invalid, i.e. the conjunction of ‘being a dentist’ and ‘being in Paris’ does not entail ‘being a dentist in Paris’, indicates that the two predicates are potentially two independent events. Consequently, nominals also must be seen as relative to events (iiib):

- (iii) a. Paul est dentiste. (Paul is dentist)
 Paul est à Paris. (Paul is in Paris)
 ∴ Paul est dentiste à Paris. (Paul is a dentist (=practices dentistry) in Paris)
 b. $\exists e$ [dentist(e, paul) & in(e, paris)]

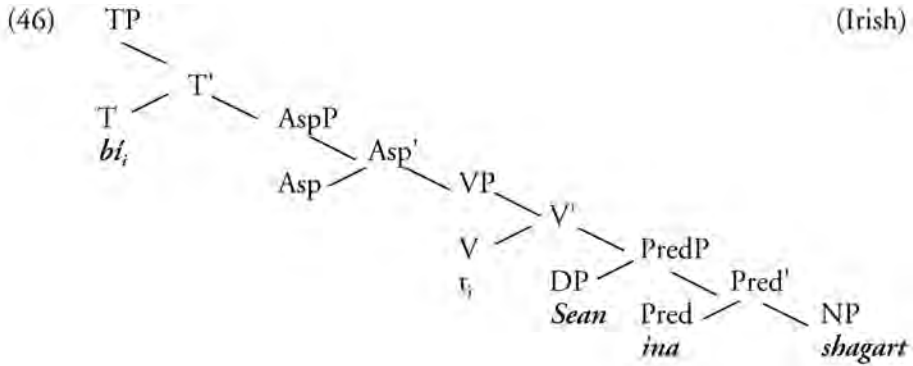
Second, if nominals were not predicated of an event, we would expect them to be incompatible with adjectives like *future* or *former*, known to be predicates of events only (cf. Larson 1995, 1998, among others). This prediction is not borne out by the data:

- (iv) John is a former lawyer.
 a. $\sqrt{\exists e}$ [lawyer(j,e) & former(e)]
 b. $\# \exists e$ [lawyer(j) & former(j)]

¹⁴ Note the role of prepositions in licensing of external arguments for eventive nominals in English predicates such as (v):

- (v) a. Paul is *(in) transit.
 c. The house is *(on) fire.
 b. The partners are *(at) war with each other.
 d. The road is *(under) construction.

Irish is a VSO language, where V raises to T and carries tense information, while the DP subject remains in the VP-internal position (cf. Chung and McCloskey 1987). The copula *bí* is generated in V and moves to T:



The event variable in the *bí*-construction is introduced by the nominal and is bound, I assume, by Aspect, which introduces existential closure over the event. As an aspect projection is obligatorily realized in eventive sentences, the compatibility of the *bí*-construction with aspectual distinctions follows. I assume that T, however, is a predicate of event.

5.2.2. Covert P

The account proposed above and carried out for Irish extends straightforwardly to Instrumental-marked predicates in Russian under the assumption that Instrumental is assigned by an empty preposition, and that Pred is realized by a covert P.

Russian predicate nominals in copular sentences can bear either Instrumental or Nominative case (47):¹⁵

- | | |
|--|--|
| <p>(47) Saša byl muzykantom.
 Sasha was musician.INST
 “Sasha was a musician.”</p> | <p>(48) Saša byl muzykant.
 Sasha was musician.NOM
 “Sasha was a musician.”
 (Bailyn & Rubin 1991)</p> |
|--|--|

Instrumental-marked predicates express properties perceived as transient. Sentence (47) receives a temporary interpretation: “Sasha was a musician temporarily, at some point” (cf. for instance Bailyn & Rubin 1991, Matushansky 2000, Filip 2001, among others) or has been claimed to involve a “change of state” (Filip 2001) (49):

¹⁵ In the past tense only (cf. for instance, Matushansky 2000 and Filip 2001): Nominative is the only option in the present tense and Instrumental is highly preferred or the only option in the future tense. The reason why Instrumental is not allowed in the present (where the copula is null) remains an issue.

- (49) On byl učitelem (potom) fotografom.
 he was teacher.INST (then) photographer.INST
"First he was a teacher, then he became a photographer."

(Geist 1999, cited by Filip 2001)

By contrast, Nominative-marked predicates (48) express properties that are seen as inherent or 'defining' (Matushansky 2002).

Assuming the distinction between eventive and non-eventive copular sentences, Instrumental case marked predicates appear in eventive sentences only, while Nominative marked predicates are restricted to non-eventive predications. Instrumental marked predicate nominals are the Russian counterpart of the variant without the article in French and German, and the PP[*in*-NP] in Irish. First, Instrumental marked predicates cannot appear in identificational sentences (50), as well as true equatives (51), which both require Nominative case:

- (50) Ivanuška-duračok byl tot brat / *tem bratom kotoryj
 Ivanushka-fool was that brother.NOM / that brother.INST which
 vseгда popadal v bedu.
 always got into trouble.
"Ivanushka the Fool was that brother that always got into trouble."

(Pereltsvaig 2001)

- (51) Mark Tven byl Samuèl Klements.
 Mark Twain.NOM was Samuel Clements.NOM
"Mark Twain was Samuel Clements."

(Matushansky 2000)

Second, Instrumental-marked predicates do not give rise to lifetime effects: in (52b) the individual Sasha is now dead, whereas in (52a), he is simply not a musician anymore (Matushansky 2000):

- (52) a. Saša byl muzykantom. (absence of lifetime effects)
 Sasha was musician.INST
"Sasha was a musician."
 b. Saša byl muzykant. (lifetime effects)
 Sasha was musician.NOM
"Sasha was a musician."

Third, as it was the case previously for French and Irish, the predicate position of small clauses can only be filled by the eventive predicate nominal, i.e. the predicate can only be marked by Instrumental case (53):

- (53) Sovremenniki sčitali Puškina *veliki poèt / velikim poètom.
 contemporaries considered Pushkin.ACC great poet.NOM / great poet.INST
"Contemporaries considered Pushkin (to be) a great poet."

(Matushansky 2000)

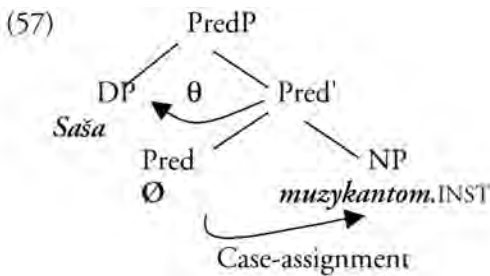
Similarly, Instrumental case is the only option in the predicate position of a small clause selected by a raising verb like *become* (54) and *turn out* (55) (cf. Bailyn & Rubin 1991):

- (54) Saša stal vračom. (55) Saša okazalsja durakom.
 Sasha became doctor.INST Sasha turned out fool.INST
 “Sasha became a doctor.” “Sasha turned out to be a fool.”

Finally, sentences (54) and (55) constitute not only evidence that Instrumental marked predicates can be small clause predicates, but also that they are compatible with aspectual distinctions. The sensitivity of case marking to aspect has been discussed by Matushansky (2000), who shows that only Instrumental marked predicates are compatible with expressed aspect. Aspect in Russian is obtained by morphological affixation (prefix or suffix) to a simple verb that is usually taken to be imperfective by default. With aspectually specified forms of the copula (56a-b), Nominative becomes impossible. This is not true of the default copula (56c):

- (56) a. Ja pobyla zavedujuščej / * zavedujuščaja dva časa.
 I was. PERF manager.INST / manager.NOM two hours
 “I have been a manager for two hours.”
 b. Ja byvala zavedujuščej / * zavedujuščaja no redko.
 I was. IMPERF manager.INST / manager.NOM but rarely
 “I have been a manager, but rarely.”
 c. Ja byla zavedujuščej / zavedujuščaja dva časa | no redko.
 I was manager.INST / manager.NOM two hours | but rarely
 “I have been a manager, for two hours | but rarely.”

The account for the Russian data can be unified with the account of Irish when we assume that Pred is realized by an empty preposition in Russian, which is responsible for assigning Instrumental case to the predicate. The covert preposition \emptyset in Russian (57) is the same functional head Pred as the overt preposition ‘in’ in Irish:

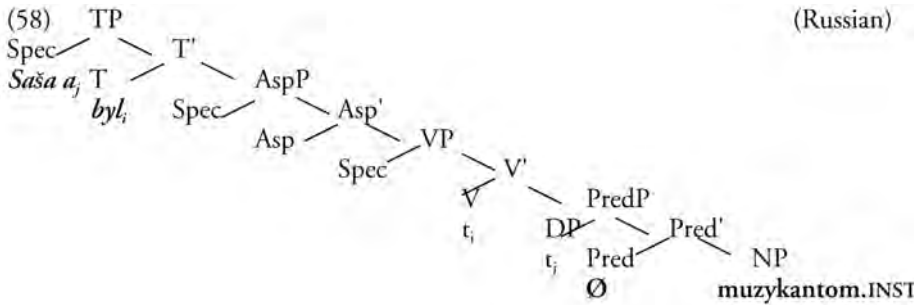


A structure similar to (57) has been argued for, independently, by Bailyn & Rubin (1991) for all Instrumental marked predicates in Russian. The only difference is that I take Pred to be a covert preposition in Russian. The reason why Russian marks eventive predicate nominals with Instrumental case, instead of a structural case, namely Accusative (*vs.* Nominative in non-eventive constructions) has often been an unanswered question. Under the view developed here, it is so because case is assigned by a preposition.

The apparent difference between case-marking in Russian and PP in Irish is, indeed, only a superficial one, linked to the fact that the former is a case language,

while the latter is not, a common variation across languages. In both cases, Pred is realized by a preposition.

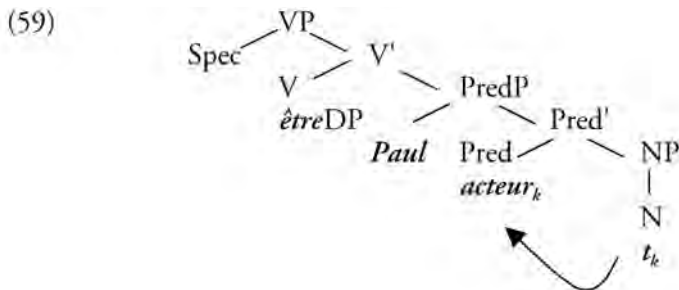
The structure of eventive nominal predication in Russian is as follows (58): the copula is generated in V and moves to T; the DP subject moves to spec-TP to get Nominative case:



5.3. N-licensing

5.3.1. Head movement

Given our premises and the unified account proposed for Irish and Russian, the fact that French and German do not exhibit an overt preposition with predicate nominals is rather unexpected. Since there is, by assumption, a functional head Pred in these languages that licenses the external argument of a predicate nominal, and since Pred needs to be realized, either it is realized as an empty preposition, as in Russian, or it is realized by some other element, functional or not. I will argue in this sub-section that the correct analysis is that Pred is realized by the predicative head N through head movement from its original position, as illustrated in (59):



This view is supported by two pieces of data: the lack of agreement on the predicative heads, and the impossibility for the predicative head to be modified by an adjective or a relative clause.

First, Noun Incorporation is subject to the Head movement Constraint (cf. Travis 1984), which prevents a head to move to another head position if there is an

intervening head. Here, N can move to Pred only if NP is the complement of Pred. More specifically, an intervening head like Number, above NP, would block incorporation. Accordingly, I predict that incorporated predicate nominals in French and German cannot bear agreement. This prediction is borne out by the data. Number agreement is absent on bare predicates (i.e. on eventive predicate nominals) in French and German.¹⁶ The lack of number agreement is illustrated with the predicate nominal *général* “general” in (60a), and the compound *avocat international* “international lawyer” in (60b):¹⁷

- (60) a. Paul et Simon sont *?généraux / général (des armées).
 Paul and Simon are general.PL / general.SG of-the armies
 “Paul and Simon are generals (of the army).”
 b. Dupont et Dupré sont *avocats internationaux / avocat international.
 Dupont & Dupré are*lawyer.PLinternational.PL /lawyer.SGinternational.SG
 “Dupont and Dupré are international lawyers.”

In German, both a bare singular and a bare plural are possible (see (61a) and (61b), respectively). However, (61b) cannot be an eventive copular sentence, but rather is a non-eventive one. In other words, the bare plural predicate *Generäle* “generals” in (61b) is the plural of the variant with article (i.e. *ein General* “a general”), and not of the bare singular (*General* “general”):

- (61) a. Paul und Robert sind General. b. Paul und Robert sind Generäle.
 Paul and Robert are general.SG Paul and Robert are general.PL
 “Paul and Robert are generals.” “Paul and Robert are generals.”

The claim that (61b) cannot be an eventive predication is supported by two pieces of evidence. First, it can only be interpreted as identificational, i.e. it constitutes an appropriate answer to the question *Who are Paul and Robert?*, and not to *What do they do?* Recall that eventive sentences cannot have an identificational reading. Second, the plural is the dispreferred option with an aspectual copula like *become*. Recall, again, that aspectual copulas take eventive predicates only:

- (62) Paul und Robert wollen Arzt / ??Ärzte werden.
 Paul and Robert want doctor.SG / doctor.PL become
 “Paul and Robert want to become doctors.”

Second, assuming that adjectives attach higher than NP (i.e. at some functional layer between NP and DP), the impossibility to modify a bare predicate by an adjective and a relative clause in French and German further supports the view that

¹⁶ This generalization seems to be subject to dialectal variations, as certain speakers of French and German allow plural agreement as well.

¹⁷ Only a very limited set of nouns exhibit a clear plural form in French. In fact, although plural is always marked in the written system, it is rarely pronounced (except for irregular plurals and in the context of a liaison, for instance). In both examples in (60) *general* and *international* are unambiguously singular. In (60b) I take the fact that the adjective *international* is singular as an indication that the whole predicate is singular.

they are bare NPs. The insertion of an adjective modifying the bare predicate nominal obligatorily triggers the appearance of the article (63-64). Note that the article is required with both pre and post-nominal adjectives in French:¹⁸

- (63) a. Paul est *(un) riche / surprenant / agréable / beau (...) médecin.
 Paul is a rich / surprising / agreeable / beautiful doctor
"Paul is a rich / surprising / agreeable / beautiful doctor."
 b. Paul est *(un) médecin âgé / intelligent / réputé ...
 Paul is a doctor aged / intelligent / reputable
"Paul is a(n) old / intelligent / well-known doctor."
- (64) Paul ist *(ein) reicher / erstaunlicher / anerkannter / angenehmer Arzt.
 Paul is a rich / surprising / well-known / agreeable doctor
"Paul is a rich / surprising / well-known / agreeable doctor."

Examples (65-66) illustrate the same fact with relative clauses:

- (65) a. Paul est *(un) médecin que tout le monde aime.
 Paul is a doctor that everybody loves
"Paul is a doctor that everybody likes."
 b. Paul est *(un) médecin qui travaille trop.
 Paul is a doctor who works too much
"Paul is a doctor who works too much."
- (66) Paul is *(ein) Arzt den alle mögen.
 Paul is a doctor that all like
"Paul is a doctor that everybody likes."

Contrary to French/German, in Irish and Russian, number agreement must be realized: the NP inside the PP selected by *bí* must bear agreement in Irish (67), and Instrumental marked predicates must bear agreement in Russian (68):

- (67) a. Tá Seán agus Máire 'na ndochtúirí.
 COP-BI Sean and Máire in.AGR doctor.PL
"Sean and Maire are doctors."
 b. Tá siad 'na gcáirde maithe ag a chéile.
 COP-BI they in.AGR friend.PL good.PL at each-other
"They are good friends to one another."
- (68) a. Saša i Miša byli muzykantami / *muzykantom.
 Sasha and Misha were musician.INST.PL / musician.INST.SG
"Sasha and Misha were musicians."
 b. Kogda oni byli studentami, ...
 when they were students. INST.PL
"When they were students, ..."

¹⁸ Both in French and in German there is a very restricted set of exceptions to this generalization. In French, for instance, the few adjectives allowed with the bare variant are: *professionnel* "professional", *amateur* "amateur", *bon* "good", *excellent* "excellent", *mauvais* "bad", certain nationality adjectives as *grec* "Greek" (under the appropriate reading).

Agreement on the predicate in Irish and Russian suggests that NP is dominated by a NumP, site of the realization of agreement. In other words, with P-licensing the predicate can involve an additional layer of functional projection above NP, while with N-licensing, incorporation blocks any intervening head above NP. Furthermore, Irish and Russian are not subject to any constraint with respect to the modification of the NP inside the PP predicate (69) and the instrumental marked predicate (70), respectively. This again supports the idea that predicates in Irish and Russian involve an additional functional layer, to which by assumption, adjectives and relative clauses are adjoined:

- (69) Tá sé ina fhear láidir (anois). (Irish)
 COP-TA he in.AGR man strong (now)
 “*He is a strong man (now).*” (The Christians Brothers 1999)
- (70) a. Makarenko byl xorošim učitelem. (Russian)
 Makarenko was good.INST teacher.INST
 “*Makarenko was a good teacher.*”
 b. Zoluška byla bednoj krest’jankoj.
 Cinderella was poor.INST peasant.INST
 “*Cinderella was a poor peasant.*” (Matushansky 2000)

5.3.2. Against an empty P

The main argument against an empty P in French and German comes from the fact that eventive predicate nominals do not bear a case that can possibly be assigned by a preposition in German. Prepositions in German can assign Accusative, Dative or Genitive case. However, bare predicate nominals can only bear Nominative, a default case. The difficulty to see what case a bare nominal bears in German is that case appears on adjectives and articles only. Hence, it is not expected to be visible at all on a bare singular as in *Paul ist Skifahrer* “Paul is a skier”. Nevertheless, the few adjectives that can modify the bare predicate are, under the traditional view, Nominative-marked. This suggests that bare predicates bear Nominative (71):¹⁹

- (71) Paul ist professioneller Skifahrer.
 Paul is professional.NOM skier.NOM
 “*Paul is a professional skier.*”

Since Nominative cannot be plausibly assigned by a preposition, there is no empty P in German, and by extension, neither in French.

5.4. English

5.4.1. Covert P

We now have three tests that help distinguishing the two strategies of licensing of the head Pred from one another: number agreement, the possible modification of the head N by an adjective and case marking.

¹⁹ The variant with the article is also marked with Nominative case. Hence, contrary to Russian, case does not discriminate between the two types of predicates.

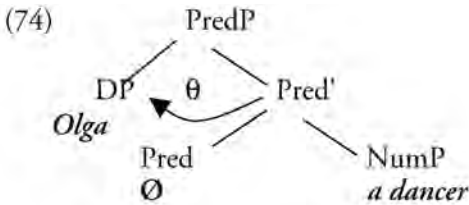
These tests applied to English (the case test is not relevant, as English is not a case language), produce the surprising result that English is an Irish/Russian like type of language, and not a French/German type of language as one would a priori expect. First, English exhibits obligatorily number agreement on predicate nominals: a plural subject triggers a plural agreement on the predicate:

(72) John and Paul are doctors / *doctor.

Second, a predicate nominal modified by an adjective can receive an eventive interpretation. Consider the famous example (73). Sentence (73) is ambiguous between two readings: it can be interpreted either as “Olga is a dancer and a beautiful person” (intersective reading) or as “Olga dances beautifully” (non-intersective reading) (Larson 1995). The second reading corresponds to the eventive predication, where *Olga is a dancer* is interpreted as *Olga dances*:

(73) Olga is a beautiful dancer.

Accordingly English patterns with Irish/Russian, and licenses its head Pred through the P-licensing strategy. Hence, contrary to French/German, we can assume that Pred is realized as an empty preposition in English. This claim is compatible with Bowers (1993), who assumes that Pred is a null head in English:



5.4.2. Obligatory indefinite article

Although the similarity between English and Irish/Russian may seem surprising at first sight, it presents some interesting results, especially in terms of the contrasts between English and French. First, it provides some understanding to why English does not have a variant without article and the article remains obligatory in eventive predication. The obligatory presence of the article *a/an* in eventive predication in English constitutes further evidence for the additional functional layer NumP above NP, assuming that the indefinite article is the spell-out of Num:

(75) Olga is *(a) dancer.

The indefinite article is obligatory in non-eventive sentences, as it is a mark of agreement. The article *a/an* is ambiguous between the indefinite article (in non-eventive sentences) and a mark of agreement (in eventive sentences).

5.4.3. ‘Olga is a beautiful dancer’

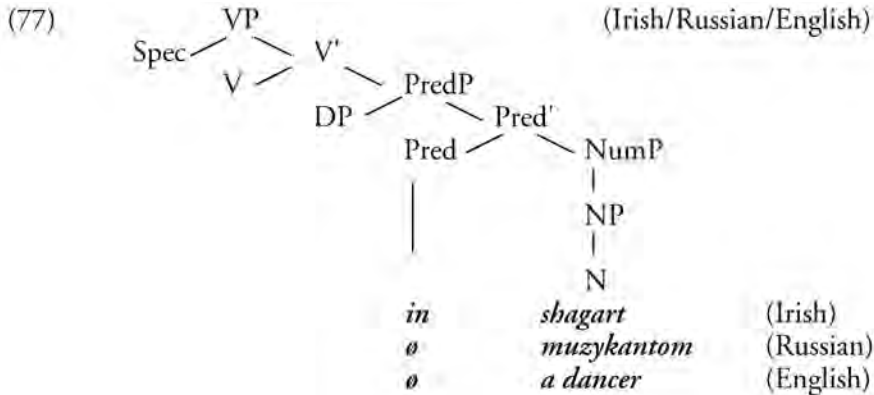
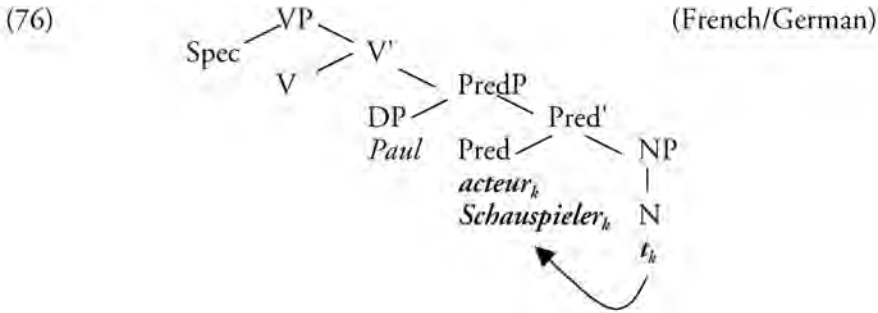
Second, it provides some insight regarding the fact that the sequence *beautiful dancer* can be interpreted as eventive in English but not in French. A sentence like *Olga est une*

belle danseuse ‘Olga is a beautiful dancer’ in French has only one reading, namely ‘Olga is a dancer and a beautiful person’, and lacks the eventive reading ‘Olga dances beautifully’.

Because the article is ambiguous between the indefinite article and the mark of agreement, and because English has only one copula, a sentence like *Olga is a dancer* is ambiguous between an eventive predication and a non-eventive predication. The ambiguity gives rise to the two interpretations mentioned above. This is clearly not the case in French, where the variant with article can only receive a non-eventive reading. The cross-linguistic variation is predicted under my account, as *a beautiful dancer* can be a property of events in English but not in French.

6. Conclusion

To conclude, the only difference between French/German on the one hand, and Irish/Russian/English on the other, is the nature of the element that realizes the functional head Pred. The former group of languages licenses the head Pred through N-licensing (76), and the latter through P-licensing (77). This difference aside, the structure of eventive predication is the same in all five languages:



This account provides a unified analysis to the fact that eventive predicates are introduced by an overt preposition in Irish, are assigned Instrumental case in Russian, and must be bare in French and German. The three phenomena are reducible

to whether Pred dominates a P or an N. This difference aside, the structure of eventive predications is identical: eventive copular sentences are constructed with an eventive copula that selects a small clause headed by a predicational head. A similar small clause structure has been, independently, argued for by Baylin & Rubin (1991) and Bowers (1993). As I have shown this analysis is easily extendable to English, and produces interesting results in terms of the variation between English and French.

References

- Adger, D. and G. Ramchand, 2003, "Predication and Equation", *LI* 34(3): 325-60.
- Bailyn, J. F. and E. J. Rubin, 1991, "The Unification of Instrumental Case Assignment in Russian". In Taborio, A. and W. Harbert (eds.), *Cornell Working Papers in Linguistics* 9: 99-126.
- Bowers, J., 1993, "The Syntax of Predication", *LI* 24(4): 591-656.
- Carlson, G., 1977, *Reference to Kind in English*. Garland Publishers, New York.
- Carnie, A., 1995, *Non-Verbal Predication and Head Movement*. Ph.D. dissertation, MIT, Cambridge.
- Christian B., 1962, 1999, *New Irish Grammar*. Fallons, Dublin.
- Chung, S. and J. McCloskey, 1987, "Government, Barriers, and Small Clauses in Modern Irish", *LI* 18(2): 173-238.
- Davidson, D., 1967, "The Logical Form of Action Sentences". In N. Rescher (ed.), *The Logic of Decision and Action*. University of Pittsburgh Press, Pittsburg.
- Doherty, C., 1996, "Clausal Structure and Modern Irish Copula", *NLLT* 14: 1-46.
- Fernald, Th. B., 1994, *On the Non-uniformity of the Individual and Stage-level Effects*. Ph.D. dissertation. UCSC, Santa Cruz.
- Filip, H. 2001, "The Semantics of Case in Russian Secondary Predication". In *Proceedings from SALT 11*: 192-211.
- Geist, L., 1999, "Russisch *by* als funktionale und/oder lexikalische Kategorie", *ZAS Papers in Linguistics* 14: 1-39.
- Hale, K. and J. McCloskey, 1984, "On the Syntax of Person-Number Inflection in Modern Irish", *NLLT* 1: 487-533.
- Higginbotham, J., 1985, "On Semantics", *LI* 16: 547-93.
- , and G. Ramchand, 1997, 'The Stage-Level/Individual-Level Distinction and The Mapping Hypothesis'. In *Oxford Working Papers in Linguistics, Philology and Phonetics* 2: 53-83.
- Higgins, F. R., 1979, *The Pseudo-cleft Construction in English*. Garland Publishers, N.Y.
- Kratzer, A., 1995, "Stage-Level and Individual-Level Predicates". In G. Carlson and F. J. Pelletier (eds.), *The Generic Book*. The University of Chicago Press, Chicago: pp. 125-75.
- Kupferman, L., 1979, "Les constructions *Il est un médecin / C'est un médecin*: essai de solution", *Cahiers de linguistique* 9: 131-164, Presse de l'Université du Québec, Montréal.
- , 1991, "Structure événementielle de l'alternance *un/ø* devant les noms humains attributs", *Langage* 102: 52-75.
- Larson, R., 1995, "Olga is a beautiful dancer". Talk presented at LSA, New Orleans.

- Matushansky, O., 2000, "The Instrument of Inversion: Instrumental Case in the Russian Copula". In Billerey and Lillehaugen (eds.), *WCCFL 19 Proceedings*, Cascadilla Press. pp. 288-301.
- , 2002, "Encasing the Time: Temporal Effects on Russian Predicative Case". Talk presented at NELS 23, MIT.
- Milsark, G., 1974, *Existential sentences in English*. Ph.D. dissertation, MIT, Cambridge.
- Musan, R., 1995, *On the Temporal Interpretation of Noun Phrases*. Ph.D. dissertation, MIT, Cambridge.
- Ó Máille, T., 1912, "Contributions to the History of the Verbs of existence in Irish", *Ériu* 4(1): 1-102.
- Ó Siadhail, M., 1989, *Modern Irish: Grammatical Structure and Dialectal Variation*. Cambridge Studies in Linguistics, Cambridge U.P., Cambridge.
- Pereltsvaig, A., 2001, "On Thematic Relations in Copular sentences". In M. Yoo and J. Steele (eds.), *McGill Working Papers in Linguistics* 15(2).
- Pollock, J.-Y., 1983, "Sur quelques propriétés des phrases copulatives en français", *Langue Française* 58: 89-125.
- Ramchand, G. C., 1996, "Two Subject Positions in Scottish Gaelic: the Syntax-Semantics Interface", *Natural Language Semantics* 4(2): 165-91.
- Rouveret, A., 1998, "Points de vue sur le verbe 'être'". In A. Rouveret (ed.), "*Etre*" et "*Avoir*": *syntaxe, sémantique, typologie*. Presses Universitaires de Vincennes, Saint-Denis. pp. 11-65.
- Roy, I., 2001, "Predicate Nominals in French". Ms. University of Southern California.
- Stenson, N., 1981, *Studies in Irish Syntax*. *Ars Linguistica* 8. Gunter Narr Verlag, Tübingen.
- Travis, L., 1984, *Parameters and effects of word order variation*. Ph.D. dissertation, MIT, Cambridge.

RESUMPTIVE PRONOUNS AND MATCHING EFFECTS IN ZURICH GERMAN RELATIVE CLAUSES AS DISTRIBUTED DELETION

Martin Salzmann

Leiden Center for Linguistics (ULCL)

Abstract

*Zurich German (ZG) relative clauses are remarkable from a Germanic point of view in that resumptive pronouns are employed instead of relative pronouns. Reconstruction effects and Strong Crossover violations show that movement is involved in the derivation of ZG relative clauses. Matching effects sensitive to case and preposition provide crucial evidence that the distribution of resumptives is determined by general licensing conditions on oblique case and prepositions. The matching/non-matching dichotomy is modeled as an instance of Distributed Deletion, which is claimed to be independently available in the language. Matching is furthermore sensitive to the actual surface form and thus favors a late insertion approach to morphology.**

1. Introduction

This paper investigates the grammar of resumptive pronouns in Zurich German (ZG)¹ relative clauses. In section 2, I will lay out the general properties of ZG relative clauses, including the distribution of resumptives. In section 3, I discuss data

* I would like to thank the audiences at the WSM in Leiden (spring 2004), the GGS 2004 in Mannheim, the TABU-dag 2004 in Groningen, and of course BIDE 2004 where earlier versions of this research were presented, as well as Rajesh Bhatt, Lisa Cheng, Aniko Liptak, Johan Rooryck, Luis Vicente, and Henrietta Yang for invaluable comments at various stages in the development of these ideas. The most important contributors, however, were my informants without whom this research would have been impossible. I am very grateful to Barbara Bächli, Silvio Bär, Kathrin Büchler, Martin Businger, Silvia Coretti, Peter Gallmann, Martin Graf, Andreas Henrici, Maja Herrmann, Beatrice Hartmann, Roland Litschor, Michael Mente, Franziska Näf, Christian Rapold, Didier Ruedin, Etienne Ruedin, Charlotte Schweri, Guido Seiler, Roman Sigg, Benjamin Stückelberger, Kathrin Würth, Tobias Zimmermann, Regula Zimmermann-Etter, Hans-Jürg Zollinger, and Serena Zweimüller.

¹ Zurich German is the dialect spoken in most parts of the canton (state) of Zurich in Switzerland. There are approximately one million native speakers.

that argue in favor of a movement analysis of resumptives. Section 4 presents hitherto undiscovered matching effects. In section 5 I show that resumptives receive the same interpretation as gaps. Section 6 reviews previous approaches to resumption and shows that they all fail to explain the ZG data. In section 7 I present a new account, and section 8 contains detailed derivations. Section 9 summarizes the paper.

2. General properties of Zurich German relative clauses

2.1. General form

Restrictive Relative Clauses² in ZG are introduced by an invariant complementizer *wo* (*won* before vowels). Relative pronouns are absent, except for adverbial relations like 'why', 'how', 'where', which I will not discuss here. Instead, there are resumptive pronouns, formally identical to the unstressed version of the personal pronoun. They either appear in the regular argument position or cliticize optionally onto the complementizer.

2.2. Distribution of resumptive pronouns

In local relativization,³ resumptive pronouns are found from the dative object⁴ on downwards on the hierarchy introduced by Comrie/Keenan (1977), including possessors, cf. van Riemsdijk (1989: 343, 345; 2003):⁵

- | | | |
|--------|---|----------------------|
| (1) a) | d=Frau, wo (* si) immer z=spaat chunt ⁶ | (subject) |
| | the=woman C (she) always too=late comes | |
| | 'the woman who is always late' | |
| b) | es Bild, wo niemert (* s) cha zale | (direct object) |
| | a picture C nobody (it) can pay | |
| | 'a picture that nobody can afford' | |
| c) | de Bueb, wo mer *(em) es Velo versproche händ | (indirect object) |
| | the boy C we (he:DAT) a bike promised have:1PL | |
| | 'the boy we promised a bike' | (van Riemsdijk 1989) |

² I will not discuss appositive relative clauses in this paper.

³ For reasons of space, I cannot go into long-distance relativization, where resumptives appear in all positions. See Salzmänn (in prep.) or van Riemsdijk (2003) for the data. However, a few examples will be used in arguments below.

⁴ This does not hold for all Swiss dialects, in several of them and even for some speakers of ZG, dative resumptives are impossible.

⁵ Unless otherwise noted the judgment represent those of my informants, mentioned in the first footnote.

⁶ The following abbreviations and symbols are used in this text: C = relative complementizer, SG = singular, PL = plural, NOM = nominative, ACC = accusative, DAT = dative, -OBLIQUE = -oblique case form, not specified for either nominative or accusative; Ø_D = empty determiner; the sign '=' stands for cliticization.

- d) D=Frau, won=i von=*(ere) es Buech übercho ha, find i müesam.
 the=woman C=I from=(she:DAT) a book got have:1SG find:1SG I annoying
 'I find the woman annoying from whom I got a book.' (P-object)
- e) Das deet isch de Typ, won=i geschter *(sini) Fründin ha wele verfüere
 that there is the guy C=I yesterday (his) girlfriend have:1SG wanted seduce
 'That's the guy whose girlfriend I wanted to seduce yesterday.' (possessor)⁷

3. Movement

In much of the literature in the 80ies and 90ies, it is tacitly assumed that the presence of resumptive pronouns automatically implies a non-movement relationship between antecedent and pronoun. Work by Demirdache (1991), Aoun et al. (2001), and Boeckx (2003), however, has shown that at least in some languages resumption is not incompatible with movement. I will argue in this section that movement is indeed involved in the derivation of ZG relative clauses. It is, however, not trivial to diagnose the presence/absence of movement when resumptive pronouns are employed because resumptives usually void locality violations. Therefore, further diagnostics are needed. I will therefore additionally discuss in some detail reconstruction and Crossover effects.

3.1. Locality

The question of movement is particularly difficult in the domain of locality because resumptive pronouns normally void island violations. In English, resumptives only occur in contexts where movement is not available. Inserting a resumptive rescues the construction:⁸

- (2) This is the man that I was wondering < whether you would like * ___/him >.

Since English does not use resumptives in non-island contexts, their presence is direct evidence that the movement operations in question are sensitive to locality. In Zurich German (and many other languages), however, resumptives also appear in positions where movement is expected to be available, e.g. in the matrix dative object position or the subordinate subject/direct object position. Consequently, the presence of a resumptive is not indicative of a locality violation and thus neither argues in favor nor against movement. Independent evidence is necessary to determine whether there is movement or not.

3.2. Reconstruction

Reconstruction, on the other hand, is a useful diagnostics for movement in resumptive structures. Crucially, we do find robust reconstruction effects in ZG even

⁷ Alternatively, possessors can also be rendered as complements of the preposition *von* 'of'. These forms are constructed like PPs.

⁸ Another term for this use is 'intrusive pronoun', cf. Chao/Sells (1983).

in the presence of resumptives. This represents clear evidence for movement. The following examples illustrate reconstruction for anaphor binding (3), principle C (4), bound pronouns (5), and idioms (6). Whether there is a gap or a resumptive, reconstruction always takes place:

- (3) a) S Bild vo *sich*_i, wo de Peter_i ___ wett verchaffe, gfallt niemertem.⁹
 the picture of self C the Peter wants sell pleases nobody
 'Nobody likes the picture of himself_i that Peter_i wants to sell.'
 b) PRO*_i/j s Grücht über *sich*_i, wo sich de Peter_i *drüber* uufregt...¹⁰
 the gossip about self C self the Peter about_it gets_worked_up
 'the gossip about himself_i that Peter_i is getting worked up about'
- (4) a)* S Bild vom Heiri_i, won er_i ___ gmaalet hät, isch sehr unvorteilhaft.¹¹
 the picture of_the:DAT Henry C he painted has is very unfavorable
 (*) 'The picture of Henry_i that he_i painted is very unfavorable.'
 b)* S Spiegelbild vom Heiri_i, won er_i ___ *devo*o verzelt, ...¹²
 the mirror_image_of_the:DAT Henry C he about_it speaks
 (*) 'The mirror image of Henry_i that he_i talks about'
- (5) a) S Bild vo *sine*_i Eltere, wo jede Schüeler_i ___ mitbracht hät,
 the picture of his parents C every pupil brought_with has
 hanget a de Wand.
 hangs on the wall
 'The picture of his_i parents that every pupil_i brought with him is hanging on the wall.'
 b) D=Periode vo *sim*_i Läbe, wo niemert_i gärn *drüber* ret, isch d=Pubertät.
 the=period of his life C nobody likes_to about_it talks is the=puberty
 'The period of his_i life that nobdy_i likes to talk about is puberty.'
- (6) a) De Sträich, wo mer em Lehrer ___ *gspilt* händ, isch echli krass gsii.
 the trick C we the:DAT teacher played have:1PL is a_bit extreme been
 'The trick we played on the teacher was somewhat extreme.'
 b) S *Fettnäppli*, won=i drii trampet bin, isch eigetli nöd z=überseh gsii.
 the faux_pas C=I in_it stepped am is actually not to=overlook been
 'The faux pas I made could in fact not be overlooked.'¹³

⁹ The invariant anaphor *sich* does not allow for logophoric use. The objections discussed in (Bhatt 2002: 50) therefore do not apply. Furthermore, the b-example shows that the availability of coreference cannot be due to a PRO within the head of the relative clause because Peter is unlikely to have spread offensive gossip about himself. If there is a PRO at all, it certainly is not the factor that explains coreference.

¹⁰ If inanimate antecedents are resumed in a PP position, pronominal adverbs are used as resumptives.

¹¹ Incidentally, ZG seems to be different from English w.r.t. reconstruction for Principle C, where coreference is often reported to be acceptable (Sauerland 1998, Safir 1999, Bhatt 2002: 85, note 24).

¹² Again, the b-example shows that coreference cannot be due to a PRO inside the head of the relative clause. *Spiegelbild* 'mirror image' does not allow for an additional argument, thus avoiding the complications discussed in Bhatt (2002: 50).

¹³ The idiom *in es Fettnäppli trampe* (lit.: to step into a fat bowl) means 'to put one's foot in one's mouth'.

Reconstruction is a means of teasing apart resumptives in languages where they are found in both island and non-island configurations. The above examples, which involve non-island contexts, have shown that movement is available, resumption is only apparent (Aoun et al. 2001). On the other hand, movement is expected to be absent if the resumptive is located inside an island. This is confirmed by the absence of reconstruction effects in the following example, which instantiates true resumption (Aoun et al. 2001):

- (7) * S Bild vo sich_i, wo ali lached, wänn de Peter_i s zeigt,
 the=picture of self C everyone laughs when the Peter it shows
 hanget i de Stube.
 hangs in the living room
 'The picture of himself_i that everyone laughs when Peter_i shows it, is hanging in the living room.'

Reconstruction is also found with the interpretation of adjectival modifiers, a diagnostic introduced in Bhatt (2002: 56-63). The adjective can apply to any of the predicates in (8), i.e. it is three-way ambiguous, showing the signs of successive cyclicity. Furthermore, ZG shows freezing effects with NPIs (9), and negation blocks reconstruction (10):

- (8) a) s **erschte Buech**, wo de Hans glaubt, [_{CP} ~~erschte Buech~~ dass ich bhaupte,
 the first book C the John thinks first book that I claim:IS
 dass es de Frisch gschriben hät]
 that it the Frisch written has
 'the first book that John thinks that I claim that Frisch wrote' → three-way
 ambiguous
- (9) b) s **einziges Buech**, wo d Susi gsäit hät, dass de Dürrenmatt s je gschriben hät
 the only book C the Susie said has that the Dürrenmatt it ever written has
 'the only book that Susie said that Dürrenmatt had ever written' → low reading
 only
- b) s **einziges Buech**, wo d Susi je gsäit hät, dass de Dürrenmatt s gschriben hät
 the only book C the Susie ever said has that the Dürrenmatt it written has
 'the only book that Susie ever said that Dürrenmatt had written' → high reading
 only
- (10) Das isch s **erschte Buech**, wo de Hans **nöd** gsäit hät,
 this is the first book C the John not said has
 dass es de Frisch gschriben hät.
 that it the Frisch written has
 'This is the first book that John didn't say that Frisch wrote' → high reading only

3.3. Strong Crossover¹⁴

Tests for Strong Crossover effects (SCO) need to be constructed with some care, as discussed in McCloskey (1990: 21 ff.). Especially, it is important that the pro-

¹⁴ I am very grateful to Rajesh Bhatt for insightful discussion of these issues.

noun that is crossed cannot be interpreted as the resumptive (i.e. the variable) and the putative resumptive as a coreferential pronoun. Therefore, one has to cross matrix subjects or direct objects which, cannot be resumptives. Under these provisions, ZG shows clear SCO effects:

- (11) * **De Bueb**_i, won **er**_i tänkt, dass d=Marie **in**_i gärn hät.
 the boy C he thinks that the=Mary him likes has
 *‘The boy_i who_i he_i thinks that Mary likes t_i.’

4. Matching effects

In this section, I will describe a property of Zurich German (and more generally Swiss German) relative clauses that so far has gone unnoticed, namely matching effects, governed by the following generalization:

- (12) The Zurich German Relative Clause Matching Generalization (ZGMG)
 resumptives and prepositions within the relative clause are deleted if the head noun
 i) bears the same case
 ii) is selected by the same preposition

4.1. The Basis of matching: identity in case/preposition

4.1.1. Prepositional relations and dative

- (13) a) Ich han **em** **Bueb**, [wo=t (***em**) es Buech versproche häsch],
 I have:1SG the:DAT boy C=you (he:DAT) a book promised have:2SG
 es schöns Exemplar ggee.
 a beautiful copy given
 ‘I gave the boy who you promised a book a beautiful copy.’
 b) Ich ha **vo** **de** **Frau**, [won=i scho geschter (***von=ere**)
 I have:1SG from the:DAT woman C=I already yesterday (from=she:DAT)
 es Buech übercho han], wider äis übercho.
 a book received have:1SG again one received
 ‘I received from the woman from whom I had already received a book yesterday another one.’

In these examples the resumptive and (where applicable) the preposition at the extraction site have to be deleted because the head-noun receives the same marking in the external context.

4.1.2. Subjects and objects

Subjects and direct objects are systemetically exempt from the matching requirement. At first sight, one might argue that the examples (1)a/b do in fact instantiate matching because the case form used for subjects and direct objects is identical in ZG (except for pronouns) and the case borne by the external element is that very

case as well. However, this would incorrectly predict the occurrence of resumptives for subject and object if the head noun is assigned dative case or governed by a preposition. In such configurations, resumptives are systematically absent as well:

- (14) a) **De Frau**, [wo (***si**) geschter cho isch], schuld i no
 the:DAT woman C (she:NOM) yesterday come is owe :1SG I still
 viel Gält.
 much money
 'I still owe the woman who came yesterday a lot of money.'
- b) **Vo de Frau**, [won=i (***si**) letschts Jahr in Kreta troffe
 from the:DAT woman C =I (she:ACC) last year on Crete met
 han], han i nie me öppis ghöört.
 have:1SG have :1SG I never anymore something heard
 'I've never heard again from the woman I met last year on Crete.'

4.2. The precise conditions for matching

In this section, I discuss the precise conditions for matching. I will look at constructions that minimally violate the generalization in (12), i.e. examples that do not share the same preposition but the same case or vice versa. Then I will further investigate whether thematic relations play a role. Lastly, I will investigate in how far different notions of case —like structural vs. inherent; abstract vs. morphological— play a role.

4.2.1. Mismatches in preposition, case, and case-assignment 1: only 1 PP

I will first discuss mismatches where only one clause contains both a P and a DP whereas the other one only contains a DP. I will only discuss cases where there is case-matching. In examples where there is no case matching, there are (of course) always resumptives. The first case combines an external P assigning dative case with relativization of the dative object within the relative clause:

- (15) Ich ha **vom Maa**, [won=i (***em**) es Buech ggee
 I have:1SG from_the:DAT man C=I (he:DAT) a book given
 han], geschter mis Gält übercho.
 have:1SG yesterday my money got
 'Yesterday I got the money from the man to whom I had given a book.'

Evidently, dative case on the external head licenses matching. In the reverse case with an external dative and a P + dative internally both the preposition and the resumptive are required in the relative clause:

- (16) Ich han **em Maa**, [won=i *(**von=em**) es Buech übercho
 I have:1SG the:DAT Man C=I (from=he:DAT) a book received
 han], zwänzg Stutz ggee.
 have:1SG twenty bucks given
 'I gave the man from whom I had received a book twenty bucks.'

4.2.2. Mismatches in preposition, case, and case-assignment 2: 2 PPs

The next class of mismatches involves PPs in both cases. In the first set of example, there is neither matching in case nor preposition. It is little surprising that both the resumptive and the preposition have to be spelled out:

- (17) Ich ha **für d=Lüüt**, [won=i *(**mit=ene**) i d=Schuel
 I have :1SG for the:ACC=people C=I (with=they:DAT) in the=school
 bin], ganz vil Schoggi gchauft.
 am very much chocolate bought
 'I bought a lot of chocolate for the people with whom I went to school.'

In the next example, there is case-matching, but the prepositions are different. Again, both the resumptive and the preposition are required in the relative clause:

- (18) Ich ha **vo de Lüüt**, [won=i *(**mit=ene**) i
 I have:1SG from the:DAT people C=I (with=they:DAT) in
 d=Schuel bin], scho lang nüüt me ghöört.
 the=school am already long nothing anymore heard
 'I have not heard for a long time from the people with whom I went to school.'

A further logical possibility involves prepositions that can assign different cases. If one combines the two different usages of one preposition, both the resumptive and the preposition must be spelled out as shown in the following example that combines the local (with dative) and the directional (with accusative) use of the preposition *in* ('in', 'into'):

- (19) Ich ha **i de Wonig**, [won=i morn *(**i=si**)
 I have:1SG in the:DAT apartment C=I tomorrow into=her:ACC
 iizieh], vil repariert.
 move much repaired
 'I have fixed a lot in the apartment into which I will move tomorrow.'

4.2.3. Mismatches in thematic relation

The previous examples suggest that the matching effects are form- and case-based. The following examples are used to test whether thematic roles also play a role:

- (20) a) Ich ha **vom Maa**, [won i *(**von=em**) gschlage
 I have:1SG from_the:DAT man C I (from=he:DAT) hit
 worde bin], nüüt me ghöört.
 was am nothing anymore heard
 'I haven't heard anything from the man by whom I was beaten.'

In this example, which combines a source and an agent relation, dropping both the resumptive and the preposition is obligatory. The same holds for the next example, which combines comitative with instrumental:

- b) De Hans hät sini Frau mit de Tusse, [won=i hüt Aabig
the John has his wife with the:DAT chick C=I today evening
(*mit=ere) is Kino gang], scho hüüfig betroge.
(with=she:DAT) into movie go:1SG already often cheated_on
'Hans has often cheated on his wife with the chick that I will go to the
movies with tonight.'

I conclude from this that the matching effect is not sensitive to thematic relations.

4.2.4. Different kinds of datives

The previous sections suggest very strongly that the matching effects are based on formal identity. The next step is to test whether all datives pattern the same. It has been suggested for German and German dialects that datives should be divided into structural and inherent datives, cf. Gallmann (1992), Wegener (1985, 1991) etc.¹⁵ Since subjects and direct objects do not show matching effects one might expect structural datives to pattern the same. However, all datives require resumptives as shown by the following examples:¹⁶

- (21) a) De Maa, wo=t *(em) geschter ghulffe häsch, isch
The:NOM man C=you (he:DAT) yesterday helped have:2SG is
immer no dankbar.
always still grateful
'The man who you helped yesterday is still grateful.'
b) Droge sind e Gfahr, wo mer *(ene) sini Chind nöd sött
drugs are a danger C one (they:DAT) one's children not should
ussetze.
expose
'Drugs are a danger that one should not expose one's children to.'
c) De Peter hät e Frau käne gleert, won er *(ire) gefallt.
the Peter has a woman know got C he (she:DAT) pleases
'Peter met a woman who likes him.'

A so-called structural dative with ditransitives can be found in (1)c, a structural dative with unaccusatives is represented by (21)c; (21) a-b are both inherent, in a) with an unergative verb and in b) with a ditransitive verb. One would therefore expect that all datives pattern the same w.r.t. matching. This is borne out:

- (22) a) Ich han em Maa, [wo=t *(em) ghulffe häsch],
I have:1SG the:DAT man C=you (he:DAT) helped have:2SG
geschter vo dir verzelt.
yesterday about you told
'I told the man that you helped about you yesterday.'

¹⁵ I will come back to diverging views on the nature of the dative in 7.1.1.

¹⁶ For unknown reasons, the unaccusative cases often sound rather awkward, and speakers resort to periphrasis.

- b) **Em Musiker**, [wo=t (*em) applaudiert häsch], würd
 the:DAT musician C =you (he:DAT) applauded have:2SG would: 1SG
 i kån Rappe gee.
 I no cent give
 'I would not give the musician who you applauded a cent.'
- c) **De Frau**, [wo=t (*ire) eusi Chind uusgliferet häsch],
 the:DAT woman C =you (she:DAT) our child exposed have:2SG
 gib ich nie me öppis z=ässe.
 give:1SG I never again something to=eat
 'I will never again give any food to the woman who we exposed our children to.'
- d) **De Frau**, [wo=t mich (*ire) vorziesch], häsch viel
 the:DAT woman C=you me (her:DAT) prefer:2SG have:2sg way
 z=vil versproche.
 too=much promised
 'You promised the woman that you prefer me to too much.'

In a) the external verb *tell* takes a structural dative whereas the verb inside the relative clause assigns inherent dative. The same holds for b). In c), both verbs are ditransitive, but the external one takes a structural dative whereas the internal verb assigns inherent dative. d) illustrates the same point. Incidentally, these examples provide further evidence that thematic relations are not at work. They also show that datives do form a coherent group in the grammar of ZG relativization.

4.2.5. *The importance of the surface form: different abstract Case but same form*¹⁷

It is a well-known fact that it is often the exact morphological form rather than the abstract case that plays a role in matching phenomena in free relatives, cf. Groos/van Riemsdijk (1981). It seems that a similar fact holds for ZG: Case is never formally marked on ZG nouns, but only on determiners and adjectives. Bare indefinite plurals without adjectives are therefore identical in all three cases. If matching is purely form-based, it can be predicted that a matching constellation always obtains with such DPs, regardless of the exact grammatical relation/abstract case of the head noun. This prediction is borne out, as the following example shows:

- (23) Ø_D Mane, won i (*ene) es Buech gib, müend intellektuell sii.
 D men(NOM) C I (they:DAT) a book give:1SG must:PL intellectual be
 'Men to whom I give a book must be intellectual.'

The head noun is the subject of the main clause and thus assigned abstract nominative case. Inside the relative clause it functions as a dative object. The form *Mane* is underspecified morphologically, it can be used in all three cases. The crucial thing here is: Since *Mane* can be interpreted as a dative object, matching is possible, and no resumptive occurs.

¹⁷ I am grateful to Kathrin Würth for drawing my attention to this fact.

4.3. Matching and movement

While non-matching configurations show unambiguous signs of movement, we still have to test whether this also holds for examples involving matching. In the following example, reconstruction occurs under matching:¹⁸

- (24) **Jedem** Artikel über sichi, wo de Peteri (*em) misstrout, begägnert
 every:DAT friend about self C the Peter (he:DAT) distrusts counters
 er mit Aggression.
 he with aggression
 'Every article about himselfi that Peteri distrusts, he counters with aggression'

This shows that matching relatives are also derived via movement.

4.4. Conclusion

I have established in the previous subsections that matching effects are form-based: Identity of Preposition and/or case is required while identity of thematic relation is not. I have furthermore shown that the difference between structural and inherent datives is irrelevant for matching, and that the matching generalization is sensitive to the actual surface form. Reconstruction effects under matching show that matching relatives must also be given a movement analysis.

5. The interpretation of resumptives

One of the crucial properties of resumptives that helps categorize the different types is their interpretation. Furthermore, there are sometimes asymmetries between gaps and resumptives that need to be explained. As shown in Chao/Sells (1983), English resumptives, which only occur to prevent island violations, are not compatible with a bound variable interpretation. Consequently, they cannot have non-referential antecedents, i.e. quantifiers like *every*, *no* etc.:

- (25) a) I'd like to meet **the linguist** that Mary couldn't remember if she had seen __/him before.
 a) I'd like to meet **every linguist** that Mary couldn't remember if she had seen __/*him before.

The readings that are available for resumptives in English have been subsumed under the E-type reading. Languages like Hebrew, Lebanese Arabic or Swedish (and many others), however, have resumptives that do allow bound variable readings (Chao/Sells 1983, Aoun et al. 2001), cf. the following example from Lebanese Arabic, where the antecedent is linked to a resumptive in the complement clause (Aoun et al. 2001: 390):

¹⁸ There are independent reasons why SCO cannot be tested in such examples: Local relativization across a dative is ruled out independently by principle B: the (resumptive) pronoun would be bound in its governing category by the relative clause subject; and long relativization is not sensitive to matching. Therefore, the necessary constellation to test SCO effects under matching does not exist.

particular languages without paying too much attention to the cross-linguistic variation. It is also partly a result of the fact that some properties of resumption — especially the absence of movement— are often taken for granted without actually testing them. Work by Demirdache (1991), Aoun et al. (2001), and Boeckx (2003) has challenged these positions and will prove at least partially useful for the analysis of the ZG data. The only explicit account of the ZG data is by Van Riemsdijk (1989, 2003); I will discuss it at the end of this section.

6.1. Non-movement Approaches

McCloskey 1990, Shlonsky (1992), Suñer (1998), Rouveret (2002), and Adger & Ramchand (2004) all propose a base-generation approach to resumption even though the languages under question have different types of resumptives. The first three deal with Irish, Hebrew and Spanish, where resumptives are not sensitive to islands. The reason for base-generation rests solely on this fact. Other diagnostics for (non-)movement are not considered or as in McCloskey's and Shlonsky's account of SCO and WCO effects receive a representational analysis. Rouveret (2002) and Adger/Ramchand (2004) on the other hand are confronted with a very different problem: In Welsh and Scottish Gaelic, the following paradoxical situation obtains: While resumptives are sensitive to strong islands, there is otherwise no unequivocal evidence for movement (no reconstruction for binding, idioms and [sometimes] scope). These properties are captured by the assumption that the A'-dependencies are established via Agree without subsequent move.

The second type of approach is geared towards languages with properties very different from those of ZG so that it need not be considered. The first class of approaches on the other hand fails to explain the reconstruction effects and more generally the properties of movement. Furthermore, the matching effects are completely unexpected under such approaches: The external context of the head noun should in no way influence the choice between movement and base-generation. Base-generation is therefore not an option for ZG. However, there is one aspect that will prove fruitful in the analysis of ZG: Resumptives appear in those languages also to prevent (illicit) preposition stranding. As we will see, some resumptives in ZG occur for the same reason.

6.2. Movement approaches

Movement approaches to resumption have become more prominent in recent years. One can distinguish at least three different types: Movement at LF (Demirdache 1991, 1997), the Big DP analyses (Aoun et al. 2001, Boeckx 2003), and those that treat resumptives as spelled out traces (Pesetsky 1998, Grohmann 2003). They make the correct prediction that resumption is compatible with movement effects. However, most of them contain certain features that fail to explain the ZG data.

Demirdache (1991/1997) argues that resumptives are in-situ operators, i.e. operators that move at LF. This assumption manages to capture the paradoxical nature of resumptives in Hebrew. Despite the absence of locality effects, resumptives show

movement properties. The trigger SCO and WCO effects (see especially Demirdache 1997), license parasitic gaps and allow for reconstruction. Since LF-movement is assumed to be insensitive to Subjacency, the non-locality of many resumptive constructions falls into place. Independent evidence for movement of resumptives comes from optional resumptive fronting (Demirdache 1997: 195). However, there are a number of reasons to reject her approach, both related to the proposal as such and the ZG data: While LF movement usually does not show subjacency effects it is still (often) assumed to be sensitive to the CED. Therefore, resumptives within adjunct and subject islands remain a problem unless more is said (cf. Aoun/Li 1993 on Chinese *wh-in-situ*). Furthermore, Parasitic Gaps are normally assumed to be licensed at S-Structure (Culicover 2001), but not at LF. Since resumptives do not move until LF, the licensing of Parasitic Gaps in Hebrew under resumption is unexpected. As for the ZG data, there is one major problem: As shown in (7) the possibility of reconstruction correlates with locality: Reconstruction into islands is not possible, suggesting that movement is absent. Demirdache, on the other hand, predicts reconstruction into islands to be freely available. Unfortunately, she does not present any reconstruction data except for one sentence (1991: 96), which involves an island, but an apparently non-local anaphor. Therefore, nothing really follows from this, and even if correct for Hebrew, the LF movement analysis makes the wrong predictions for ZG.

Of the big DP analyses, I will first discuss Boeckx (2003), although such a brief overview cannot do full justice to his complex account. His basic idea is that resumptives are first merged with their antecedents which in the course of the derivation move away from them. The availability of resumptives is correlated with non-agreeing (roughly: non-inflecting) complementizers. In these cases, movement can take place under Match, but without Agree. This explains the island-insensitivity in many languages. Apart from the circular reasoning (non-agreeing complementizers are those that appear with resumptives and can span islands) and some other inconsistencies (see Salzmann in prep.), there are a number of descriptive facts that cannot be accounted for: The major problem is that Boeckx predicts reconstruction into islands. However, this is at least not correct for ZG: (7). Secondly, the matching effects are completely unexpected: Under Boeckx' approach it is only the complementizer that determines the possibility of resumption. The external context of the head noun should not play a role. Lastly, as discussed in 5, resumptives in ZG have the interpretation of variables, they are interpreted just like gaps. A crucial ingredient of Boeckx' approach is, however, that the resumptive is a pronoun and as such affects the interpretation of the antecedent (D-linked, specific etc.). This is not observed in ZG.

Aoun et al. (2001) also assume a Big DP analysis. The major difference from Boeckx (2003) is the assumption that the antecedent cannot move out of islands. When a resumptive is found inside an island, the antecedent is base-generated in the operator position and linked to the resumptive via binding. This correctly accounts for the movement properties of resumption in ZG and for the non-availability of reconstruction in island contexts. There are at least three problems, the first one general: Aoun et al.'s approach simply does not explain the MI distribution of resumptives: Why is such a complex DP created at all if there is no island? Why isn't

it sufficient to just move the antecedent, leaving a gap? This is sometimes obligatory (e.g. dative object in ZG), sometimes optional (direct object in Hebrew)? Two problems particular to ZG remain: The matching effects are unexpected, especially under the assumption that antecedent and resumptive agree in all relevant features. The ZG facts, however, require obligatory disagreement in case. Furthermore, it is again unexpected that the external context should play a role. Finally, resumptives in ZG do not have the interpretation of a pronoun, but that of a gap.

As for approaches that consider resumptives the spell-out of a trace/copy, I will not review them in much detail because most approaches that are available are either not very detailed (Pesetsky 1998) or are based on different data (left-dislocation, Grohmann 2003). Since my approach to be presented below can be subsumed under this general idea, I will only briefly discuss some of the advantages of these approaches and the questions that they raise. The major advantage of such approaches is that they explain the movement properties and the interpretive properties (at least in ZG). If a resumptive is just the spell-out of a copy, it is not expected to behave like a pronoun. However, if resumptives are indeed the spell-out of a copy, one has to explain a) why the trace is spelled out at all and b) why it is spelled out as a pronoun. The first question is difficult in those cases where movement is possible, i.e. in cases of apparent resumption. Some independent property of the language must account for this. As for b), one can argue that for reasons of economy (Pesetsky 1998), it is sufficient to spell out an element that realizes the phi and case features only. However, one might object that a determiner would do the same job. This connects to a further problem: If one distinguishes between true and apparent resumption, it is purely accidental that the shape of resumptives is the same in both cases. I will show below that for ZG these difficulties can be overcome in a straightforward and explanatory way.

6.3. On Swiss German: van Riemsdijk (1989, 2003)

To conclude this section, I will discuss in more detail van Riemsdijk's work on ZG relative clauses. The distribution of resumptives is said to follow from an independently available process of cliticization: subject, direct and indirect object pronouns often cliticize onto C. This brings resumptives 'close enough to the head of the relative to permit deletion' (van Riemsdijk 1989: 347). This 'explains' the obligatoriness of resumptives in prepositional relations: Since there is no preposition stranding, the pronoun cannot cliticize onto C.

The fact that the dative clitic must not be deleted in some dialects is stipulated to follow from the fact that datives are in fact PPs. Van Riemsdijk derives this from the phonological similarity between datives and locative expressions. More specifically, van Riemsdijk analyzes dative clitics as amalgamations of the preposition *a* 'to' plus pronoun. Deletion of the entire complex is then prohibited by the ban on recoverability of deletion, and moving only the clitic is impossible because it is in some way (which van Riemsdijk does not specify) not independent enough to move on its own.

There are a number of problems with this proposal, one conceptual, several empirical. As for the conceptual problem, van Riemsdijk has to assume that cliticiza-

tion is obligatory in relativization while it is optional elsewhere. The obligatoriness is derived from the Avoid Pronoun Principle, a transderivational constraint. The movement takes place so that the pronoun can later be deleted. Clearly, this involves non-trivial look-ahead: the grammar somehow has to know that it first HAS to move the clitic so it can later be deleted. Needless to say, such an approach is in stark contrast with the tendency within Generative Grammar to move away from transderivational evaluation.

There is a large number of empirical problems: The first involves the absence of A'-movement. Van Riemsdijk (1989: 344) explicitly states that Swiss German relatives — also those involving matrix subjects and direct objects do not involve A'-movement. This seems to imply that clitic movement is not an A'-movement process. Consequently, there is no A'-dependency in relative clauses. All he assumes is some co-indexing mechanism between the resumptives and the head-noun (perhaps mediated by C or Spec, CP). Such an approach makes strong predictions: Since there is no operator-variable relation, we expect the extraction site to have the semantics of a pronoun, and we do not expect any movement properties. The second point has been shown to be incorrect: Reconstruction effects and SCO effects clearly argue in favor of movement. The fact that resumptives also occur in islands does not mean that movement is never involved. Van Riemsdijk —like the other base-generation approaches— fails to distinguish between true and apparent resumption. Furthermore, the SCO effects and (28) and (29) show that resumptives are interpreted like gaps, not like pronouns. Even more problematic, it is not clear how such an analysis derives the correct semantics for restrictive relatives. It is normally assumed that movement inside the relative clause derives a predicate which combines with the head-noun via intersective modification. It is unclear to me how this can be achieved given van Riemsdijk's analysis. A further problem concerns datives. The explanation for the failure to delete the dative clitic is difficult to evaluate. There are many Swiss dialects that express dative with the additional help of a preposition-like element, *a* 'at' or *i* 'in', cf. Seiler (2001); this even holds though only sporadically for ZG. But if it is possible in ZG, it is highly unlikely that the forms we find without the extra element also represent PPs. Furthermore, in those dialects that make systematic use of this dative marker, it is incompatible with dative clitics (and unstressed pronouns more generally), cf. Seiler (2001: 251); it would be strange if ZG were an exception to this stable restriction. Also, Van Riemsdijk has to assume that it is possible to have a preposition governing prepositions e.g. when a preposition like *mit* 'with' assigns dative to a clitic: *mit em* 'with him'. According to him it would actually govern a PP. Interestingly, this happens to be impossible in those dialects which unambiguously use a preposition-like element, cf. Seiler (2001: 251). Furthermore, van Riemsdijk has to assume for those dialects which do not use dative resumptives that there the very same string *em* 'to him' does not have the status of a PP. While not impossible, such a solution is ad hoc and in the absence of independent evidence a restatement of the facts. There are also technical problems: If dative clitics are indeed PPs, one has to explain how they can actually cliticize onto a head in the left periphery. It is unclear why this option does not exist for normal PPs. Van Riemsdijk seems to assume that cliticization is rather phonological in nature, i.e. dative clitics are the only PP-elements that are light

enough to undergo this process. But then, it is unclear why in the case of the other PPs it is impossible to move only the light clitic and strand the preposition. If the movement is phonological, then there is nothing like the ECP that rules out preposition stranding.

The matching effects are generally unaccounted for in van Riemsdijk. They show that (some form of) deletion does play a role in the derivation of relative clauses in ZG, yet crucially does not involve subjects and direct objects. This suggests that properties other than the phonological weight of resumptives must be at stake but some sort of identity requirement that licenses the deletion of resumptives. I conclude that van Riemsdijk's approach is inadequate, both conceptually as well as empirically.

7. The account

In this section I present the assumptions necessary for a formal account. I first discuss the distribution of resumptives, which shows that they occur for reasons of morphological licensing of oblique case/prepositions and to prevent preposition stranding. Then, I discuss the derivation of relative clauses concluding that the head raising analysis is the most useful one for the data at hand, in particular because it provides a means (via incorporation) to link the relative clause internal context with the relative clause external context. This will be shown to underly the matching phenomenon. Non-matching configurations, on the other hand, are linked to the independently available mechanism of Distributed Deletion.

7.1. The distribution of resumptives in ZG relatives

7.1.1. *Subject/direct object vs. oblique*

From the data presented at the beginning, it becomes clear that there is a division between subject and direct object on the one hand and the other relations on the other. The distinction between subject/direct object (which I will refer to as direct arguments) and PP is unproblematic. The two direct arguments are licensed via abstract case, whereas PPs do not have to be case-licensed. The reason why there are resumptives in the latter case will be discussed in 7.1.2, in this subsection, I want to focus on the contrast between the direct arguments and datives. The division is, of course, reminiscent of the difference between structural and inherent case. This distinction correlates with a morphological distinction: While nominative and accusative are identical except for certain pronouns, the dative, which is the major (and almost only) case in oblique relations (some prepositions assign accusative), is clearly distinct. But is this correlation meaningful?

While it is undisputed that there are different types of datives and that some of them show certain properties reminiscent of structural arguments (predictability of their position, get-passive, cf. Wegener 1985, 1991, Gallmann 1992), all datives also differ systematically from nominative and accusative as shown convincingly in Vogel/Steinbach (1998) and Bayer et al. (2001). I will not review all of their arguments, but will simply mention two: Datives cannot bind anaphors while direct ob-

jects can, (30) (Vogel/Steinbach 1998: 73), and datives are barriers for extraction (31) (Vogel/Steinbach 1998: 74f.):

- (30) a) dass der Arzt_i **den** Patienten_j sich_{i/j} im Spiegel zeigte
 that the:NOM doctor the:ACC patient self:DAT in _the mirror showed
 'that the doctor showed the patient to himself in the mirror.'
 b) dass der Arzt_i **dem** Patienten_j sich_{i/j} im Spiegel zeigte
 that the:NOM doctor the:DAT patient self:ACC in _the mirror showed
 'that the doctor showed the patient to himself in the mirror.'
- (31) * [Über wen]_i hat der Verleger [einem Buch_t] keine Chance gegeben?
 about whom has the editor self:DAT book no chance given
 Lit.: , About whom has the editor given a book no chance?'

This oblique behavior correlates with special morphological licensing conditions. Like the oblique case genitive, dative requires overt case marking to be licensed as the following four asymmetries show: First, complement clauses in German cannot directly fill the slot of a dative argument (Bayer et al. 2001: 471):

- (32) a) Wir bestritten, [dass wir verreisen wollten]. ACC
 we denied that we travel_away wanted
 'We denied that we wanted to go away.'
 b)* Wir widersprachen, [dass wir verreisen wollten]. DAT
 we objected that we travel_away wanted
 'We denied that we wanted to go away.'
 c) Wir widersprachen [der **Behauptung**, [dass wir verreisen wollten]]. DAT
 we objected the:DAT claim that we travel_away wanted
 'We denied that we wanted to go away.'

CPs cannot realize morphological case. A DP has to be inserted to rescue the example. The structural cases nominative and accusative do not require this extra licensing, abstract case is sufficient. Certain indefinite quantifiers in German do not inflect for case. Interestingly, they can function as bare subjects or direct objects but not as datives (Bayer et al. 2001: 472):

- (33) a) Wir haben **genug** / **nichts** / **allerlei** / **etwas** / **wenig** **erlebt** ACC
 we have enough nothin_a_lot something little experienced
 'We have experienced enough/nothing/a lot/something/little.'
 b)* Feuchtigkeit schadet **genug** / **nichts** / **allerlei** / **etwas** / **wenig** DAT
 humidity harms enough nothing_a_lot something little
 'Humidity harms enough/nothing/a lot/something/little.'

Some of these adjectives have an inflected form, which is optional for the structural cases, but obligatory for datives (Bayer et al. 2001: 472):

- (34) a) Wir haben schon viel-(es) / nur wenig-(es) erlebt.
 we have already much-(ACC) only little-(ACC) experienced
 'We have experienced much already/only little.'
 b) Das schadet/ gleicht/ ähnelt viel-*(em)/ wenig-*(em).
 that harms equals resembles much-(DAT) little-(DAT)
 'This harms equals/resembles much/little.'

There are two further arguments from recoverability which show that the dative is subject to specific licensing conditions: Topic drop is only possible with direct arguments, but not with datives (35) Bayer et al. (2001: 489), and in comparatives, only direct arguments can be deleted, datives require resumptives (36) (Bayer 2002: 15):

- (35) a) [] Hab' ich schon gesehen b)* [] Würde ich nicht vertrauen
 have I already seen would I not trust
 'I already seen (it).' ACC 'I wouldn't trust (him)' DAT
- (36) a) Mehr Patienten sind gekommen als [NOM_] behandelt werden konnten.
 more patients are come than treated become could
 'More patients showed up than could be treated.'
- b) Mehr Patienten sind gekommen als der Arzt [DAT *(ihnen)]
 more patients are come than the doctor they:DAT
 Medikamente geben konnte.
 medicine give could
 'More patients showed up than the doctor could give medicine to.'

All these observations hold for ZG as well. The fact that the dative is also special in ZG relativization thus comes as no surprise. It is simply another instance where morphological licensing requires dative case to be spelled out. I conclude from all these facts that datives are indeed crucially different from nominative and accusative, and that what causes dative resumptives is a condition on the licensing of oblique cases. The fact that dative resumptives can be dropped under matching suggests that under specific circumstances, oblique cases CAN be recovered. I will formally implement this in 7.3.

7.1.2. Resumptives to prevent Preposition Stranding

Prepositions are similar to datives in that they are normally not recoverable if not expressed morphologically (Bayer et al. 2001: 489), i.e. the same arguments for clausal licensing, topic drop, and comparatives apply here as well. Consequently, it comes as no surprise that they also cannot be dropped in relative clauses.¹⁹ But this still does not explain why resumptives occur as well, as in the following example ((1)d, repeated here):

- (37) D=Frau, won=i von=*(ere) es Buech übercho han,
 the:ACC=woman C =I from=(she:DAT) a book received have:1SG
 find i müesam.
 find:1SG I annoying
 'I find the woman annoying from whom I got a book.'

I argue that this follows from a general ban on preposition stranding in ZG (cf. Fleischer 2001: 123f.). In this area, resumptives act as a last resort to prevent a local-

¹⁹ It is tempting to capture this similarity with the KP hypothesis (Bayer et al. 2001), where both oblique morphological case and prepositions license a KP layer on top of oblique DPs.

ity violation. Just like datives, P can be dropped in very specific constellations, namely when the head noun of the relative clause is governed by the same preposition. In this constellation, the content of P is recoverable. A formal account is presented in section 8.2.

7.2. The derivation of relative clauses in ZG

I assume that restrictive relative clauses in ZG are derived via head raising. The head raising analysis (HRA) goes back to Brame (1968), Schachter (1973), and Vergnaud (1974). More recently, it has been revived by Kayne (1994), Bianchi (1999, 2000a, 2000b), Bhatt (2002), and De Vries (2002). The crucial argument in the present context in favor of the head raising analysis comes from reconstruction: Since the head noun starts out inside the relative clause, it comes as now surprise that it can be interpreted inside the relative clause via reconstruction/interpretation of the lower copy.²⁰ On the head external analysis (Chomsky 1977) reconstruction effects are at least problematic because the head noun is not directly linked to a relative clause internal position, but only via the *wh*-operator.²¹ For present purposes I simply follow recent work that takes reconstruction effect to be decisive evidence in favor of the HRA. When adopting the HRA, there are still a number of options that have been discussed in the literature:

The head NP stays inside the relative CP (Kayne 1994, De Vries 2002) or the head NP moves out of the relative CP (Bianchi 1999/2000a-b, Bhatt 2002). On Kayne's/De Vries' approach, if the relative operator is zero (as in *that*-relatives), there is just movement to Spec, CP, and nothing further happens, (38)a. If the operator is complex, the head noun moves to the specifier of the relative operator, (38)b, in de Vries (2002: 123ff.) with subsequent feature movement from N to the external D (38)c. On Bianchi's and Bhatt's approach, the head noun moves out of the relative clause to adjoin to the CP ((39)a, cf. Bhatt 2002) or moves to the specifier of some higher functional head ((39)b, cf. Bianchi 1999/2000a-b, Bhatt 2002: 84):

- (38) a) the [_{CP} [N_P book]_i that John likes t_i]
 b) the [_{CP} [NP book]_j [_{N'} which t_j]]_i John likes t_i]
 b) FF_j + the [_{CP} [NP book]_j [_{N'} which t_j]]_i John likes t_i]
 (39) a) the [book]_j [_{CP} [NP Op/which t_j]]_i John likes t_i]
 b) the U [book]_j [_{X'} X° [_{CP} [NP Op/which t_j]]_i John likes t_i]]

The last two derivations are necessary to account for extraposition because the external determiner and the head noun form a constituent to the exclusion of the relative CP. Since I will not discuss extraposition here, I will ignore this complica-

²⁰ There is a large number of other (crosslinguistic) evidence for the HRA that I cannot review here, but see Kayne (1994), Bianchi (2000a), and De Vries (2002).

²¹ Admittedly, this assumption is to some extent theory internal and certainly has to do with the way reconstruction is handled in the Minimalist Program, i.e. as the interpretation of a lower copy. Reconstruction via binding of operators does not find a natural place within the Minimalist system even though it is by no means implausible as such.

tion and simply assume movement to an operator position within the relative clause. Another point where the approaches differ is the type of category that is raised. Kayne (1994) originally proposed that relatives only involve raising of an NP. There are a number of facts that lend some initial credibility to this proposal, for instance, there are no definiteness effects if a definite head noun combines with a relative clause involving there: *the book that there was on the table*, scope reconstruction with a definite head noun, and the exceptional compatibility of definiteness with proper names and idioms like *the Paris* *(I like). On the other hand, Bianchi (1999, 2000b), convincingly demonstrated that assuming the raising of only an NP is problematic: First, it has been shown (e.g. Longobardi 1994) that arguments are DPs while NPs can only serve as predicates. Second, the XP that is moved behaves like a referential phrase (in Cinque's 1990 terms) w.r.t. locality, i.e. it can be extracted across weak islands, and it can license PRO. So there are good reasons to assume both NP and DP raising at the same time. The paradox can be resolved in two ways: Movement to Spec, CP is movement of a DP, but then, the step that moves the NP out of the CP applies only to the NP as in Bianchi (2000b) and Bhatt (2002). Alternatively, there is DP-movement to Spec, CP with subsequent incorporation of an underspecified D into the external D (Bianchi 1999/2000b).

I will largely follow Bianchi (1999, 2000b) because there is hardly any evidence in ZG that only an NP is raised. The definiteness effect does not exist in ZG, the examples with idioms and proper names can be explained semantically (the relative clause affects to head noun so that it no longer denotes a unique individual) and finally, there is no scope reconstruction:

- (40) Ich ha de zwei Patienten aagglüüte, wo jede Tokter morn
 I have:1SG the:DAT two patients called C every doctor tomorrow
 untersucht.
 examines
 'I called the two patients that every doctor will examine tomorrow.'
 2 > every *every > 2

Such examples only have an individual reading, and crucially no distributive reading. This follows if what is reconstructed is a fully specified DP (this is where I differ from Bianchi as she assumes that D is underspecified for definiteness) so that the same reading obtains as in simple clauses.

Although there are no relative pronouns in ZG I assume that the DP that is moved is headed by a D with an operator feature, so that it corresponds to an empty relative pronoun (cf. De Vries 2002: 126). Movement is triggered by an operator feature on C against which the respective feature on D is checked, thereby avoiding some of the complications of Bianchi's (2000b) system, cf. De Vries (2002: 115). The final (simplified) derivation looks as follows (the incorporation of D will be discussed in the next section):

- (41) [_{DP}D + D_I[_{CP} [_{DP} [_{DP}t_iNP]_j] [_{IP} [_{VP} t_j V]]]]

7.3. Matching as incorporation

The major reason why datives and PPs have to be spelled out in ZG relative clauses is recoverability: As oblique phrases, they cannot be structurally licensed, but instead require morphological licensing. Under very specific circumstances, this licensing requirement seems to be lifted, namely when the head noun receives exactly the same type of morphological marking, i.e. under matching. I conclude from this that the oblique marking is recoverable under matching. For a formal account, we need a link between the external context, i.e. the external D and P and the internal context, i.e. the moved DP/PP. I propose that matching is to be understood as incorporation of relative clause internal material (i.e. D/P) into external material, i.e. D/P. If the complex heads derived via incorporation have compatible case features, the oblique case/the P is accessible for the relative clause internal copy because it is part of a (modified) chain that includes the required morphological expression, namely on the complex D/P head. This is why dative resumptives and prepositions can be dropped under matching. Incorporation leads to the following representations:

- (42) a) $[_{DP} D_i + D [_{CP} [_{DP} t_i NP]_j, C [_{IP} [_{VP} t_j V]]]]$
 b) $[_{PP} P_k + P [_{DP} D_i + D [_{CP} [_{PP} t_k [_{DP} t_i NP]]_j, C [_{IP} [_{VP} t_j V]]]]]$

These derivations raise three questions: a) What triggers this movement? b) Why aren't the PP-examples out because of a violation of the Head Movement Constraint (HMC)?²², and c) how are these complex heads spelled out?

Ad a): I assume that D (and P) can carry features that attract a head of the same category, but only if they (D, or, in the case of P, their complement) select a relative clause. This can be stated economically in the lexical entry of D (and P).

Ad b): The PP-derivations violate the HMC (Travis 1984) because D moves across P, and P moves across the external D: However, under the Minimal Link Condition (MLC, Chomsky 1995) a different interpretation is possible: An intervening head will only block movement if it could check the same feature, i.e. if it in some relevant sense of the same type. But since the attracting feature is only sensitive to the exact grammatical category, a D will not block movement of P, and neither will P block the movement of D. Similar arguments have been used for instances of long head movement, cf. Carnie et al. (2000).²²

Ad c) I assume a late insertion approach to morphology, in the spirit of Halle/Marantz (1993). This means that the syntax only manipulates features. When Vocabulary Insertion takes place at PF, the complex heads derived via head movement are spelled out as one lexical item if their parts agree in phi and case features (dative) and lexical features (prepositions). If insertion is successful, i.e. if a lexical item can be found that is compatible with the feature requirements, we obtain matching. If there are conflicting features (e.g. different case features) insertion fails, and the derivation crashes.

²² In those cases, the relevant distinction is usually between A vs. A'-head position. A possible analogy to the present case might be the requirement of T to have a DP in its specifier (i.e. the EPP). Intervening maximal categories like VP, vP, NegP etc. do not block movement of a DP because they simply belong to a different category type.

So far we know how matching cases are derived, but we still need to explain the non-matching cases. The next section shows that they are based on a mechanism that is independently available in ZG: Distributed Deletion.

7.4. A'-splits as Distributed Deletion

ZG A'-movement generally allows for a peculiar way of spelling out operator and case information: In addition to regular full category movement, it is optionally possible to spread operator and case information over two copies: A case-unmarked DP appears in the operator position while case (including prepositions) is realized in the base position. This is an instance of Distributed Deletion. I will refer to these constructions as A'-splits. The following examples illustrate this for wh-movement:

- (43) a) **Wer** häsch gsäit, dass ich *(em) das Buech cha verchaffe?
 who:-OBLIQUE have:2SG said that I he:DAT the book can sell
 'To whom did you say that I can sell the book?'
 b) **Wer** häsch gsäit, dass=t *(mit=em) wettsch go tanze?
 who:-OBLIQUE have:2SG said that=you (with=he:DAT) would_like go dance
 'With whom did you say that you would like to go dancing?'

These constructions are derived as follows: The case feature can optionally be deleted after checking. As a consequence, the moved phrase will only have an operator feature on D but no longer a case feature. Both case and operator information must be spelled out (their features are strong), but because they are not present in the same copy, parts of both copies are spelled out, expressing the respective feature content. There is a certain amount of overlap (D is realized twice) because D hosts the two crucial features.²³

It remains to be explained why case is realized as a pronoun and not just as a D element, i.e. as a determiner. I suggest that this follows from a constraint that requires the resulting copies to conform to the normal structure of DPs. This type of regeneration is generally found with split DPs, cf. Fanselow/Cavar (2002).

What determines the availability of this type of splitting? I assume that ZG has a crucial property that makes this possible: Case is only represented on D, but no longer on N.

The crucial point for the current discussion is that such a derivation lies at the heart of resumption in ZG: The case feature is only present in the base position and has to be spelled out there. The DP that moves on is underspecified for case so that as a consequence there will never be a feature clash on the complex D head. The following section provides an explicit account of all the derivations.

²³ The precise mechanism assumed here for Distributed Deletion is somewhat different from e.g. Fanselow/Cavar (2002).

8. ZG relative clauses and distributed deletion

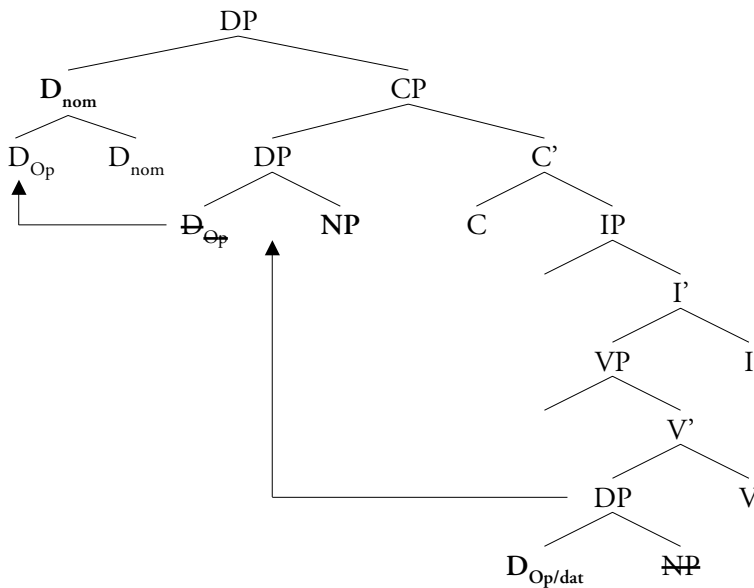
8.1. Dative

I will first discuss the non-matching examples, cf. (1)c, repeated here:

- (44) **De** **Bueb**, [wo mer *(em) es Velo versproche händ], isch tumm.
 The:NOM boy C we (he:DAT) a bike promised have:1PL is stupid
 'The boy who we promised a bike is stupid.'

Suppose a normal A'-derivation with movement of a fully specified DP to Spec, CP. Subsequent incorporation of D into the external D will lead to a case clash (nom vs. dat) so that insertion fails and the derivation crashes. An A'-split derivation, however, derives the desired result: the fronted DP is underspecified so that incorporation of D and insertion succeed. Spelling out the case feature in the base position leads to a resumptive. The following structure represents the converging derivation (bold-faced constituents are spelled out, strike-through means non-pronunciation, irrelevant parts are omitted):

- (45) [_{DP} D_{Op} + D_{nom} [_{CP} [_{DP} ~~D_{Op}~~ NP] C [_{IP} [_{VP} [_{DP} D_{Op/dat} ~~NP~~] V]]]]

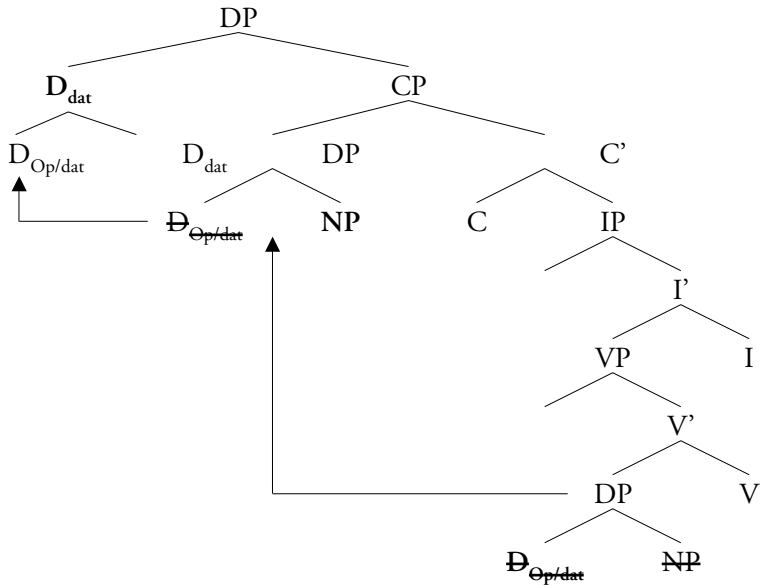


Tree 1

A matching derivation simply involves moving of a fully specified DP. Subsequent incorporation leads to case compatibility so that insertion succeeds. No case feature is left inside the relative clause and as a consequence no resumptives appears, cf. (13)a, repeated here:²⁴

²⁴ The same derivation applies to (15).

- (46) Ich han **em** **Bueb**, [wo=t (*em) es Buech versproche häsch],
 I have:1SG the:DAT boy C=you (he:DAT) a book promised have:2SG
 es schöns Exemplar ggee.
 a beautiful copy given
 'I gave the boy who you promised a book a beautiful copy.'
 (47) [_{DP} D_{Op/dat} + D_{dat} [_{CP} [_{DP} ~~D_{Op/dat}~~ NP] C [_{IP} [_{VP} [_{DP} ~~D_{Op/dat}~~ NP] V]]]]



Tree 2

8.2. PPs

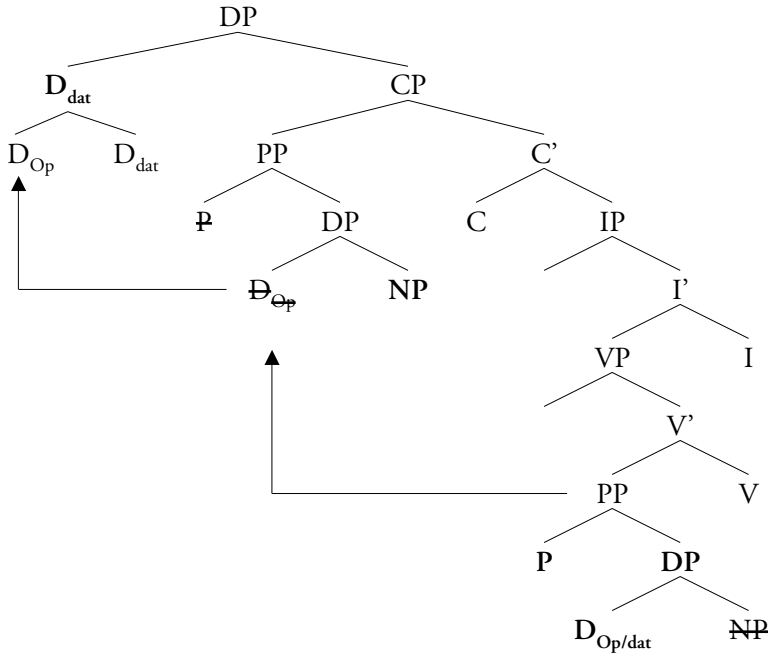
I will first discuss a case where there is only an external D but no P and case matching, cf. (16), repeated here:

- (48) Ich han **em** **Maa**, [won=i *(von=em) es Buech übercho
 I have:1SG the:DAT Man C=I (from=he:DAT) a book received
 han], zwänzg Stutz ggee.
 have:1SG twenty bucks given
 'I gave the man from whom I had received a book twenty bucks.'

Movement of a fully specified DP with subsequent incorporation leads to compatible case features so that no resumptive is expected. However, this leads to preposition stranding, and the derivation crashes. The only converging derivation involves Distributed Deletion. As a consequence, case is spelled out in the base position. It remains to be explained why the preposition is also spelled out in the base position and not in Spec, CP. It is again the ban preposition stranding which favors pronunciation of the lower copy (P would be

without a D in Spec, CP). The final output is thus due to a conspiracy of factors.²⁵

- (49) [_{DP} D_{Op} + D_{dat} [_{CP} [_{PP} P [_{DP} D_{Op} NP]]] C [_{IP} [_{VP} [_{PP} P [_{DP} D_{Op/dat} NP]] V]]]]



Tree 3

The same derivation applies if there is a non-matching external case. More interesting are cases with a PP both externally and internally. In the following example, there is case matching, but the prepositions differ, cf. (18), repeated here:

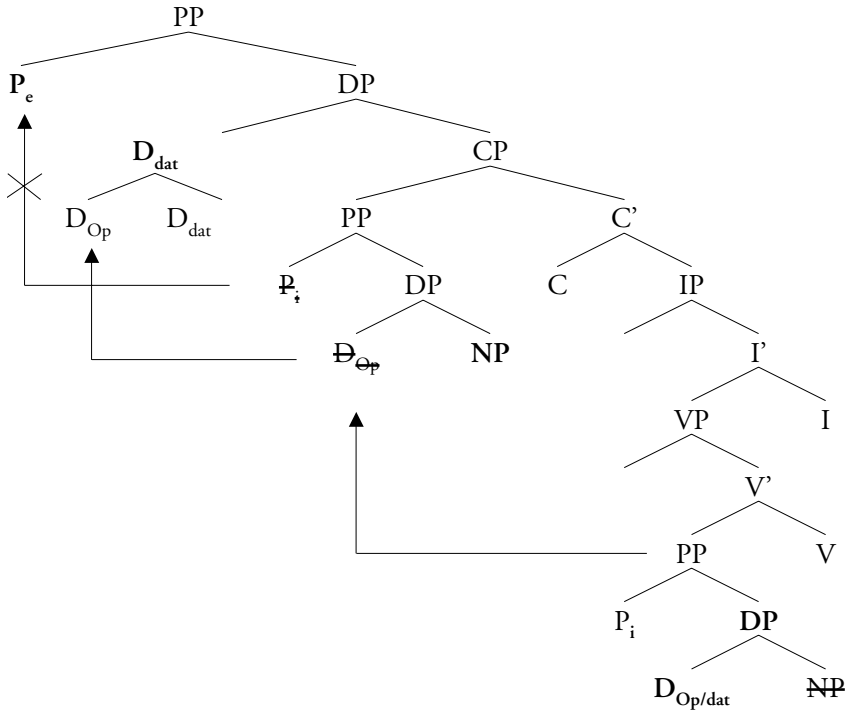
- (50) Ich ha vo de Lüüt, [won=i *(mit=ene) i
 I have:1SG from the:DAT people C=I (with=they:DAT) in
 d=Schuel bin], scho lang nüüt me ghöört.
 the=school am already long nothing anymore heard
 'I have not heard for a long time from the people with whom I went to school.'

Incorporation of the internal D does not lead to a case clash so that no resumptive is expected. However, P-incorporation leads to a clash in lexical features. Conse-

²⁵ Note that the notion of preposition stranding employed here is to be understood as a PF constraint, which rules out representations where P does not have a DP complement. If P were spelled out in Spec, CP, it would have an NP complement, but crucially, D would be empty. It might be possible to derive this restriction from the morphological licensing requirement on oblique case, assuming that P always assigns oblique case.

quently, P needs to be realized inside the relative clause.²⁶ This in turn requires case to be spelled out to prevent preposition stranding. Both P and D are realized in the base position due to a conspiracy of factors: Case can only be spelled out in the base position (due to the A'-split derivation) so that the preposition is spelled out there as well:

$$(51) [_{PP} P_e [_{DP} D_{Op} + D_{dat} [_{CP} [_{PP} P_i [_{DP} D_{Op} NP]] [_{IP} [_{VP} [_{PP} P_i [_{DP} D_{Op/dat} NP] V]]]]]]]$$



Tree 4

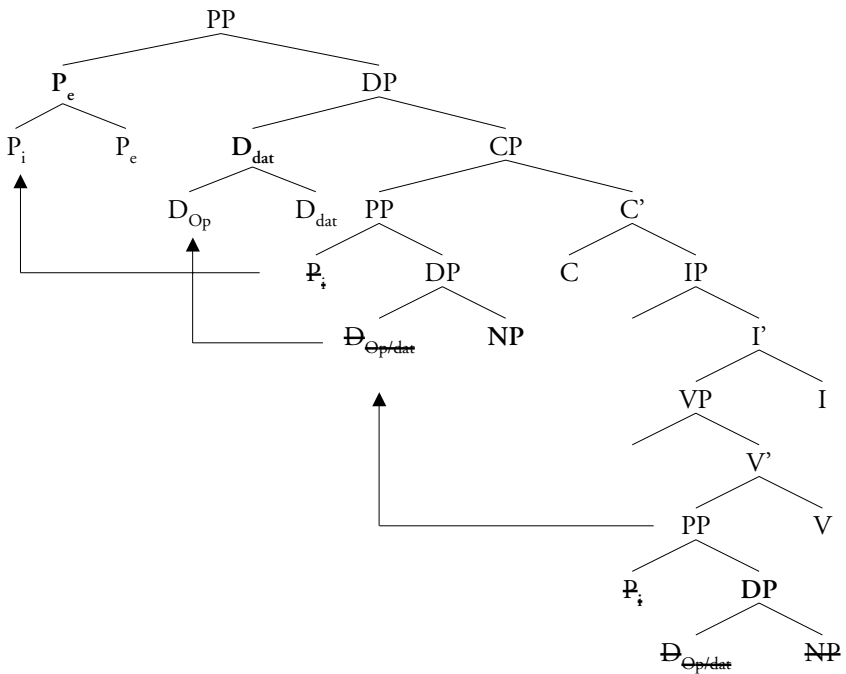
The same derivation applies to the configuration where both case and preposition are different, cf. (17) and the derivation in (19), where the prepositions are identical but the cases differ. In the latter case, P-incorporation and fusion fails because homophonous prepositions that can assign different cases differ in their lexical features. The last case to consider involves identical prepositions and case matching cf. (13)b, repeated here:

²⁶ For this account to work, I need to assume that P-incorporation (as opposed to D-incorporation) is optional.

- (52) Ich ha vo de Frau, [won=i scho geschter (*von=ere)
 I have:1sg from the:DAT woman C=I already yesterday (from=she:DAT)
 es Buech übercho han], wider äis übercho.
 a book received have:1SG again one received
 'I received from the woman from whom I had already received a book yester-
 day another one.'

Here both D and P incorporation result in compatible features. Neither D nor P needs to be spelled out inside the relative clause:

- (53) $[_{PP} P_i + P_e [_{DP} D_{Op/dat} + D_{dat} [_{CP} [_{PP} P_i [_{DP} D_{Op/dat} NP]]] C [_{IP} [_{VP} [_{PP} P_i [_{DP} D_{Op/dat} NP]] V]]]]]$



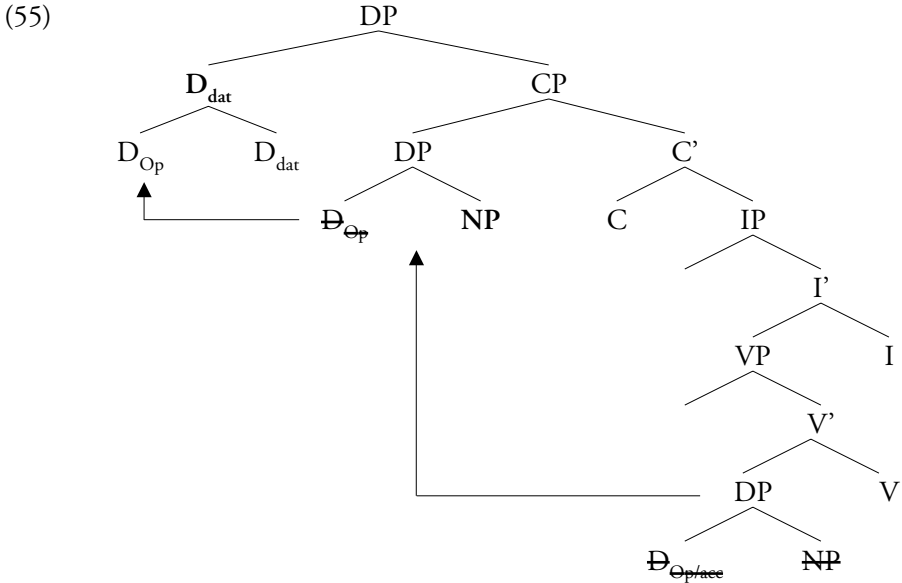
Tree 5

8.3. Subject and direct object

Relativization of subjects and direct objects with an oblique external D is problematic:

- (54) Ich ha de Frau, won=i geschter __ käne gleert ha,
 I have :1SG the:DAT woman C=I yesterday ACC know got have :1SG
 Blueme gschänkt.
 flowers given
 'I gave flowers to the woman who I met yesterday.'

Since there is a case clash, insertion fails and the only converging derivation would be of the A'-split type. However, we do not find resumptives, the stranded case feature does not have to be realized:



Tree 6

In the following section, I will discuss solutions to this and one further problem.

8.4. Conditions on PF chains, Spell-out, and recoverability

We are not in a position to formulate the licensing conditions for oblique case and prepositions on the one hand and for structural cases on the other. Cases normally have to be realized unless they are recoverable. Crucially, the conditions on recoverability are stricter for oblique cases and PPs.

Structural cases do not need morphological licensing, they are always structurally recoverable in relative clauses because a part of its (modified) chain receives case, namely the external D. This requirement overrules the spelling out of the stranded case feature in (55): Dative and PPs, however, always need morphological licensing. Crucially, oblique case must be unique within a given (modified) chain. It is either realized in the base position as a resumptive or on the head noun, which also forms part of the chain. This uniqueness condition is necessary to rule out a further possible derivation: Nothing so far ruled out applying Distributed Deletion under matching. At the point where the case feature is erased, the computational system does not yet know that eventually a matching configuration obtains. Preventing Distributed Deletion in this case would involve non-trivial look-ahead. Instead, Distributed Deletion is always an option. If it applies under matching, a representation results

where the (modified) chain contains two occurrences of oblique case/prepositions. The uniqueness condition on oblique case at PF rules out such a case.

8.5. The Importance of the surface forms

So far, I have presented no evidence in favor of a late insertion approach to the matching phenomena discussed in this paper. Example (23), repeated here, crucially showed that the surface form is crucial for matching:

- (56) \emptyset_D Mane, won=i (*ene) es Buech gib, müend intellektuell sii.
 D men(NOM) C =I (they:DAT) a book give:1s must:PL intellectual be
 'Men to whom I give a book must be intellectual.'

This follows under the approach advocated here: The fully specified internal D incorporates into the external D. This results in a case conflict: nominative vs. dative. One expects that insertion fails. However, there is an underspecified lexical item that can resolve this conflict: For indefinite plural, the empty determiner is inserted. It seems unreasonable to posit three homophonous (i.e. empty) determiners with different case specifications. Instead, it is much more plausible that there is only one, and it is underspecified for case. Insertion is subject to the specificity principle. Since the empty determiner is the most specific form available and since it does not conflict with the feature specifications, insertion is successful and the derivation converges. A similar reasoning applies to cases where there is a nominative-accusative clash. Since these cases are no longer morphologically different (except for certain pronouns), one can safely assume that most lexical items will be underspecified. If there is a nominative-accusative clash, insertion is still possible due to underspecified forms.²⁷

8.6. Overview over the matching configurations

The following table summarizes all configurations discussed in this paper:

	external case	internal spell-out
	nom/acc	nom/acc
= (1)c, (44)	nom/acc	dat
= (1)d, (37)	nom/acc/dat	P + nom/acc/dat
= (13)a, (46) <i>matching!</i>	dat	dat
= (54)	dat	nom/acc
= (18), (50)	P _a + dat	P _b + dat
= (15) <i>matching!</i>	P + dat	dat
= (13)b, (52) <i>matching!</i>	P _a + acc/dat	P_a + acc/dat
= (19)	P _a + acc/dat	P _a + dat/acc
= (17)	P _a + acc	P _b + dat

²⁷ Underspecification of nominative/accusative must not mean that they do not have any features whatsoever. They still need to be differentiated from dative case. One can assume that they are only specified for [- oblique], the second feature that distinguishes between nominative and accusative being neutralized.

9. Conclusion

The study of resumptives in ZG is very important because it reveals a new pattern of resumption that must be made available by UG. Resumptives in ZG are crucially linked to the licensing of oblique case and PPs. They occur to license oblique relations unless the case assigned by the external case of the head noun makes the very same morphological information available via a modified chain. These properties follow straightforwardly under the proposal advanced here: Restrictive relatives in ZG are derived via head raising and incorporation of relative clause internal material into relative clause external material. Matching is formalized as incorporation under identity while resumption is a result of Distributed Deletion.

References

- Adger, D. and G. Ramchand, 2004, *Merge and Move: Wh-Dependencies Revisited*. London/Oxford, ms.
- Aoun, J., L. Choueiri, and N. Hornstein, 2001, "Resumption, Movement, and Derivational Economy", *LI* 32. 371-403.
- Aoun, Y. and Y.-h. A. Li, 1993, "Wh-Elements in Situ: Syntax or LF?", *LI* 24. 199-238.
- , and —, 2003, *Essays on the representational and derivational Nature of Grammar: the Diversity of Wh-constructions*. MIT: MIT Press (= Linguistic Inquiry Monographs 40).
- Bayer, J., 2002, "Dative Pertinacity". Freudental: Handout of Workshop on Pertinacity.
- , M. Bader, and M. Meng, 2001, "Morphological Underspecification meets oblique Case: Syntactic and processing Effects in German", *Lingua* 111. 465-514.
- Bhatt, R., 2002, "The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification", *Natural Language Semantics* 40. 43-90.
- Bianchi, V., 1999. *Consequences of Antisymmetry*. Berlin/New York: Mouton de Gruyter.
- , 2000a, "Some issues in the Syntax of Relative Determiners". *The syntax of relative clauses*, ed. by A. Alexiadou, P. Law, A. Meinunger and C. Wilder, 53-81. Amsterdam/Philadelphia: John Benjamins (= Linguistics Today 32).
- , 2000b, "The Raising Analysis of Relative Clauses: A Reply to Borsley", *LI* 31. 123-40.
- Boeckx, C., 2003, *Islands and Chains. Resumption as Stranding*. Amsterdam/Philadelphia: John Benjamins.
- Brame, M. K., 1968, "A New Analysis of the Relative Clause: Evidence for an Interpretative Theory". MIT, ms.
- Carnie, A., H. Harley, and E. Pyatt, 2000, "VSO order as Raising to out of IP". *The Syntax of Verb Initial Languages*, ed. by A. Carnie and E. Guilloyle, 39-60. Oxford: OUP.
- Chao, W. and P. Sells, 1983, "On the Interpretation of Resumptive Pronouns", *NELS* 13. 47-61.
- Chomsky, N., 1977, "On Wh Movement". *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian, 71-132. New York: Academic Press.
- , 1995, *The Minimalist Program*. Cambridge: MIT Press.
- Cinque, G., 1990, *A'-Dependencies*. Cambridge: MIT Press (= LI Monographs 17).
- Comrie, B., and E. L. Keenan, 1977, "Noun Phrase Accessibility and Universal Grammar", *LI* 8. 63-99.

- Culicover, P. W., 2001, "Parasitic Gaps: A History". *Parasitic Gaps*, ed. by P. W. Culicover and P. M. Postal, 3-68. Cambridge: MIT Press (= Current Studies in Linguistics 35).
- Demirdache, H., 1991, *Resumptive Chains in Restrictive Relatives, Appositives and Dislocation Structures*. MIT: Ph.D. Dissertation.
- , 1997, "Dislocation, Resumption and Weakest Crossover". *Materials on Left-Dislocation*, ed. by E. Agnostopoulou, H. van Riemsdijk, and F. Zwart, 193-231. Amsterdam/Philadelphia: John Benjamins (= Linguistics Today 14).
- Fanselow, G., and D. Cavar, 2002, "Distributed Deletion". *Theoretical Approaches to Universals*, ed. by A. Alexiadou, 65-107. Amsterdam/Philadelphia: John Benjamins (= Linguistics Today 49).
- Fleischer, J., 2001, "Preposition Stranding in German Dialects". *Syntactic Microvariation*, ed. by S. Barbiers, Leonie C., and S. van der Kleij, 116-151. Amsterdam: Meertens Instituut.
- Gallmann, P., 1992, "Dativanhebung?", *Groninger Arbeiten zur Germanistischen Linguistik* 35. 92-122.
- Grohmann, K. K., 2003, *Prolific Domains. On the Anti-Locality of Movement Dependencies*. Amsterdam/Philadelphia: John Benjamins (= Linguistics Today 66).
- Groos, A. and van Riemsdijk, H., 1981, "Matching Effects in Free Relatives: A Parameter of Core Grammar". *Theory of Markedness in Generative Grammar. Proceedings of the 1979 GLOW Conference*, ed. by A. Belletti, L. Brandi, and L. Rizzi, 171-216. Pisa: Scuola Normale Superiore.
- Halle, M., and Marantz, A., 1993, "Distributed Morphology and the Pieces of Inflection". *The View from building 20: essays in linguistics in honor of Sylvain Bromberger*, ed. by K. Hale and S. J. Keyser, 111-176. Cambridge: MIT Press (= Current Studies in Linguistics 24).
- Kayne, R. S., 1994, *The Antisymmetry of Syntax*. Cambridge/London: MIT Press (= Linguistic Inquiry Monographs 25).
- Longobardi, G., 1994, "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form", *LI* 25. 609-65.
- McCloskey, J., 1990, "Resumptive Pronouns, A'-Binding, and Levels of Representation in Irish". *The Syntax of modern Celtic Languages*, ed. by R. Hendrick, 199-248. New York: Academic Press.
- Müller, G., 1999, "Imperfect Checking", *The Linguistic Review* 16. 359-404.
- Pesetsky, D., 1998, *Some Optimality Principles of Sentence Pronunciation*. In *Optimality and Competition in Syntax*, ed. by P. Barbosa, D. Fox, M. McGinnis, and D. Pesetsky, 337-383. Cambridge: MIT Press.
- Riemsdijk, H. van, 1989, "Swiss Relatives. Sentential Complementation and the Lexicon". *Studies in Honour of Wim de Geest*, ed. by D. Jaspers, W. Klooster, Y. Putseys, and P. Seuren, 343-354. Dordrecht/Providence: Foris (= Linguistic Models 13).
- , 2003, "Some Remarks on Aboutness Relatives". Handout from *NWO-OTKA Meeting*, Collegium Budapest, December 5/6.
- Rouveret, A., 2002, "How are resumptive Pronouns linked to the Periphery?", *Linguistic Variation Yearbook*. Volume 2, ed. by P. Pica and J. Rooryck, 123-184. Amsterdam/Philadelphia: John Benjamins.
- Safir, K., 1999, "Vehicle Change and Reconstruction in A'-Chains", *LI* 30. 587-620.
- Salzmann, M., in prep., "Resumption as Distributed Deletion. Resumptive pronouns in Zurich German".

- Sauerland, U., 1998, *The Meaning of Chains*. MIT: Ph.D. dissertation.
- Schachter, P., 1973, "Focus and Relativization", *Lg* 49. 19-46.
- Seiler, G., 2001, "Prepositional Dative Marking in Upper German: A Case of Syntactic Microvariation". *Syntactic Microvariation*, ed. by S. Barbiers, L. Cornips, and S. van der Kleij, 243-279. Amsterdam: Meertens Instituut.
- Sharvit, Y., 1999, "Resumptive Pronouns in Relative Clauses", *NLLT* 17. 587-612.
- Shlonsky, U., 1992, "Resumptive Pronouns as a Last Resort", *LI* 23. 443-68.
- Suñer, M., 1998, "Resumptive Restrictive Relatives: A crosslinguistic Perspective", *Lg* 74. 335-64.
- Travis, L. deMena, 1984, *Parameters and Effects of Word Order Variation*. MIT: PhD dissertation.
- Vergnaud, J.-R., 1974, *French Relative Clauses*. MIT: PhD dissertation.
- Vogel, R. and M. Steinbach, 1998, "The Dative — an Oblique Case", *Linguistische Berichte* 173. 65-90.
- Vries, M. de, 2002, *The Syntax of Relativization*. Utrecht: LOT (= LOT Dissertation Series 53).
- Wegener, H., 1985, *Der Dativ im heutigen Deutsch*. Tübingen: Narr.
- , 1991, "Der Dativ — ein struktureller Kasus?". In *Strukturen und Merkmale syntaktischer Kategorien*, ed. by G. Fanselow and S. Felix, 77-107. Tübingen: Gunter Narr.

THE COMPILATION OF A SAMPLE PFR CHINESE CORPUS OF SKELETON-PARSED SENTENCES

May Lai-Yin Wong

Department of Linguistics and Modern English Grammar (Lancaster University)

Abstract

*The approach taken in this paper for the construction of a treebank is inspired by the skeleton parsing approach. From the PFR Chinese Corpus, a sample text of some 100,000 word tokens was chosen for the production of the treebank. A clear account of the 17 non terminal constituents that are defined and instantiated in the corpus texts will be provided in a parsing scheme. A set of parsing guidelines on practical issues related to map any parses on to sentences in the application of the parsing scheme will also be considered. It is noteworthy also to discuss the major difficulties encountered in the course of skeleton parsing, as this illuminates some of the peculiarities of the Chinese language. The conclusion is an evaluation of the success of the treebank compilation.**

1. Introduction

Treebanks are simply corpora in which syntactic constituent structure is made explicit by a process of corpus annotation (Leech and Garside 1991: 15; Abeillé 2003: xiv). My major concern here is not with software to achieve this annotation automatically (as at the time of writing, there are no effective available parsers designed for the Chinese language), but with the establishment of a parsing scheme and its manual application to written Chinese corpus data. More specifically, the approach taken here is inspired by the skeleton parsing approach (Eyes and Leech 1993; Garside 1993; Black et al. 1996; Leech and Eyes 1997). Skeleton parsing seeks to produce simplified constituent-structure annotations.

2. PFR sample skeleton treebank: text selection

From the PFR Chinese Corpus (Yu 1999), a sample text of some 100,000 word tokens, yielding approximately 2,500 sentences was chosen for the production of my treebank. I contend that a standard block of about one hundred thousand

* I am indebted to the audience in the Bilbao-Deusto Student Conference 2004, held in Bilbao, Spain, from July 8 to 10 for their insightful comments. All the errors are, of course, mine.

words is a unit of about the right size for skeleton parsing and handling by a human treebanker. Furthermore, my choice of text was governed by the need to produce a parsed sample corpus of reasonable length which could not only be manageable for hand-parsing but also represent a typical chunk of the data that would eventually be treebanked by me.

3. PFR sample skeleton treebank: parsing scheme

As Sampson (1995: 2ff) puts it, the process of parsing refers to the ability to extract from a linear sequence of words the underlying hierarchical grammatical structure, and thus a parsing scheme “is a set of categories and notational conventions allowing the grammatical properties of a text to be made explicit”. In other words, it is a guideline document which helps the human analyst parse sentences (Leech and Garside 1991: 15-16). A clearly defined parsing scheme is essential for the production of a satisfactorily parsed text.

3.1. UCREL skeleton parsing annotation scheme

As most of the existing treebanks are primarily based upon English texts, it does not come as a surprise that the annotation schemes used on those treebanks chiefly reflect the syntactic categories which are directly relevant to English grammar. A case in point is the UCREL skeleton parsing scheme, as illustrated in Table 1.¹

TABLE 1: The UCREL skeleton parsing annotation scheme

UCREL Skeleton Parsing Annotation Scheme	
Fa	Adverbial Clause
Fc	Comparative Clause
Fn	Noun Clause
Fr	Relative Clause
G	Genitive
J	Adjective Phrase (predicative)
N	Noun Phrase
Nr	Adverbial Noun Phrase (temporal)
Nv	Adverbial Noun Phrase (non-temporal) (not in AP or SEC corpora)
P	Prepositional Phrase
S	Sentence (used eg in quoted speech, also with + and & as co-ordinates)
Tg	<i>-ing</i> Clause
Ti	Infinitive Clause
Tn	Past Participle Clause
V	Verb Phrase
(null)	Unlabelled Constituent

¹ The table was adapted from UCREL's website <http://www.comp.lancs.ac.uk/computing/research/ucrel/skeletontags.html>.

As stated, some of the syntactic constituent labels in the UCREL skeleton parsing scheme are specially designed to suit English grammar. When I attempted to adopt this scheme wholesale for my own research, I found that some modifications were needed to accommodate the syntactic properties of the Chinese language. Since there is no conclusive morphological evidence that motivates the postulation of infinitival clauses in Chinese (Xue et al. 2000: 32), the non-finite clauses, including the *-ing* clause (Tg), infinitive clause (Ti) and past participle clause (Tn) were wnot taken over from the UCREL parsing scheme to my parsing scheme. Similarly, I did not include noun clause (Fn)² and relative clause (Fr)³ as they are also not compatible with Chinese syntax. Furthermore, owing to the fact that different languages tend to employ different strategies in signalling the same grammatical relations, the parsing label of comparative clause (Fc) was also not adopted in my parsing scheme. While English makes use of a clause to give comparisons, the comparative constructions in Chinese are expressed by means of a prepositional phrase.⁴ Lastly, I did not take the genitive (G) as one of the constituent labels in my parsing scheme so as to avoid terminological controversy.⁵

In view of the differences between the English and Chinese grammatical systems, new constituent labels that are not used in the UCREL skeleton parsing scheme had to be invented for the purposes of this research. These are: adverb phrase, correlative clause, adverbial idiom/set phrase, adverbial adjective phrase, adverbial prepositional phrase, adverbial verb phrase and verbal object.

² Nominal clauses are also difficult in the Chinese linguistics. The notion of the nominal clause may be useful in English as the clauses used as sentential subject or object are formally distinctive from those used independently: they are typically introduced by the conjunction *that* as in “That John stole my book was totally incredible” (Quirk and Greenbaum 1973: 316-322). Chinese clauses or sentences, however, do not vary morphosyntactically when they are used as subject or direct object (Liu 1996: 245, 253). Hence, there is no justification for adopting this grammatical category in a parsing scheme designed for the Chinese language. The lack of previous accounts of nominal clauses in Chinese tends to serve as counterevidence to the notion that nominal clauses exist in Chinese.

³ Relative clauses have long been a source of controversy in Chinese linguistics. Some scholars (e.g. Li and Thompson 1989: 579ff; Aoun and Li 1993; Chiu 1993; Wu 2000) believe that a nominalisation (whereby a verb, verb phrase, or sentence, followed by the particle *de*, functions as a noun phrase) can be called a relative clause in Chinese if the head noun that it modifies refers to an unspecified element involved in the situation described by the nominalisation. On the other hand, Chao (1968), among others, does not adopt the notion of relative clause in his descriptive grammar of Chinese.

⁴ The comparative marker 比 *bi* “than” and the phrase that immediately follows it form a prepositional phrase which serves as a preverbal adjunct (Zhao 1989; Liu 1999: 204ff). Compare the following contrived sentences from English and Chinese respectively, both of which express the same meaning:

(a) He does the assignment [Fc better than I do Fc].

(b) 他做的功课做得[P 比我 P]好。 *Ta zuo de gongke zuo de bi wo hao*

⁵ *The genitive constructions in English roughly* correspond to those constructions marked by the particle 的 *de* in Chinese. However, this apparent correspondence is complicated by two issues. Firstly, possessive constructions in Chinese do not necessarily take the particle 的 *de* (Li and Thompson 1989: 115) as in 他爸爸 *ta baba* and 他的爸爸 *ta de baba* which both indicate the same meaning “his father”. Secondly, apart from marking possessions, the particle 的 *de* can be an adjectival marker and a marker of nominalisation and explicit modification (Zhu 1982 and 2000). Since the term “genitive” is confusing as far as Chinese linguistics is concerned, I decided to exclude this constituent from my parsing scheme.

3.2. PFR skeleton parsing labels

As Leech and Eyes (1997: 37) note, Sampson's (1995) annotation scheme demonstrates three key components of a clearly specified parsing scheme:

- A list of symbols used in the annotation: non-terminals, terminals, and other symbols;
- A basic definition of the symbols: e.g. N = noun phrase;
- A description, which is as detailed as possible, of how the symbols are actually applied to text sentences. For example, how do annotators recognise a noun phrase when they see one, and how do they distinguish noun phrase tokens from words or word sequences which are not noun phrases?

With respect to the first and second points, in keeping with Sampson, I intend to provide these 3 sets of data for my treebank: (a) the non-terminal labels and (b) their definitions with illustrative examples are given for the PFR treebank in Table 2.⁶

TABLE 2: The list of constituent labels for the PFR Sample Skeleton Treebank parsing scheme

Nonterminal Category	Symbol	Example
Adverbial Clause	Fa	<Fa>只要_c <N>我们_r</N> <R>进一步_d</R> <解放思想_i , _w 实事求是_i> , _w <V>抓住_v <N>机遇_n</N></V> , _w <开拓进取_l></Fa> , _w <Fa>zhiyao_c <N>women_r</N> <R>jinyibu_d</R> <jiefangsixiang_i , _w shishiqiushi_i> , _w <V>zhuazhu_v <N>jiyu_n</N></V> , _w <kaituojinqu_l></Fa> , _w "If we become more open-minded and down-to-earth and make every effort to explore new possibilities, ..."
Correlative Clause	Fc	<Fc>越_d 走_v 越_d 宽广_a</Fc> <Fc>yue_d zou_v yue_d kuanguang_a</Fc> "the more we walk, the broader (the road) will be"
Main Clause (to which the adverbial clause is subordinated)	Fm	<Fm><Fa>只要_c <N>我们_r</N> <R>进一步_d</R> <解放思想_i , _w 实事求是_i> , _w <V>抓住_v <N>机遇_n</N></V> , _w <开拓进取_l></Fa> , _w <N><V>建设_v <V>有_v <N>中国_ns 特色_n 社会主义_n</N></V></V> 的_u 道路_n</N> 就_c <V>会_v <Fc>越_d 走_v 越_d 宽广_a</Fc></V> 。 _w</Fm> <Fm><Fa>zhiyao_c <N>women_r</N> <R>jinyibu_d</R> <jiefangsixiang_i , _w shishiqiushi_i> , _w <V>zhuazhu_v <N>jiyu_n</N></V> , _w <kaituojinqu_l></Fa> , _w <N><V>jianshe_v <V>you_v <N>Zhongguo_ns tese_n shehuizhuyi_n</N></V></V> de_u daolu_n</N> jiu_c <V>hui_v <Fc>yue_d zou_v yue_d kuanguang_a</Fc></V> 。 _w</Fm>

⁶ All of the examples in this paper are given in the Chinese characters, followed by pinyin romanisations and English translations.

Nonterminal Category	Symbol	Example
		“If we become more open-minded and down-to-earth and make every effort to explore new possibilities, we are in a better position to set up an ideology which can fully represent Chinese characteristics.”
Adverbial Idiom/ Set Phrase	Ia	<Ia>坚定不移_地_u</Ia> <Ia> <i>jiandingbuyi_i de_u</i> </Ia> “persistently” <Ia>满怀信心_地_u</Ia> <Ia> <i>manhuaixinxin_i de_u</i> </Ia> “confidently”
Adjective Phrase	J	<J>非常_d 重要_a</J> <J> <i>feichang_d zhongyao_a</i> </J> “very important”
Adverbial Adjective Phrase	Ja	<Ja>成功_a 地_u</Ja> <Ja> <i>chenggong_a de_u</i> </Ja> “successfully”
Noun Phrase	N	<N>百年_m 历史_n</N> <N> <i>bainian_m lishi_n</i> </N> “a hundred years’ history”
Adverbial Noun Phrase	Na	<Na>今天_t 上午_t</Na> <Na> <i>jintian_t shangwu_t</i> </Na> “this morning” <Na>现在_t</Na> <Na> <i>xianzai_t</i> </Na> “at present”
Prepositional Phrase	P	<P>对_p <N>香港_ns</N></P> <P> <i>dui_p <N>Xianggang_ns</i> </N></P> “to Hong Kong”
Adverbial Prepositional Phrase	Pa	<Pa>在_p <N>这_r 一_m 年_q</N></Pa> <Pa> <i>zai_p <N>zhe_r yi_m nian_q</i> </N></Pa> “in this year”
Adverb Phrase	R	<R>还_d 不_d</R> <R> <i>hai_d bu_d</i> </R> “not...though”
Sentence (including direct speech quotation, also with & and + as co-ordinates)	S	<S N=>1>><Pa>在_p <N>这_r 一_m 年_q 中</N>_f</Pa> , _w <N>中国_ns 的_u 外交_n 工作_vn</N> <V>取得_v 了_u <N>重要_a 成果_n</N></V> °_w</S> <S N=”1”>><Pa> <i>zai_p <N>zhe_r yi_m nian_q zhong</N>_f</i> </Pa> , _w <N> <i>Zhongguo_ns de_u waijiao_n gongzuo_vn</i> </N> <V> <i>qude_v le_u <N>zhongyao_a chengguo_n</i> </N></V> °_w</S> “In this year, the Chinese government has gained great success in its diplomacy.”

Nonterminal Category	Symbol	Example
Verb Phrase	V	<V>展望_v <N>新_a的_u 世纪_n</N></V> <V>zhanwang_v <N>xin_a de_u sheji_n</N></V> “have hope in the new era”
Adverbial Verb Phrase	Va	<Va>迈向_v <N><充满_v 希望_n的_u> 1 9 9 8年_t 之际_f</N></Va> <Va>maixiang_v <N><chongman_v xiwang_n de_u> 1 9 9 8 nian_t zhiji_f</N></Va> “at the moment we are looking forward to the prosperous year of 1998” <Va>退休_v 前_f</Va> <Va>tuixiu_v qian_f</Va> “before retirement”
Verbal Object	Vo	<V>希望_v <Vo>依靠_v <N>大家_r</N></Vo></V> <V>xiwang_v <Vo>yikao_v <N>dajia_r</N></Vo></V> “wish to rely on you”
Initial Conjunct	&	<N&>中国_ns 改革_v</N&> 和_c <N&>Zhongguo_ns gaige_v</N&> he_c “China’s revolution”
Non-initial Conjunct	+	和_c <N+>发展_v的_u 全局_n</N+> he_c <N+>fazhan_v de_u quanju_n</N+> “and the entire development”

4. Guidelines of skeleton parsing

Having reviewed my annotation scheme in some depth, I can now present my guidelines for annotation, in keeping with my desire to match the advantages of Sampson’s SUSANNE scheme (see section 4.2). The parsing scheme matches features (a) and (b) of a clear and explicit parsing scheme (Leech and Eyes 1997: 37). Nonetheless, an annotation scheme is more than (a) and (b) above. Feature (c), a set of parsing guidelines should also be provided in order to explain how the parsing symbols are actually applied to text sentences to avoid undesirable inconsistency.

During the course of annotation, as more data was analysed, the guidelines took shape and were recorded and updated. Hence, after I finished the task of parsing a sample text taken from the PFR Chinese Corpus, I had a set of guidelines ready to be consolidated into a document to be available to users of the treebank and to future annotators who might want to adopt the same scheme. It is advisable, as Kahrel et al. (1997: 241ff) note, to document explicitly all of the decisions taken in the development of an annotation scheme, as well as its application so that future users can apply the scheme in a manner consistent with that of the originators of the scheme. The decisions were then adhered to consistently in the annotation of similar cases thereafter. In the following subsections, I will discuss the issues that arose and illustrate them with examples drawn from the sample treebank.

4.1. Underspecification — Use of unlabelled bracketings

Brackets may be left unlabelled in cases where a particular grouped sequence of words cannot fit into any of the existing phrase or clause categories. Examples of constituents enclosed in unlabelled brackets are given below from (a) to (e).

- (a) Multi-word premodifiers of noun phrases marked by the particle 的 *de* (see also section 5.2.3):

e.g. <N><全国_n 各族_r 的_u> 人民_n</N> <N><*quanguo_n gezu_r de_u renmin_n*</N> “people from different ethnic groups throughout the country”;

e.g. <N><同_p 联合国_nt 和_c 其他_r 国际_n 组织_n 的_u> 协调_vn</N> <N><*tong_p Lianheguo_nt he_c qita_r guoji_n zuzhi_n de_u xietiao_vn*</N> “the compromise between the United Nations and other international organisations”;

e.g. <N><最近_t 一个_m 时期_n 一些_m 国家_n 和_c 地区_n 发生_v 的_u> 金融_n 风波_n</N> <N><*zuijin_t yige_m shiqi_n yixie_m guojia_n he_c diqu_n fasheng_v de_u jinyong_n fengbo_n*</N> “the recent financial crises happened in some countries and districts”;

- (b) Serial verb constructions which are used as if they were compound verbs (see also section 5.2.4):

e.g. <坚持_v 奉行_v> <*jianchi_v fengxing_v*> “insist on following”;

e.g. <指挥_v 演奏_v> <*zhibui_v yanzou_v*> “lead and perform”;

e.g. <看望_v 慰问_v> <*kanwang_v weiwen_v*> “visit and send regards to ...”.

- (c) Serial adjective constructions:

e.g. <团结_a 一致_a> <*tuanjie_a yizhi_a*> “be united together”;

e.g. <圆满_a 成功_a> <*yuanman_a chenggong_a*> “perfectly successful”.

- (d) Idioms/set phrases which are used idiosyncratically as if they were single-word nouns or verbs (see also section 5.2.2):

e.g. <大势所趋_i , _w 民心所向_l> <*dashisuoqu_i , _w minxinsuoxiang_l*> “urged by the trend, supported by general public”;

e.g. <大气磅礴_i , _w 波澜壮阔_i> <*diqibangbo_i , _w bolanzhuangkuo_i*> “powerful wind, fierce waves”;

e.g. <流光溢彩_l , _w 火树银花_i> <*liuguangyicai_l , _w huoshuyinhua_i*> “filled with colourful lights, magnificent”.

- (e) Coordinated verbs with shared direct object:

e.g. <V><学习_v 和_c 掌握_v> <N>党_n 的_u 十五大_j 精神_n</N></V> <V><*xuexi_v he_c zhangwo_v*> <N>*dang_n de_u shiwuda_j jingshen_n*</N></V> “learn and master the Communist Party’s 15 principles”;

e.g. <V><尊重_v 、 _w 认识_v 和_c 掌握_v> <N>客观_a 规律_n</N></V> <V><*zhuanzhong_v 、 _w renshi_v he_c zhangwo_v*> <N>*keguan_a guilü_n*</N></V> “respect, understand and master what we learn in our daily life”.

4.2. Bracketing of multi-word constituents

The unlabelled bracketing facility evidently has its uses in skeleton parsing as it allows analysis to proceed where labelling decisions are not obvious or straightforward. Nevertheless, for some multi-word adverb phrases containing two adverbs (e.g. <R>还_d 不_d</R> <R>hai_d bu_d</R> “not...though”; <R>永远_d 不再_d</R> <R>yongyuan_d buzai_d</R> “never forever”; <R>一直_d 都_d</R> <R>yizhi_d dou_d</R> “constantly”), and multi-word attributive adjectival phrases containing an adjective premodified by at least one adverb (e.g. <J>非常_d 重要_a 的_u</J> <J>feichang_d zhongyao_a de_u</J> “very important”; <J>很_d 不_d 平凡_a 的_u</J> <J>hen_d bu_d pingfan_a de_u</J> “very extraordinary; <J>十分_m 高兴_a</J> <J>shifen_m gaoxing_a</J> “very happy”), though Eyes and Leech (1993: 53) chose to put them into unlabelled brackets, they were labelled in my treebank. The reason for this is that their internal structure is clear, having a head (adjective or adverb) being modified by another adverb.

4.3. Bracketing of single-word constituents

As suggested in the EAGLES Recommendations for the Syntactic Annotation of Corpora, Version of 11th March 1996 (Leech et al. 1996), it is considered preferable to bracket single-word constituents where they show their phrasal status by the possibility of adding modifiers or replacing them by a multi-word phrase as in example 1, or where they are in coordination with other multi-word constituents as in example 2.

- (1) <N>人民_n 生活_vn</N> <R>进一步_d</R> <V>改善_v</V>
 <N>renmin_n shenghuo_vn</N> <R>jinyibu_d</R> <V>gaishan_v</V>
 “the life of the citizens is further improved”
- (2) <N><N&>全党_n</N&> 和_c <N+>全国_n 各族_r 人民_n</N+></N>
 <N><N&>quandang_n</N&> he_c <N+>quanguo_n gezu_r renmin_n</N+></N>
 “the Communist Party and the citizens of varied ethnic groups throughout the country”

4.4. Punctuation

Generally speaking, I included punctuation within the bracketing. As for phrase/sentence-initial and phrase/sentence-final punctuations, I enclosed them within the parsing bracketing, as in example 3:

- (3) <P>为_p <N>“_w 两手抓_l 、_w 两手_m 都_d 要_v 硬_a ”_w</N></P>
 <P>wei_p <N>“_w liangshouzhua_l 、_w liangshou_m dou_d yao_v ying_a ”_w</N></P>
 <V>提供_v 了_u <N>新_a 的_u 理论_n 根据_n</N></V>
 <V>tigong_v le_u <N>xin_a de_u lilun_n genju_n</N></V>
 “provide new theoretical evidence to the principle of perseverance in an undertaking”

As regards medial punctuation marks, typically commas, I attached them to the highest available node in the parse tree, thus these punctuation marks can be used as delimiters of major constituents, as in example 4:

- (4) <S N="5"><S&><N>[中国_ns 政府_n]nt</N> <Ja>顺利_ad</Ja>
 <V>恢复_v <P>对_p <N>香港_ns</N></P> <Vo>行使_v <N>主权_
 n</N></Vo></V></S&> , _w 并_c <S+><P>按照_p <N><“_w 一国
 两制_j”_w 、_w “_w 港人治港_l”_w 、_w 高度_d 自治_v 的_u> 方
 针_n</N></P> <V>保持_v <N>香港_ns 的_u 繁荣_an 稳定_an</N></
 V></S+> 。_w</S>
 <S N="5"><S&><N>[Zhongguo_ns zhengfu_n]nt</N> <Ja>shunli_ad</
 Ja> <V>hui fu_v <P>dui_p <N>Xianggang_ns</N></P> <Vo>xingshi_v
 <N>zhuquan_n</N></Vo></V></S&> , _w bing_c <S+><P>anzhao_p
 <N><“_w yiguoliangzhi_j”_w 、_w “_w gangrenzhigang_l”_w 、_w gaodu_
 d zizhi_v de_u fangzhen_n</N></P> <V>baochi_v <N>Xianggang_ns de_u
 fanrong_an wending_an</N></V></S+> 。_w</S>
 “The Chinese government has succeeded in resuming its sovereignty over
 Hong Kong and maintaining the prosperity and stability of Hong Kong,
 according to the objectives on ‘one country two systems’, ‘Hong Kong
 people ruling themselves’ and ‘high degree of independence’.”

4.5. Ambiguity

Linguistic forms are often ambiguous. My annotation scheme, however, did not contain any notation for representing ambiguity explicitly with which the human analyst selects one possible sense for a form and represents it. I decided not to explicitly mark an ambiguous form because even if a given item has more than one reading, the human analyst will not recognise this in the course of parsing and just annotate the item with the interpretation that seems initially most plausible.

With such a detailed and carefully articulated parsing scheme and guidelines, I can now proceed to describe the actual process of skeleton parsing and difficulties encountered in the process of parsing.

5. The process of skeleton parsing

5.1. The basic concept of skeleton parsing

The basic idea of skeleton parsing, as Garside and McEnery (1993: 19) demonstrate, is that the treebanker marks only those syntactic structures which seem “intuitively obvious”, rather than keeping track of a particular reference grammar. In the course of skeleton parsing, I inserted a nested set of brackets around a sequence of word tokens which appeared to be intuitively correct to group as a single unit. I then assigned to each of these units (i.e. sentence constituents) a label from the set of categories specified in my parsing scheme. An excerpt of the PFR Skeleton Treebank is provided in Figure 1.

<ORIGINALMP>19980101-01-001-013_m</ORIGINALMP> <S N=" 28" ><V>环顾_v <N>全球_n</N></V> ,_w <N><日益_d 密切_a 的_u 世界_n 经济_n 联系_vn</N> ,_w <N>日新月异_j 的_u 科技_n 进步_vn</N> ,_w <R>正在_d</R><P>为_p <N><各国_r 经济_n 的_u 发展_vn</N></P> <V>提供_v <N>历史_n 机遇_n</N></V> 。_w</S> <S N=" 29" >但是_c ,_w <N>世界_n</N> <R>还_d 不_d</R> </>安宁_a</I> 。_w</S> <S N=" 30" ><N><南北_n 之间_f 的_u 贫富_n 差距_n</N> <V>继续_v <Vo>扩大_v</Vo></V> ;_w <N>局部_b 冲突_vn</N> <时>有发生_l> ;_w <N><<不_d 公正_a 不_d 合理_a 的_u 旧_a 的_u 国际_n 政治_n 经济_n 秩序_n</N> <R>还_d</R> <V>没有_v <N>根本_a 改变_vn</N></V> ;_w <N>发展中国家_l</N> <P>在_p <N>激烈_a 的_u 国际_n 经济_n 竞争_vn 中_f</N></P> <R>仍_d</R> <V>处于_v <N>弱势_n 地位_n</N></V> ;_w <N>人类_n 的_u <N><N&&生存_vn</N&& 与_c <N+>发展_vn</N+></N></N> <R>还_d</R> <V>面临_v <N>种种_q <N><N&&威胁_vn</N&& 和_c <N+>挑战_vn</N+></N></N></V> 。_w</S> <S N=" 31" ><N><和平_n 与_c 发展_vn 的_u 前景_n</N> <V>是_v </>光明_a 的_u</I></V> ,_w <N>21_m 世纪_n</N> <R>将_d</R> <V>是_v <N><充满_v 希望_n 的_u 世纪_n</N></V> 。_w</S> <S N=" 32" >但_c <N>前进_v 的_u 道路_n</N> <V><V><V&&不会_v</V&& <R>也_d</R> <V+><R>不_d</R> 可能_v</V+></V> <一帆风顺_j></I> ,_w <N>关键_n</N> <V>是_v <S><N>世界_n 各国_r 人民_n</N> <V>要_v </>进一步_d 团结_a 起来_v</I></V></S></V> ,_w <V><V><V&&共同_d</R> 推动_v</V&& <V+><R>早日_d</R> 建立_v</V+></V> <N><公正_a 合理_a 的_u 国际_n 政治_n 经济_n 新_a 秩序_n</N></V> 。_w</S>

FIGURE 1: An excerpt of the PFR Sample Skeleton Treebank

5.2. Difficulties in skeleton parsing Chinese text

It is noteworthy here to discuss the major difficulties that I encountered in the course of skeleton parsing a sample text taken from my corpus, as this illuminates some of the peculiarities of the Chinese language.

5.2.1. *Ba* constructions

Firstly, 把 *ba* constructions make the parse of a verb phrase incomplete. The *ba* construction is a widely discussed topic in the grammar of Chinese (see, for instance, Li and Thompson 1989: 463-491; Chen 1990; Kit 1992; Zou 1993; Xia and Wu 1996; Li 1997; Xue et al. 2000; Li 2001). In general, the structure of the *ba* construction is expressed and underlined in example 5: a *ba* sentence has a subject, followed by *ba* and the *ba* noun phrase (i.e. the NP directly following *ba*) followed by a verb.

(5) *subject ba NP verb*

In my PFR treebank, 104 instances of the *ba* construction were found, which fall into three types of this construction. The general pattern of *ba* sentences is to place the direct object of the following verb immediately after *ba* as in example 6.

(6) <P>把_p <N_i>电厂_n</N_i></P> <V>建_v 好_a</V> Ø_i <P>ba_p <N_i>dianchang_n</N_i></P> <V>jian_v hao_a</V> Ø_i “build an electricity supply station”

Typically, a transitive verb should follow the pattern “<V>...<N>...</N></V>”. However, the verb, 建_v 好_a, *jian-hao*, “build well, build in good shape” that

follows the *ba* construction lacks a direct object, which is equivalent to the *ba* noun phrase, i.e. *ba* NP_i verb Ø_i, where Ø_i denotes the empty position of the preposed verb object that shares the same reference as the prepositional complement of *ba*.⁷ In other words, the structure of the verb phrase following the *ba* construction is in the form of <V>...</V> rather than the canonical form mentioned before.

More complicated *ba* constructions involve the occurrence of two following verbs and a passivised verb. In cases like example 7 where two different transitive verbs follow the *ba* construction, it is not immediately obvious whether that the prepositional complement of *ba* co-refers to the object of the first verb (改编 *gaibian* “be adapted for”) or that of the second one (为 *wei* “be changed as”). Since the second verb already takes a direct object (器乐曲 *qiyuequ* “acoustics of musical instrument”), the *ba* noun phrase must be co-referential with the object of the first verb.

- (7) <P>把_p <N_i><广大_b 听众_n 耳熟能详_i 的_u> 歌曲_n</N_i></P>
 <V>改编_v Ø_i<V>为_v <N>器乐曲_n</N></V></V>
 <P>*ba*_p <N_i><*guangda*_b *tingzhong*_n *ershunengxiang*_i 的_u> *gequ*_n</N_i></P>
 <V>*gaibian*_v Ø_i<V>*wei*_v <N>*qiyuequ*_n</N></V></V>
 “change those popular songs into acoustics of musical instrument”

In less obvious cases like example 8, however, it is impossible to locate any empty position that co-refers to the *ba* complement. The verb 带入 *dairu* “bring to” that follows the *ba* construction is used causatively without any visible passivisation. As Norman (1988: 164) notes, Chinese verbs do not make any distinction between the active (or unaccusative) and passive (or causative). The *ba* noun phrase 人们 *renmen* “people” therefore actually refers to the logical subject of the verb.

- (8) <P>把_p <N>人们_n</N></P> <V>带入_v <N>迷人_a 的_u 艺术_n 境地_n</N></V>
 <P>*ba*_p <N>*renmen*_n</N></P> <V>*dairu*_v <N>*miren*_a *de*_u *yishu*_n *jingdi*_n</N></V>
 “bring people into a fascinating imaginary place”

5.2.2. Idioms or set phrases

The use of idioms (tagged “i”) or set phrases (tagged “l”) as if they were nouns and verbs is also problematic. Noun-like idioms and set phrases are illustrated in example 9 and verb-like set phrases in example 10. To my knowledge, the grammatical categories of this kind of idiomatic expressions have not been documented so far.

- (9) <N>今晚_t 的_u 长安街_{ns}</N> <流光溢彩_l , _w 火树银花_i>
 <N>*jintwan*_t *de*_u *Changanjie*_{ns}</N> <*liuguangyicai*_l , _w *huoshuyinhua*_i>
 “Tonight the Changan Street was filled with colourful lights and really looked magnificent.”

⁷ 建_v好_a, *jian-hao* “build well, build in good shape” is a compound verb. More specifically, it is a verb-complement (V-R) compound (Chao 1968: 435ff). The resultative complement 好 *hao* “good” is bound to and follows the verb 建 *jian* “build” and expresses the result of the action of the verb.

- (10) <N>国民经济_n</N> <稳中求进_l>
 <N>guominjingji_n</N> <wenzhongqiujin_l>
 “The national economy is progressing steadily.”

That they can be used rather idiosyncratically as a noun or a verb makes it almost impossible for even a human analyst to determine the phrasal category of a given idiomatic expression: whether it is a noun phrase or a verb phrase. As in the above two examples, it is unclear whether the idiom/set phrase placed after the subject noun phrase is intended to function as a nominal expression or a verbal one. Unlike English, in which the subject must be followed by a verbal predicate, a Chinese predicate can be a verbal predicate, an adjectival predicate or a nominal predicate (Chao 1968: 90). In the absence of further evidence of the categorial status of such segments, those idioms and set phrases occurring in the predicate position were left unlabelled in my treebank.

5.2.3. Lengthy premodifiers of a noun phrase

Unlike English, which favours the use of postmodification if a modifier of a noun phrase is long (Quirk and Greenbaum 1973: 425; de Haan 1991), Chinese prefers premodification to postmodification, regardless of the length of the modifier (cf. Liu 1996: 265-274). It is thus common in the PFR treebank that a noun is qualified by a grammatical unit of over six words which is marked by the particle 的 *de* at the end, as in example 11. The particle 的 *de* is traditionally treated as a marker of modification (Chao 1968: 285).

- (11) <N><中国_ns 与_p 周边_n 国家_n 和_c 广大_b 发展中国家_l 的_u>
 友好_a 合作_vn</N>
 <N><Zhongguo_ns yu_p zhoubian_n guojia_n he_c guangda_b fazhan-
 zhongguojia_l de_u youhao_a hezuo_vn</N>
 “the co-operation between China and her surrounding countries and
 developing countries”

These lengthy premodifiers make the structure of the noun phrase in which they occur extremely difficult to interpret. Some premodifiers of this sort are complicated by the fact that they are further modified by another element marked by *de* in their internal structure, as in example 12.

- (12) <N>党_n 的_u <<基本_a 路线_n 提出_v 的_u> 党_n 在_p 社会主义_n
 初级_b 阶段_n 经济_n 、_w 政治_n 、_w 文化_n 的_u> 基本_a 纲领_n</N>
 <N>dang_n de_u <<jiben_a luxian_n tichu_v de_u> dang_n zai_p
 shehuizhuyi_n chuji_b jieduan_n jingji_n 、_w zhengzhi_n 、_w wenhua_n
 de_u> jiben_a gangling_n</N>
 “the primary principles of the Communist Party on economy, politics and
 culture, which are also on a par with the Party’s basic directions”

5.2.4. Serial verb constructions

Serial verb constructions in Chinese also increase the complexity of parsing. There is an immense literature on Chinese serial verb constructions (see, for

instance, Li and Thompson 1989: 594ff; Lin and Soo 1994; Liu 1996). Generally speaking, a serial verb construction refers to a succession of two or more actions that share the same subject, as illustrated in the following concocted example.

- (13) <N>我</N> <V>去 <N>朋友 家</N></V> <V>吃 <N>晚饭</N></V>
 <N>wo</N> <V>qu <N>pengyou jia</N></V> <V>chi <N>wanfan</N></V>
 “I went to my friend’s house to have dinner.”

However, some of the serial verb constructions in my treebank do not conform to this general pattern of two successive verbs, each of which has a different direct object. Unlike ordinary serial verbs, the serial verbs, as shown in examples 14 and 15, do not take a direct object separately. They are more like compound verbs than serial verbs, though it is not clear that they can be fully assimilated to the former category. Evidence in support of this analysis comes from the fact that these verbs (i.e. 指挥_v 演奏_v *zhìhuī yǎnzòu* “lead and perform” as in example 14, and 坚持_v 奉行_v *jiānchí fēngxíng* “insist and follow” as in example 15), functioning as if they were a single unit, take the same object, i.e. the following noun phrase.

- (14) <V><指挥_v 演奏_v> 了_u <N>一_m 批_q 中外_j 名曲_n</N></V>
 <V><zhìhuī_v yǎnzòu_v> le_u <N>yī_m pī_q zhōngwài_j míngqǔ_n</N></V>
 “led and performed a variety of Chinese and western popular songs”
 (15) <坚持_v 奉行_v> <N>独立自主_l的_u 和平_n 外交_n 政策_n</N>
 <jiānchí_v fēngxíng_v> <N>dulìzìzhǔ_l de_u héping_n wàijiāo_n zhèngcè_n</N>
 “insist on adopting an independent diplomatic policy in maintaining peace”

Besides sharing the same direct object, another clue that tends to prove that the two verbs are actually used as a compound verb is the suffixation of the morpheme 了 *-le*, as highlighted in example 14. The verbal *-le* has generally been taken as an aspect marker, indicating completion (Norman 1988: 163; Xiao 2002), and it is attached to verbs and not to the objects of verbs (Chao 1968: 247), excluding the possibility that the first verb takes the second verb (and the following noun phrase) as its object. Further research on clarifying their subcategorisation (whether they are serial or compound verbs) ought to be done in order to give a more precise parse.

6. Conclusion: Quality control of the skeleton parsing process

In evaluating the success of an annotation project, Eyes and Leech (1993: 37-42) provide six essential criteria that can be used for evaluating my skeleton parsing scheme.

1. Consensual categories: The linguistic categories that were employed in my parsing scheme have been demonstrated, by comparison to seven syntactic theories, to represent grammatical features largely agreed upon by linguists, rather than features which are theory-specific or deeply controversial.
2. Overall coverage: My sample treebank represents a reasonable length of text (comprising about 100,000 word tokens or 2,500 sentences) to be manually parsed and could be re-used in future research.

3. Productivity: Productivity was satisfactory with the simplified syntactic analysis provided by skeleton parsing.
4. Accuracy: The output of the parsed sentences was cross-checked by several posteditors with a background in linguistics. While one can never guarantee 100% accuracy, I believe the sample treebank to be highly accurate.
5. Uniformity of analysis: To demonstrate consistency of analysis, a concordance of the verb 要 *yao* “need” was drawn from my skeleton treebank. This verb always takes a verbal object, i.e. a verb functioning as the direct object of another verb, which is represented as *Vo* in my parsing scheme and is distinct from *V*, which stands for an independent verb phrase (see Table 2 for a description of the symbols *Vo* and *V*). There are 252 instances of the verb *yao* in my treebank. In each case, it is followed by a verbal object consistently marked as *Vo* not *V*, as highlighted in Figure 2.⁸

```

<V>注定_v <Vo>要_v <Vo>失败_v</Vo></Vo></V>
<V>要_v <I>进一步_d 团结_a 起来_v</I></V>
<V>要_v <Ja>更_d 好_a 地_u</Ja> <Vo>坚持_v <N><解放思想_i 、_w 实事求是_i 的_u 思想_n 路线_n</N></Vo></V>
<V>要_v <Ja>更_d 好_a 地_u</Ja> <Vo>坚持_v <P以_p <N>经济_n 建设_vn</N></P> <Vo>为_v <N>中心_n</N></Vo></Vo></V>
<V>要_v <Vo>看_v <Vo>能否_v <P>把_p <N>经济_n 工作_vn</N></P> 搞_v 上去_v</Vo></Vo></V>
<V>要_v <Ja>切实_ad</Ja> <P>把_p <N>精力_n</N></P> <Vo>集中_v <Vo>到_v <N><贯彻_v 落实_v 好_a 中央_n 关于_p 今年_t 经济_n 工作_vn 的_u <N><N&>总体_n 要求_n</N&> 和_c <N+>各项_r 重要_a 任务_n</N+> 上_f 来_v</N></N></Vo></Vo></V>
<V>要_v <Ja>更_d 好_a 地_u</Ja> <Vo>坚持_v <N> “<_w 两手抓_l 、_w 两手_m 都_d 要_v 硬_a ” _w 的_u 方针_n</N></Vo></V>
<V>要_v <Ja>更_d 好_a 地_u</Ja> <Vo>发扬_v <N>求真务实_l 、_w <密切_ad 联系_v 群众_n 的_u 作风_n</N></Vo></V>
<V>要_v <Vo>发扬_v <N>这样_r 的_u 好_a 作风_n</N></Vo></V>
<V>要_v <Vo>尊重_v <N>群众_n 的_u 意愿_n</N></Vo></V>
<V>要_v <R>大力_d</R> <Vo>倡导_v <N><说实话_l 、_w 办_v 实事_n 、_w 鼓_v 实劲_n 、_w 讲_v 实效_n 的_u 作风_n</N></Vo></V>
<V>要_v <Vo>坚守_v <N>岗位_n</N></Vo></V>
<V>要_v <Vo>让_v <S><N>他们_r</N> <Ja>健康_ad</Ja> <V>成长_v</V></S></Vo></V>

```

FIGURE 2: A concordance of the verb *yao*

⁸ The verb *yao* can take an adjectival direct object, as shown in one instance of the concordance extracted. A vast majority of adjectives may function as verbs in Mandarin Chinese by taking aspect markers (e.g. *-guo* “experiential aspect”, *-le* “perfective aspect”, etc.) or directional complements (e.g. *qilai* “inchoative”). See Li and Thompson (1989: 141-147).

6. Linguistic validity: One of the aims of carrying out a skeleton parsing on a sample text of the PFR Chinese Corpus is to gain a better understanding of how to precisely locate adverbial clauses in a piece of POS tagged text. Further research will be conducted into adverbial clauses in written Chinese.

References

- Abeillé, A. (ed.), 2003, *Treebanks: Building and Using Parsed Corpora*. (Text, Speech and Language Technology, vol. 20.) Dordrecht, Boston and London: Kluwer Academic Publishers.
- Aoun, J. and A. Y.-h. Li, 1993, *Syntax of Scope*. Cambridge: MIT Press.
- Black, E., S. Eubank, H. Kashioka, D. Magerman, R. Garside and G. Leech, 1996, "Beyond skeleton parsing: producing a comprehensive large-scale general-English treebank with full grammatical analysis". *Proceedings of the 16th International Conference on Computational Linguistics (COLING '96)*, 107-12. Copenhagen.
- Chao, Y., 1968, *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chen, H.-H., 1990, "A logic-based government-binding parser for Mandarin Chinese". *Proceedings of the 13th International Conference on Computational Linguistics (COLING '90)*, ed. H. Karlgren, 48-53. Helsinki, Finland.
- Chiu, H.-ch., 1993, *The inflectional structure of Mandarin Chinese*. Unpublished Ph.D. thesis, University of Los Angeles.
- de Haan, P., 1991, "On the exploration of corpus data by means of problem-oriented tagging: postmodifying clauses in the English noun phrase". *English Computer Corpora: Selected Papers and Research Guide*, ed. S. Johansson and A.-B. Stenström, 52-65. Berlin: Mouton de Gruyter.
- Eyes, E. and G. Leech, 1993, "Syntactic annotation: linguistic aspects of grammatical tagging and skeleton parsing". *Statistically-driven Computer Grammars of English: The IBM/Lancaster Approach*, ed. E. Black, R. Garside and G. Leech, 36-61. Amsterdam: Rodopi.
- Garside, R., 1993, "The large-scale production of syntactically analysed corpora", *Literary and Linguistic Computing* 8: 39-46.
- and A. McEnery, 1993, "Treebanking: the compilation of a corpus of skeleton-parsed sentences". *Statistically-driven Computer Grammars of English: The IBM/Lancaster Approach*, ed. E. Black, R. Garside and G. Leech, 17-35. Amsterdam: Rodopi.
- Kahrel, P., R. Barnett and G. Leech, 1997, "Towards cross-linguistic standards or guidelines for the annotation of corpora". *Corpus Annotation: Linguistics Information from Computer Text Corpora*, ed. R. Garside, G. Leech and A. McEnery, 231-42. London and New York: Longman.
- Kit, Ch., 1992, "Parsing Chinese *ba* and *bei* constructions - an LFG approach". *Proceedings of the 3rd International Conference on Chinese Information Processing*, 12-6. Beijing, China.
- Leech, G. and E. Eyes, 1997, "Syntactic annotation: treebanks". *Corpus Annotation: Linguistics Information from Computer Text Corpora*, ed. R. Garside, G. Leech and A. McEnery, 34-52. London and New York: Longman.

- Leech, G. and R. Garside, 1991, "Running a grammar factory: the production of syntactically analysed corpora or 'treebanks'". *English Computer Corpora: Selected Papers and Research Guide*, ed. S. Johansson and A.-B. Stenström, 15-32. Berlin: Mouton de Gruyter.
- , R. Barnett, and P. Kahrel, 1996, "Recommendations for the Syntactic Annotation of Corpora". EAGLES Document EAG-TCWG-SASG/1.8 Version of 11th March 1996. Available on-line at <http://www.ilc.pi.cnr.it/EAGLES96/segsasg1/html>.
- Li, A. Y.-h., 2001, "The *ba* construction". University of Southern California, ms. Available on-line at <http://www.usc.edu/dept/LAS/ealc/chinling/audreyli.htm>.
- Li, Ch. and S. Thompson, 1989, *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles: University of California Press.
- Li, J., 1997, "The *ba* construction in Mandarin Chinese: a serial verb analysis". *Clitics, Pronouns and Movement*, ed. J. Black and V. Motapanyane, 175-216. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Lin, K. H. C. and V.-w., Soo, 1994, "Hypothesis scoring over theta grids information in parsing Chinese sentences with serial verb constructions". *Proceedings of the 15th International Conference on Computational Linguistics (COLING '94)*, 942-948. Kyoto, Japan.
- Liu, Ch.-sh., 1999, *Anaphora in Mandarin Chinese and binding at the interface*. Unpublished Ph.D. thesis, University of California, Irvine.
- Liu, G., 1996, "On serial verbs in Chinese and their representation in HPSG". *Proceedings of International Chinese Computing Conference (ICCC-96)*, 129-35. Singapore.
- Liu, Y.-h., 1996, *Modern Chinese Grammar for Teachers of Chinese as a Second Language and Advanced Learners of Modern Chinese*. Taiwan: Shi Da Shu Yuan.
- Norman, J., 1988, *Chinese*. Cambridge and New York: Cambridge U.P.
- Quirk, R. and S. Greenbaum, 1973, *A University Grammar of English*. London: Longman.
- Sampson, G., 1995, *English for the Computer: The SUSANNE Corpus and Analytic Scheme*. Oxford: Clarendon Press.
- Wu, Z. X.-z., 2000, *Grammaticalisation and the development of functional categories in Chinese*. Unpublished Ph.D. thesis, University of Southern California.
- Xia, X.-y. and De-kai Wu, 1996, "Parsing Chinese with an almost-context-free grammar". Paper presented at Conference on Empirical Methods in Natural Language Processing (EMNLP-96), May 17-18. University of Pennsylvania, Philadelphia.
- Xiao, Z.-h., 2002, *A corpus-based study of aspect in Mandarin Chinese*. Unpublished Ph.D. thesis, Lancaster University.
- Xue, N. Fei Xia, Shi-zhe Huang and A. Kroch, 2000, *The Bracketing Guidelines for the Penn Chinese Treebank (3.0)*. Available on-line at <http://www.cis.upenn.edu/~chinese/>.
- Yu, S.-w., 1999, *Xiandai Hanyu Yuliaoku Jiagong: Ciyu Qiefen Yu Cixing Biaozhu Guifan Yu Shouce* [Corpus annotation of Modern Chinese Corpora: Guidelines and Handbook on Word Segmentation and Part-of-Speech Tagging]. Beijing: Institute of Computational Linguistics, Peking University. Available on-line at <http://www.icl.pku.edu.cn/research/corpus/copus-annotation.htm>.

- Zhao, Er-pin, 1989, *The categorical status of THAN and BI in English and Chinese comparative constructions*. Unpublished Ph.D. thesis. University of Texas, Austin.
- Zhu, De-xi, 1982, *Yufa Xiangyi* [Lecture Notes on Chinese Grammar]. Beijing: Commercial Press.
- , 2000, *Xiandai Hanyu Yufa Yanjiu* [A Grammatical Study of Modern Chinese]. Beijing: Commercial Press.
- Zou, Ke, 1993, “The syntax of the Chinese *ba* construction”, *Linguistics* 31: 715-736.

«Julio Urkixo» Euskal Filologi Mintegiaren Aldizkariaren Gehigarriak
Anejos del Anuario del Seminario de Filología Vasca «Julio de Urquijo»
Publications of «Julio de Urquijo» Seminar of Basque Philology

- I. *El Seminario "Julio de Urquijo". Antecedentes y constitución*, 1955. Agotado.
- II. JOSÉ MARÍA LACARRA, *Vasconia medieval. Historia y filología*, 1957. Agotado.
- III. MANUEL AGUD - LUIS MICHELENA, *Landuccio, Dictionarium Linguae Cantabricae (1562)*, 1958. Agotado.
- IV. LUIS MICHELENA, *Fonética histórica vasca*, 1990 (1961¹, 1977², 1985³). 18 €.
- V. NILS M. HOLMER, *El idioma vasco hablado. Un estudio de dialectología vasca*, 1991 (1964¹). 12 €.
- VI. LUIS VILLASANTE, *Pedro A. de Añibarro, Gramática vascongada*, 1970. 8 €.
- VII. CANDIDO IZAGUIRRE, *El vocabulario vasco de Aránzazu-Oñate y zonas colindantes. (Índice inverso de Gerardo Markuleta)*, 1994 (1970¹). 12 €.
- VIII. *Papers from the Basque Linguistics Seminar. University of Nevada. Summer 1972, 1974.* 10 €.
- IX. JULIEN VINSON, *Essai d'une bibliographie de la langue basque. Con las anotaciones del ejemplar de Julio de Urquijo*, 1984. 18 €.
- X. LUIS MICHELENA, *Sobre historia de la lengua vasca*, 1988. 2 vols. Agotado.
- XI. LUIS MICHELENA-IBON SARASOLA, *Textos arcaicos vascos. Contribución al estudio y edición de textos antiguos vascos*, 1989. 12 €.
- XII. HUGO SCHUCHARDT, *Introducción a las obras de Leizarraga. Sobre el modo de disponer la reimpresión, en particular sobre las erratas y variantes en el texto de Leizarraga*, 1989. 8 €.
- XIII. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, I. A-Ardui*, 1989, 1993. Agotado.
- XIV. JOSEBA A. LAKARRA (ed.), *Memoriae L. Mitxelena magistri sacrum*, 1991. 36 €.
- XV. RICARDO GÓMEZ - JOSEBA A. LAKARRA (arg.), *Euskalaritzaren historiaz I: XVI-XIX. mendak*, 1992. 18 €.
- XVI. BEÑAT OYHARÇABAL, *La pastorale souletine: édition critique de "Charlemagne"*, 1990. 18 €.
- XVII. RICARDO GÓMEZ - JOSEBA A. LAKARRA (arg.), *Euskalaritzaren historiaz II: XIX-XX. mendak*. Prestatzen.
- XVIII. JOSEBA A. LAKARRA, *Harrieten Gramatikako biztegiak (1741)*, 1994. 10 €.
- XIX. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, II. Ardun-Beuden*, 1990, 1993. Agotado.
- XX. LUIS MICHELENA, *Lenguas y protolenguas*, 1990 (1963, 1986). 8 €.

- XXI. ARENE GARAMENDI, *El teatro popular vasco. Semiótica de la representación*, 1991. 12 €.
- XXII. LASZLO K. MARACZ, *Asymmetries in Hungarian*, 1991. 15 €.
- XXIII. PETER BAKKER, GIDOR BILBAO, NICOLAAS G. H. DEEN, JOSÉ I. HUALDE, *Basque pidgins in Iceland and Canada*, 1991. 10 €.
- XXIV. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, III. Beule-Egileor (Babarraso-Bazar)*, 1991. Agotado.
- XXV. JOSÉ M.^a SÁNCHEZ CARRIÓN, *Un futuro para nuestro pasado*, 1991. 15 €.
- XXVI. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, IV. Egiluma-Galanga*, 1991. Agotado.
- XXVII. JOSEBA A. LAKARRA - JON ORTIZ de URBINA (eds.), *Syntactic theory and Basque syntax*, 1992. 18 €.
- XXVIII. RICARDO GÓMEZ - JOSEBA A. LAKARRA (arg.), *Euskal dialektologiako kongresua (Donostia, 1991ko irailaren 2-6)*, 1994. 21 €.
- XXIX. JOSÉ I. HUALDE - XABIER BILBAO, *A phonological study of the Basque dialect of Getxo*, 1992. 8 €.
- XXX. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, V. Galani-Iloza*, 1992. 8 €.
- XXXI. KARLOS OTEGI, *Lizardi: lectura semiótica de "Biotz-begietan"*, 1993. 18 €.
- XXXII. AURELIA ARKOTXA, *Imaginaire et poésie dans "Maldan behera" de Gabriel Aresti (1933-1975)*, 1993. 18 €.
- XXXIII. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, VI. Ilpiztu-Korotz*, 1993. 8 €.
- XXXIV. JOSÉ I. HUALDE - GORKA ELORDIETA - ARANTZAZU ELORDIETA, *The Basque dialect of Lekeitio*, 1994. 18 €.
- XXXV. GEORGES REBUSCHI, *Essais de linguistique basque*, 1997. 18 €.
- XXXVI. XABIER ARTIAGOITIA, *Verbal projections in Basque and minimal structure*, 1994. 12 €.
- XXXVII. MANUEL AGUD - ANTONIO TOVAR, *Diccionario etimológico vasco, VII. Korpa-Orloi*, 1994. 8 €.
- XXXVIII. PATXI GOENAGA (ed.), *De grammatica generativa*, 1995. 18 €.
- XXXIX. ANTONIO CID, *Romancero y balada oral vasca. (Literatura, historia, significado)*. En preparación.
- XL. AMAIA MENDIKOETXEA - MYRIAM URIBE-ETXEBARRIA (eds.), *Theoretical issues at the morphology-syntax interface*, 1997. 21 €.
- XLI. BERNARD HURCH - MARÍA JOSÉ KEREJETA, *Hugo Schuchardt - Julio de Urquijo: Correspondencia (1906-1927)*, 1997. 21 €.
- XLII. JOSÉ I. HUALDE, *Euskararen azentuerak*, 1997. 15 €.
- XLIII. RUDOLF P. G. de RIJK, *De lingua Vasconum: Selected Writings*, 1998. 15 €.
- XLIV. XABIER ARTIAGOITIA - PATXI GOENAGA - JOSEBA A. LAKARRA (arg./eds.), *Erramu Boneta: Festschrift Rudolf P. G. de Rijk*, 2002. 30 €.

- XLV. JOSEBA A. LAKARRA, *Ikerketak euskararen historiaz eta euskal filologiaz*. Argitaratzeko.
- XLVI. BEÑAT OYHARÇABAL, *Inquiries into the lexicon-syntax relations in Basque*, 2003. 18 €.
- XLVII. BLANCA URGELL, *Larramendiren "Hiztegi Hirukoitza"-ren Eranskina: saio bat hiztegi-gintzaren testukritikaz*. Argitaratzeko.
- XLVIII. ÍÑIGO RUIZ ARZALLUZ, *"Aitorkizunen" historia eta testua: Orixeren eskuizkributik Le-kuonaren ediziora*, 2003. 21 €.
- XLIX. GOTZON AURREKOETXEA - XARLES VIDEGAIN (arg.), *Haur prodigoaren parabola Ipar Euskal Herriko 150 bertsiotan*, 2004. 21 €.
- L. JOSEBA A. LAKARRA, *Raíz y reconstrucción del protovasco*. En prensa.

MONUMENTA LINGUAE VASCONUM
STUDIA ET INSTRUMENTA

- I. BLANCA URGELL, *Larramendiren "Hiztegi Hirukoitza"-ren Eranskina: saio bat hiztegi-gintzaren testukritikaz* (= Gehigarriak XLVII). Argitaratzeko.
- II. ÍÑIGO RUIZ ARZALLUZ, *"Aitorkizunen" historia eta testua: Orixeren eskuizkributik Le-kuonaren ediziora*, 2003, (= Gehigarriak XLVIII). 21 €.
- III. OROITZ JAUREGI, *Correspondencia de Gerhard Bähr con R. M. Azkue, H. Schuchardt y J. Urquijo (1920-1944)*, (= ASJU XXXVI-2), 21 €.
- IV. CÉLINE MOUNOLE HIRART-URRUTY, *C. H. de Belsunce Bizkondea Tableau analytique et grammatical de la langue basque (1858) azterketa eta edizioa* (= ASJU XXXVII-2). Argitaratzeko.

BIBLIOGRAFIA-LABURDURA GOMENDATUAK
ABREVIATURAS BIBLIOGRAFICAS RECOMENDADAS
RECOMMENDED BIBLIOGRAPHICAL ABBREVIATIONS

Hemen agertzen ez denerako, erabil bitez *Orotariko Euskal Hiztegia*-n agertzen direnak.

Para las obras no citadas abajo, se emplearán las abreviaturas del Diccionario General Vasco.

For any works which do not appear below, the abbreviations given in the *Diccionario General Vasco* should be used.

- AEF* = *Anuario de Eusko Folklore*, Vitoria-Gasteiz, 1921-1936; Donostia-San Sebastián, 1956-
- AION* = *Annali dell'Istituto Orientale di Napoli*, Napoli, 1979-
- ASJU* = *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*. *International Journal of Basque Linguistics and Philology*, Donostia-San Sebastián, 1954-1955, 1967-
- Azk* = Resurrección M.^a de Azkue, *Diccionario vasco-español-francés*, Bilbao, 1905-1906 [1969², 1984³].
- Azk Morf* = Id., *Morfología vasca (Gramática básica dialectal del euskera)*, Bilbao, 1923-1925 [1969²].
- BAP* = *Boletín de la Real Sociedad Vascongada de Amigos del País*, Donostia-San Sebastián, 1945-
- BGS* = *Beitrag zur Geschichte der Sprachwissenschaft*, Münster, 1991-
- BISS* = *Boletín de la Institución "Sancho el Sabio"*, Vitoria-Gasteiz, 1957-81. Vide *Sancho el Sabio*.
- BMB* = *Bulletin du Musée Basque*, Baiona, 1924-43, 1964-
- BRAE* = *Boletín de la Real Academia Española*, Madrid, 1914-
- BRAH* = *Boletín de la Real Academia de la Historia*, Madrid, 1877-
- BSL* = *Bulletin de la Société de Linguistique de Paris*, Paris, 1884-
- BLS* = (*Proceedings of the*) *Berkeley Linguistics Society*, Univ. of California, Berkeley, 1975-
- CAJ* = *Central Asiatic Journal*, Wiesbaden, 1955-
- Campión* = Arturo Campión, *Gramática de los cuatro dialectos literarios de la lengua éuskara*, Iruñea/Pamplona, 1884 [1977²].
- CEEN* = *Cuadernos de Etnografía y Etnología de Navarra*, Pamplona, 1969-
- CIL* = *Corpus Inscriptionum Latinarum*, Berlin, 1863-
- CLAO* = *Cahiers de Linguistique - Asie Orientale*, Paris, 1971-
- CLS* = (*Proceedings of the*) *Chicago Linguistics Society*, Univ. of Chicago, 1965-

- Contr* = Ibon Sarasola, "Contribución al estudio y edición de textos antiguos vascos", *ASJU* 17 (1983): 69-212. L. Michelena & I. Sarasola, *Textos arcaicos vascos. Contribución...*, Anejos de *ASJU* 11, Donostia-San Sebastián, 1989.
- DCECH* = Juan Corominas & José Antonio Pascual, *Diccionario crítico etimológico castellano e hispánico*. Madrid, Gredos, 1980-1991.
- DELL* = Alfred Ernout & Antoine Meillet, *Dictionnaire étymologique de la langue latine. Histoire des mots*, Paris, 1932 [1939², 1951³, 1959⁴].
- DGV* = vide *OEH*.
- Diachronica* = *Diachronica. International Journal for Historical Linguistics*, Amsterdam-Philadelphia, 1984-
- DRA* = Manuel de la Sota, Pierre Lafitte, Lino de Akesolo. et al., *Diccionario Retana de Autoridades de la Lengua Vasca*, Bilbao, 1976-1989.
- Euskera* = *Euskera. Euskaltzaindiaren lan eta agiriak*, Bilbao, 1920-1936, 1953-
- EAA* = *Estudios de Arqueología alavesa*, Vitoria-Gasteiz, 1966-
- EFDA* = Luis Michelena, *Estudio sobre las fuentes del diccionario de Azkue*, Bilbao, 1970 [= Azk 1984].
- EFOu* = *Études finno-ougriennes*, Paris, 1964-
- EH* = Ibon Sarasola, *Euskal hiztegia*, Donostia-San Sebastián, 1996.
- EI* = Ana M.^a Echaide (arg.), *Erizkizundi irukoitzea*, Bilbao, 1984.
- EJ* = *Eusko Jakintza*, Baiona, 1947-1957.
- ELH* = *Enciclopedia Lingüística Hispánica*, Madrid, 1959-
- FEW* = W. von Wartburg, *Französisches Etymologisches Wörterbuch*, Bonn, 1928-
- FHV* = Luis Michelena, *Fonética histórica vasca*, Anejos de *ASJU* 4, Donostia-San Sebastián, 1961, 1977² [1985, 1990].
- FLV* = *Fontes Linguae Vasconum. Studia et documenta*, Iruñea/Pamplona, 1969-
- FL* = *Folia Linguistica. Acta Societatis Linguisticae Europaeae*, Berlin-New York, 1967-
- FLH* = *Folia Linguistica Historica. Acta Societatis Linguisticae Europaeae*, Berlin-New York, 1980-
- GH* = *Gure Herria*, Baiona, 1921-
- HEL* = *Histoire, Epistémologie, Langage*, Paris, 1979-
- HL* = *Historiographia Linguistica: International Journal for the History of the Language Sciences*, John Benjamins, 1974-
- HLEH* = Ibon Sarasola, *Hauta-lanerako euskal hiztegia*, Donostia-San Sebastián, 1984-1995. Vide *EH*.
- HLV* = Luis Michelena, *Historia de la literatura vasca*, Madrid, 1960 [1988].
- HLV* = Luis Villasante, *Historia de la literatura vasca*, Bilbao, 1961, 1979².
- HomUrq* = *Homenaje a don Julio de Urquijo e Ybarra*, Donostia-San Sebastián, 1949-1951.
- HSLV* = Ibon Sarasola, *Historia social de la literatura vasca*, Madrid, 1976 [1982].

- IEW* = Julius Pokorny, *Indogermanisches Etymologisches Wörterbuch*, Berna, 1951-1969.
- IF* = *Indogermanische Forschungen*, Berlin, 1892-
- IJAL* = *International Journal of American Linguistics*, Chicago, 1917-
- IL* = *Indian Linguistics. Journal of the Society of India*, Pune (India), 1931-
- IMU* = *Italia medioevale e umanistica*, Padova, 1958-
- Incipit* = *Incipit. Seminario de edición y crítica textual*, Buenos Aires, 1981-
- JALL* = *Journal of African Languages and Linguistics*, Berlin-New York, 1979-
- JEAL* = *Journal of East Asian Linguistics*, Berlin, etc., 1992-
- JWAL* = *Journal of West African Languages*, Dallas, 1964-
- Lexicographica* = *Lexicographica. Internationales Jahrbuch für Lexikographie*, Tübingen, 1985-
- Lg* = *Language*, Baltimore, 1924-
- Lb* = P. Lhande, *Dictionnaire Basque-Français*, Paris, 1926.
- LH* = Luis Michelena, *Lengua e historia*, Madrid, Paraninfo, 1985.
- LI* = *Linguistic Inquiry*, Cambridge (Mass.), 1971-
- MDEV* = Manuel Agud & Antonio Tovar, *Materiales para un diccionario etimológico vasco*, Anejos de *ASJU* 13, Donostia-San Sebastián, 1989-
- Memoriae* = Joseba A. Lakarra (ed.), *Memoriae L. Mitxelena Magistri Sacrum*, Anejos de *ASJU* 14, Donostia-San Sebastián, 1991.
- NLLT* = *Natural Language and Linguistic Theory*, Dordrecht, 1983-
- NTS* = *Norks Tidsskrift for Sprogvidenskap*, Oslo, 1928-
- OEH* = Luis Michelena, *Diccionario General Vasco. Orotariko Euskal Hiztegia*, Donostia-San Sebastián, 1987-2005.
- OL* = *Oceanic Linguistics*, Univ. of Hawaii, 1962-
- Phonology* = *Phonology*, Cambridge, 1984-
- PT* = Luis Michelena, *Palabras y textos*, Bilbao, UPV/EHU, 1987.
- PV* = *Príncipe de Viana*, Pamplona, 1940-
- RDTP* = *Revista de Dialectología y Tradiciones Populares*, Madrid, 1944-
- REW* = W. Meyer-Lübke, *Romanisches etymologisches Wörterbuch*, Heidelberg, 1930³.
- RFE* = *Revista de Filología Española*, Madrid, 1914-
- RIEV* = *Revista Internacional de los Estudios Vascos*, Paris-San Sebastián, 1907-1936, 1983-
- RLPhC* = *Revue de Linguistique et Philologie Comparée*, Paris, 1867-1916.
- RPh* = *Romance Philology*, Berkeley (CA), 1947-
- SAL* = *Studies in African Linguistics*, Bloomington (Indiana) / Columbus (Ohio), 1970-
- Sancho el Sabio* = *Sancho el Sabio: revista de cultura e investigación vasca = euskal kultura eta ikerketa aldizkaria*, Vitoria-Gasteiz, 1991-

- SHLV = Luis Michelena, *Sobre historia de la lengua vasca*, Anejos de ASJU 10, J. A. Lakarra (ed.), Donostia-San Sebastián, 1988.
- Symbolae = José Luis Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*, Vitoria-Gasteiz, Instituto de Ciencias de la Antigüedad-Antzinate-Zientzien Instituta, 1985.
- Syntax = *Syntax, A Journal of Theoretical, Experimental and Interdisciplinary Research*, Oxford, etc., 1998-
- TAV = Luis Michelena, *Textos arcaicos vascos*, Madrid, Minotauro, 1964 [= Luis Michelena - Ibon Sarasola, *Textos arcaicos vascos. Contribución*, Anejos de ASJU 11, Donostia-San Sebastián, 1989].
- TPhS = *Transactions of the Philological Society*, London, 1842-
- UAJ = *Ural-Altische Jahrbücher*, Wiesbaden, 1981-
- Vinson = Julien Vinson, *Essai d'une bibliographie de la langue basque*, Paris, 1891-1898 [vide Vinson-Urquijo].
- Vinson-Urquijo = Julien Vinson, *Essai... con las anotaciones del ejemplar de Julio de Urquijo*, Anejos de ASJU 9, Donostia-San Sebastián, 1984.
- ZRPb = *Zeitschrift für romanische Philologie*, Halle, 1877-

EGILEENTZAKO OHARRAK

ASJU-n euskaraz edo nazioarteko zientzi elkarteetan ohiko diren hizkuntzetako batean idatzirik euskal linguistika eta filologiazko lanak argitaratzen dira, baita eremu ezberdin edo zabalago bati atxikiak izan arren, euskalaritzarako interesgarri izan daitezkeenak ere. Originalak helbide honetara bidali behar dira: Joseba A. Lakarra, Hizkuntzalaritza eta Euskal Ikasketak Saila, Filologia eta Geografi-Historia Fakultatea, Unibertsitateko ibilbidea 5, 01006 Gasteiz (joseba.lakarra@ehu.es).

ASJU-ra igorritako artikulua gutxienez bi aztertzailek irakurriko dituzte, haien iruzkinak kontuan izanik atera edo ez erabakitzeko; erabakia ahalik eta lasterrenik gaztigituko zaie egileei. Artikulua onartzekotan, oztopo, akats edo aldabeharren zerrenda ere emango zaie. Egileek lehendabiziko inprenta probak jasoko dituzte (eta originalarekin batera itzuli beharko dituzte); eskuratzen dituztenetik astebeteko epea izango dute zuzentzeko. Argitaratzailearen baimenik gabe ezingo dute garrantzizko aldaketa, gehiketa edo kenketarik egin. Egileei *ASJU*-ko zenbakiaren ale bana eta lanaren 25 separata emango zaizkie (10, liburu iruzkinak bada); gehiago nahi izanez gero, kostu prezioan agin ditzakete aurretiaz.

Ez da inongo murrizketarik originalen luzeraz, baina ez lukete izan behar berez behar baino gehiago; lanek zehatzak eta argiak beharko dute izan. Berariazko abegia egingo zaie ohar laburrei, batez ere dagoeneko argitaratu beste lanen bat kritikatzeko edo garatzen dutenean.

Originalen hasieran egilearen/egileen helbidea, telefonoa eta helbide elektronikoa ezarriko dira; biko espazioan, orrialde bakarrean, eta zein-nahi argitasun edo zuzenketarako albo guztietan zuriune zabalekin idatzirik aurkeztuko dira lanak. Orrialdeak eta oin-oharrak segidako zerrendan zenbatuko dira. Lana euskarri elektronikoa (programa erabilienetako batean) eta paperean (3 kopia) bidaliko da. Horrekin batera 10-20 lerroko laburpena ere erantsiko da. Aurkeztu baino lehen zuzen bedi ahalik eta hobekienik originala, inprenta hutsak gutxitzeko; orobat, argazki, lauki, mapa, grafiko, taula, irudi, etab. eman ez gero, izan bidez kalitatekerik handienekoak gardentasunik gal ez dezaten. Hauek guztiak zenbatuko dira eta ezagutzeko oin-perpau laburra erantsiko zaie, testuan ere non jarri behar diren argiro markaturik. Adibideak zenbatu egingo dira: (1), (2)a, (2)b, etab.; testuan aipatzerakoan egin bedi era honetan: (2a), (2b), (2a, b), (4d-h), etab. Inprentan ohiko ez den zein-nahi zeinu, letra edo diakritikoren azalpen argia ezarriko da lehendabiziko agerraldiaren testu aldameneko zuriunean.

Testua honako arauok beteaz aurkeztuko da: Aipu luzeak ahalik berezian joango dira, sangratuta, hasiera eta amaiera kakotzik gabe, letra borobilean; aipu laburrak ere borobilean, testuan bertan eta kakotz bikoitzen artean (“ ” edo « »). Kakotz bakunak (‘ ’) adierak edo hitz solteen itzulpenak emateko baliatuko dira. Metalinguistikoki erabilitako edota artikulua idazteko erabili den hizkuntzaz beste bateko hitzak letra etzanean ezarriko dira.

Liburu eta aldizkariaren izenei letra etzana dagokie, eta kakotzak artikuluei. Aldizkariaren zenbaki, urte eta orrialdeak eta liburuen argitaletxe eta edizio (ez inprimatze) tokia emango dira. Hala dagokionean, berrinprimatzea, berrargitalpena edo itzulpena den zehaztuko da. Aipuetarako erabil bedi urte-egile sistema, ahal den neurrian, eta urte bereko egile baten lan bat baino gehiago aipatu bada, a, b... hurrenkeran bereiziko dira: adib. (Vinson 1897a: 35-38), (ikus Lacombe 1924, Azkue 1923-25, Unlenbeck 1947b). Amaierako bibliografiarik ez bada, eman bidez bibliografia zehaztasunak oro soilik lehen agerraldian, eta ondokoetan egilearen deitura eta lanaren izenburu laburtua bakarrik, *op. cit.* eta *ibidem* direlakoak saihestuaz: adib. Guerra, *Cantares*, 22-24. Bibliografia ere biko espazioan idatziko da, eta honako formatu honi atxikiko zaio:

Mitxelena, K., 1950b, “La aspiración intervocálica”, *BAP* 6, 443-449. Berrarg. bere *Sobre historia de la lengua vasca*, *ASJU*-ren Gehigarriak 10, Donostia 1988, I, 191-202.

—, 1981a, “Lengua común y dialectos vascos”, *ASJU* 15, 291-313. Berrarg. bere *Palabras y Textos*, UPV/EHU, Vitoria-Gasteiz 1987, 35-55.

Ortiz de Urbina, J., 1989, *Some parameters in the grammar of Basque*, Foris, Dordrecht.

Rijk, R. P. de, 1985, “Un verbe méconnu”, in J. L. Melena (arg.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*, UPV/EHU, Vitoria-Gasteiz, II, 921-935.

Sarasola, I., 1986, “Larramendiren eraginaz eta”, *ASJU* 20: 1, 203-216.

Bibliografi laburduretarako erabil bedi ale honetan bertan erantsi den laburdura gomendatuen zerrenda. Beharrezkoa balitz, egileak besterik ere erabili ahalko luke, beti ere esangura lehendabiziko agerraldian azalduz.

INFORMATION FOR AUTHORS

Papers on Basque linguistics and philology, and more general fields related or of interest to Basque studies are accepted, provided they are written in the languages most used by the international scientific community. Submissions should be sent to: Joseba A. Lakarra, Department of Linguistics and Basque Studies, Faculty of Philology and Geography and History, Unibertsitate Etorbidea/Paseo de la Universidad 5, 01006 Vitoria-Gasteiz (joseba.lakarra@ehu.es).

Papers received by *ASJU* are submitted to at least two reviewers; the decision on publication is communicated to the author(s) within as short a time as possible. Should a paper be accepted, a list of objections or changes deemed necessary will be sent to the author(s). When the authors receive the first proofs of their work, these should be returned to the editor together with the original within one week. No changes, additions or deletions may be made without the permission of the editor. Authors receive a copy of the *ASJU* volume in which their article appears and 25 offprints of their papers (10 in the case of reviews). Further additional offprints may be ordered at cost price.

There is no restriction on the maximum length of submissions, but they should be no longer than is necessary; authors must be concise and clear. Preference will be given to short notes, especially when they offer critique or elaborate on previously published papers.

The originals, which should include the address, telephone number(s) and e-mail of the author(s) in the first lines, must be typed and double-spaced throughout on single-sided sheets; this also applies to the notes. Wide margins for possible corrections or clarifications are required. Pages are to be numbered serially, as are the notes. Manuscripts must be submitted in digital format (in one of the commonly used program formats) together with three printed copies, and must include an abstract of 10-20 lines. It is recommended that the paper be carefully corrected before presentation to avoid possible errors, and that photographs, pictures, maps, graphs, tables, figures, etc. be of the best possible quality to avoid loss of detail in reproduction. These graphics should all be numbered and have a short footnote or key for identification; their approximate position in the text should also be indicated. Examples should be labelled consecutively by numbers enclosed in brackets: (1), (2a), (2b), etc.; these bracketed numbers should be used in the body of the text when referring to examples, like so: (2a), (2b), (2a, b), (4d-h), etc. A clear description of any unusual symbols, characters or diacritics should be provided in the margin on their first occurrence.

The text must be formatted as follows: long quotations must be indented, without inverted commas at the beginning and end of the text, in plain type; short quotations, also in plain type, must be enclosed by double inverted commas (“ ” or « »). Single quotation marks (‘ ’) are to be used to denote translations of isolated terms. Terms used metalinguistically or in a language different to that of the text should be in italics.

The titles of books and journals should be in italics and those of papers between inverted commas. The issue, year and page numbers of journals should be given, and for books, the publisher's name and place of edition; where relevant, state whether the quotation is from a reprint, reedition or translation. Where possible use the author-year system for quotation, e.g. (Lafitte 1976a: 35-38), (see Schuchardt 1900, Azkue 1923-25, 1935). Otherwise, the complete bibliographical information should be given only on the first occurrence, limiting any subsequent references to the surname of the author and the abbreviated title (avoiding notations such as *op cit.* and *ibidem*), e.g. Altuna, *Versificación*, pp. 43-57. The bibliography must also be double-spaced, with the following format:

- Mitxelena, K., 1950b, “La aspiración intervocálica”, *BAP* 6, 443-449. Reed in *Sobre historia de la lengua vasca*, Supplements of *ASJU* 10, Donostia 1988, I, 191-202.
- , 1981a, “Lengua común y dialectos vascos”, *ASJU* 15, 291-313. Reed in *Palabras y Textos*, UPV/EHU, Vitoria-Gasteiz 1987, 35-55.
- Ortiz de Urbina, J., 1989, *Some parameters in the grammar of Basque*, Foris, Dordrecht.
- Rijk, R. P. de, 1985, “Un verbe méconnu”, in J. L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*, UPV/EHU, Vitoria-Gasteiz, 921-935.
- Sarasola, I., 1986, “Larramendiren eraginaz eta”, *ASJU* 20: 1, 203-216.

For abbreviations of secondary sources the “Abbreviation Index” published in this issue should be used. If necessary, other abbreviations may be used, and these should be made explicit on their first appearance in the text.

INFORMACIÓN PARA LOS AUTORES

ASJU publica artículos sobre lingüística y filología vasca y campos más generales relacionados con o de interés para la vascofilia, escritos en euskera o en cualquiera de los idiomas utilizados por la comunidad científica internacional. Los originales se enviarán a: Joseba A. Lakarra, Depto. de Lingüística y Estudios Vascos, Facultad de Filología y Geografía e Historia, Paseo de la Universidad 5, 01006 Vitoria-Gasteiz (joseba.lakarra@ehu.es).

Los artículos recibidos en *ASJU* son examinados al menos por dos revisores, cuyos informes condicionan su aceptación; la decisión será comunicada a los autores lo antes posible. En caso de aceptarse el artículo, se enviará al autor una lista de objeciones o cambios necesarios. Los autores recibirán las primeras pruebas de su trabajo (que deberán devolver junto con el original), y dispondrán para corregirlas de un plazo no superior a una semana desde su recepción. No podrán hacerse cambios, adiciones o supresiones importantes sin permiso del editor. Se entregará a cada autor un ejemplar del *ASJU* y 25 separatas de su artículo (10 en el caso de las reseñas), si bien es posible encargar con antelación otras adicionales, que le serán facturadas a precio de coste.

No existe restricción alguna sobre la longitud máxima de los originales, pero no deben ser más largos de lo necesario; los autores han de ser concisos y claros. Se dará preferencia a notas breves, particularmente cuando critiquen o desarrollen artículos publicados con anterioridad.

Los originales, al comienzo de los cuales se hará constar la dirección, el teléfono y la dirección electrónica del(os) autor(es), se presentarán escritos a doble espacio y por una sola cara, con márgenes amplios para posibles correcciones y aclaraciones. Las páginas irán numeradas correlativamente, así como las notas. Los manuscritos se enviarán en soporte electrónico (en alguno de los programas más usados) y en papel por triplicado, e irán acompañados de un resumen de entre 10 a 20 líneas. Se recomienda que el original sea corregido antes de su presentación para minimizar las erratas, y que fotografías, cuadros, mapas, gráficos, tablas, figuras, etc., sean de la mejor calidad posible para evitar pérdidas de detalle en la reproducción; todos ellos irán numerados y llevarán un breve pie o leyenda para su identificación; se indicará asimismo el lugar aproximado de colocación en el texto. Los ejemplos irán numerados: (1), (2a), (2b), etc.; al referirse a ellos en el texto se usará el formato (2a), (2b), (2a, b), (4d-h), etc. Se aclarará al margen en su primera aparición en el texto cualquier símbolo, carácter o marca diacrítica inusual.

El texto se presentará de acuerdo con las siguientes normas:

1) Las citas largas irán en texto sangrado, sin comillas al principio y final, en letra redonda; las citas breves, también en redonda, irán entre comillas dobles (“ ” o « »). Se usarán los ápices (‘ ’) para denotar acepciones o traducciones de términos aislados. Los términos utilizados metalingüísticamente o en una lengua distinta a aquella en la cual está redactado el texto irán en cursiva.

2) Los títulos de libros y revistas irán en cursiva, y los de los artículos entre comillas. Se indicará el n.º, año y páginas correspondientes de las publicaciones periódicas, y editorial y lugar de edición de los libros; en su caso se indicará si se cita de una reimpresión, reedición o traducción. Úsese en la medida de lo posible el sistema autor-año para las citas, p. ej. (Lafitte 1967a: 35-38), (véase Schuchardt 1900, Azkue 1923-25, 1935). En su defecto, se darán datos bibliográficos completos sólo en la primera ocasión, limitándose en las siguientes a señalar apellido del autor y título abreviado, evitando *op. cit.* e *ibidem*: Altuna, *Versificación*, pp. 43-57. La bibliografía irá también a doble espacio, ajustándose al siguiente formato:

Mitxelena, K., 1950b, “La aspiración intervocálica”, *BAP* 6, 443-449. Reproducido en su *Sobre historia de la lengua vasca*, Anejos del *ASJU* 10, Donostia 1988, I, 191-202.

—, 1981a, “Lengua común y dialectos vascos”, *ASJU* 15, 291-313. Reproducido en su *Palabras y Textos*, UPV/EHU, Vitoria-Gasteiz 1987, pp. 35-55.

Ortiz de Urbina, J., 1989, *Some parameters in the grammar of Basque*, Foris, Dordrecht.

Rijk, R. P. de, 1985, “Un verbe méconnu”, in J. L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*, UPV/EHU, Vitoria-Gasteiz, II, 921-935.

Sarasola, I., 1986, “Larramendiren eraginaz eta”, *ASJU* 20: 1, 203-216.

3) Para las abreviaturas de fuentes primarias o secundarias se recurrirá al índice de abreviaturas recomendadas publicado en este mismo número. En caso necesario el autor podrá utilizar otras, cuyo valor explicará en la primera aparición.

